donor_bias

Sean Maden

2022-12-21

This vignette walks through a series of deconvolution analyses evaluating the impact of between-donor variations on results.

Experiment series 1: ndonors = 2

Let's run a series of experiments keeping `ndonor = 2`. We start with the simplest example of `k = 2` types and `G = 2` total markers, one for each type. We will vary the offset from small (2) to medium (5) to large (10). This offset magnitude should be proportional to observed between-donor variances.

```
ndonor = 2
gindexv = c(1, 2)
offsetv <- c(2, 5, 10)
num.sim <- 50
lexpt <- lapply(offsetv, function(offi){
  title.append <- paste0("Offset = ", offi, "\n")
  donor_marker_experiment(gindexv = c(1,2), num.sim = num.sim, ndonor = ndonor,
                          plot.title.append = title.append,
                          sd.offset.pos = offi,
                          sd.offset.neg = offi)
})
```

```
## Loading required package: dplyr
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:Biobase':
##
##     combine
```

```
## The following objects are masked from 'package:GenomicRanges':
##
##     intersect, setdiff, union
```

```
## The following object is masked from 'package:GenomeInfoDb':
##
##     intersect
```

```
## The following objects are masked from 'package:IRanges':
##
##     collapse, desc, intersect, setdiff, slice, union

## The following objects are masked from 'package:S4Vectors':
##
##     first, intersect, rename, setdiff, setequal, union

## The following objects are masked from 'package:BiocGenerics':
##
##     combine, intersect, setdiff, union

## The following object is masked from 'package:matrixStats':
##
##     count

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
names(lexpt) <- paste0("offset:", offsetv)
```

Randomized marker data

```
knitr::kable(lexpt$`offset:2`$marker.table)
```

donor1	donor2	donor.combn.all.mean	donor.combn.all.median	type	marker	marker.type
28	24	26.0	26.0	type1	marker1	type1
4	4	4.0	4.0	type1	marker2	type2
3	3	3.0	3.0	type2	marker1	type1
23	20	21.5	21.5	type2	marker2	type2

```
knitr::kable(lexpt$`offset:5`$marker.table)
```

donor1	donor2	donor.combn.all.mean	donor.combn.all.median	type	marker	marker.type
32	23	27.5	27.5	type1	marker1	type1
8	7	7.5	7.5	type1	marker2	type2
6	6	6.0	6.0	type2	marker1	type1
27	19	23.0	23.0	type2	marker2	type2

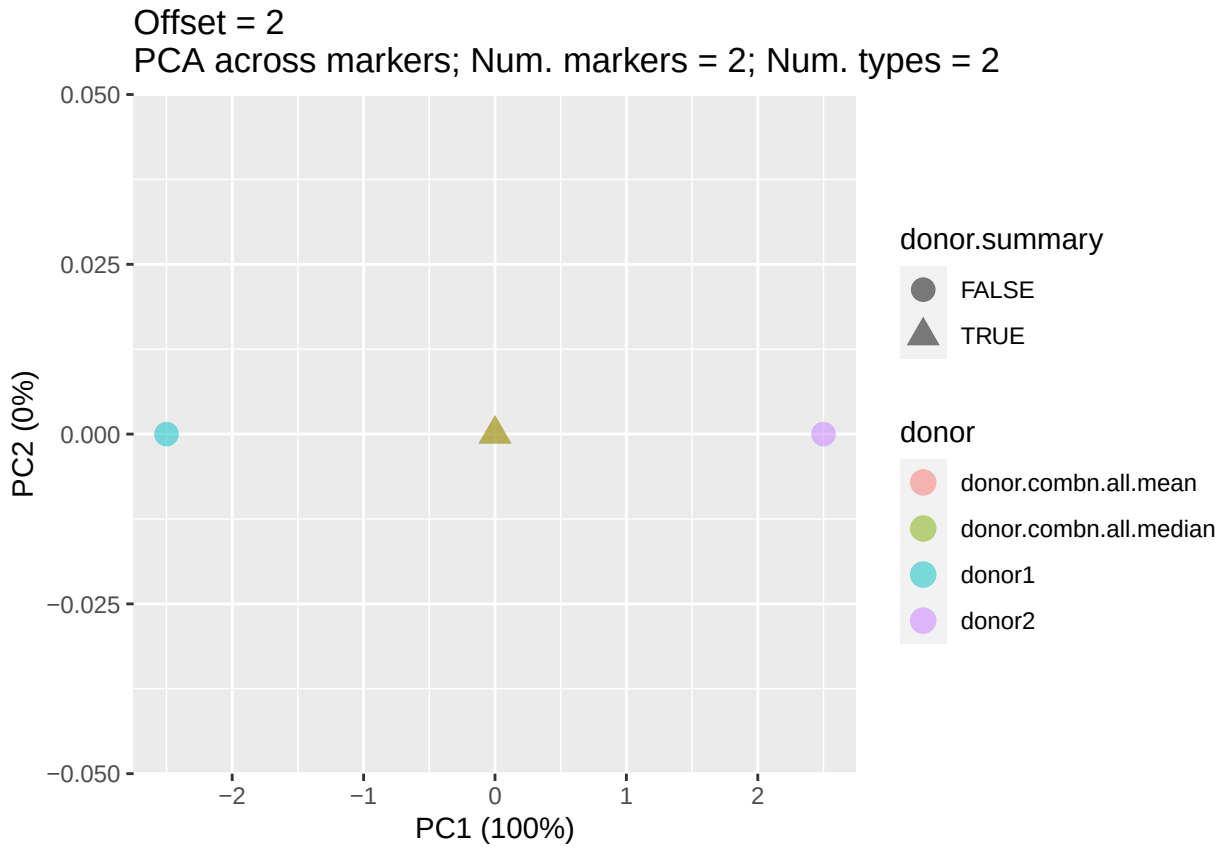
```
knitr::kable(lexpt$`offset:10`$marker.table)
```

donor1	donor2	donor.combn.all.mean	donor.combn.all.median	type	marker	marker.type
39	21	30.0	30.0	type1	marker1	type1
14	13	13.5	13.5	type1	marker2	type2
20	19	19.5	19.5	type2	marker1	type1
40	22	31.0	31.0	type2	marker2	type2

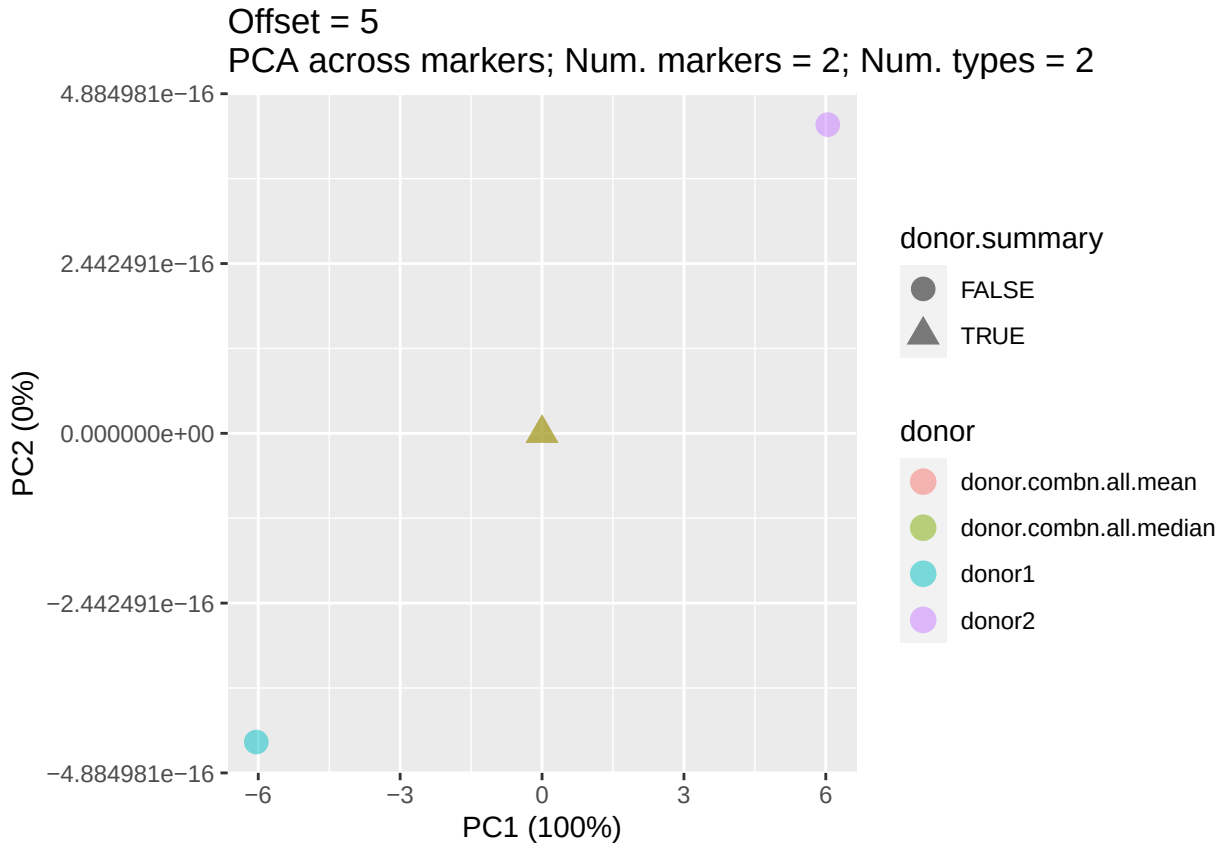
Marker bias plots

PC1 vs. PC2 of donor bias

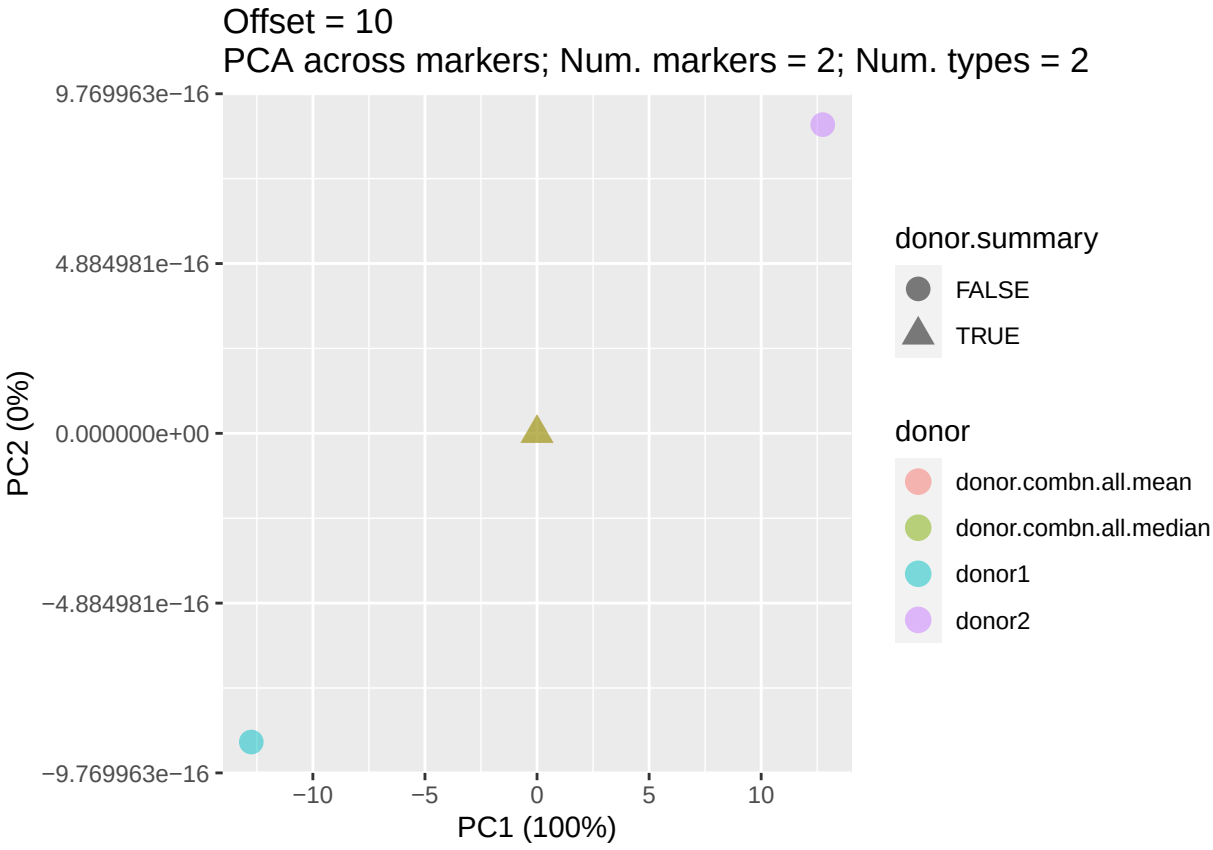
```
lexpt$`offset:2`$lpca.markers$pca.bydonor$scatterplot.pc1.pc2
```




```
lexpt$`offset:5`$lpca.markers$pca.bydonor$scatterplot.pc1.pc2
```

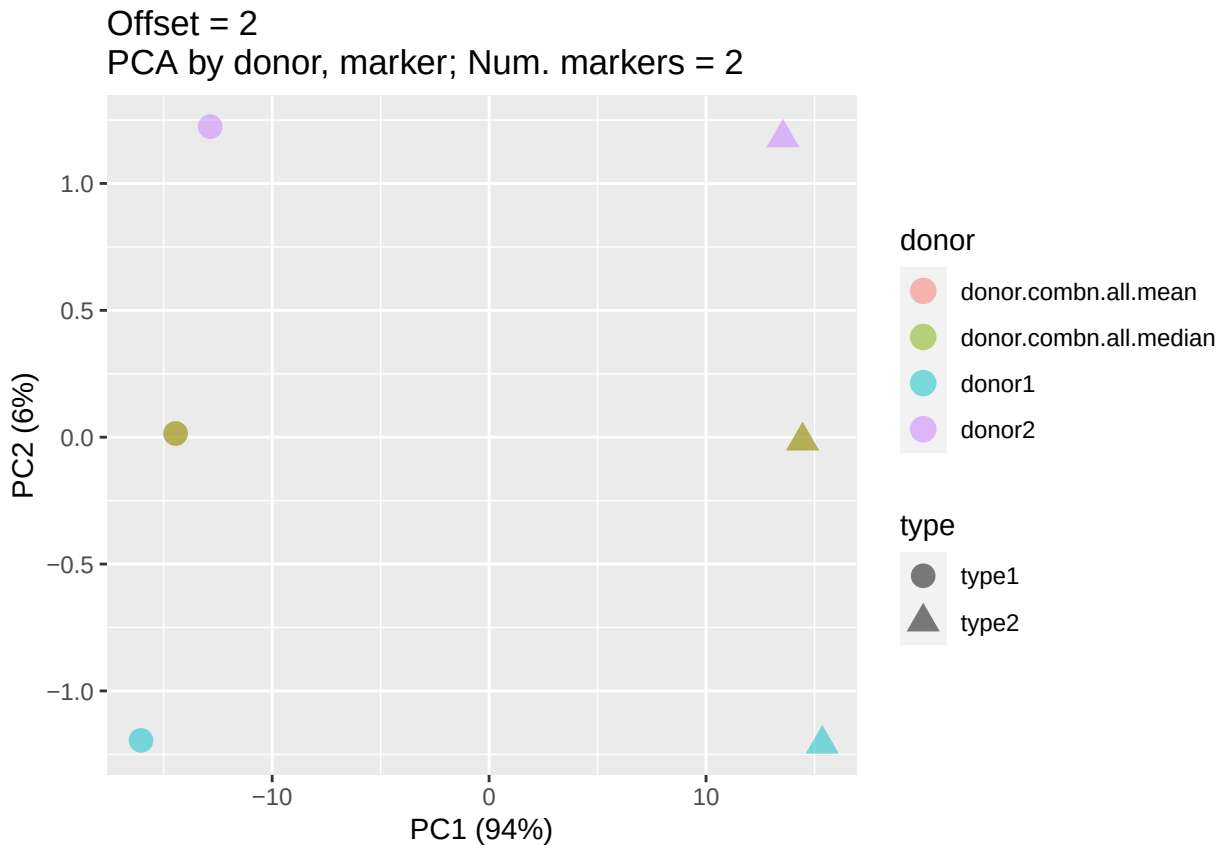



```
lexpt$`offset:10`$lpca.markers$pca.bydonor$scatterplot.pc1.pc2
```

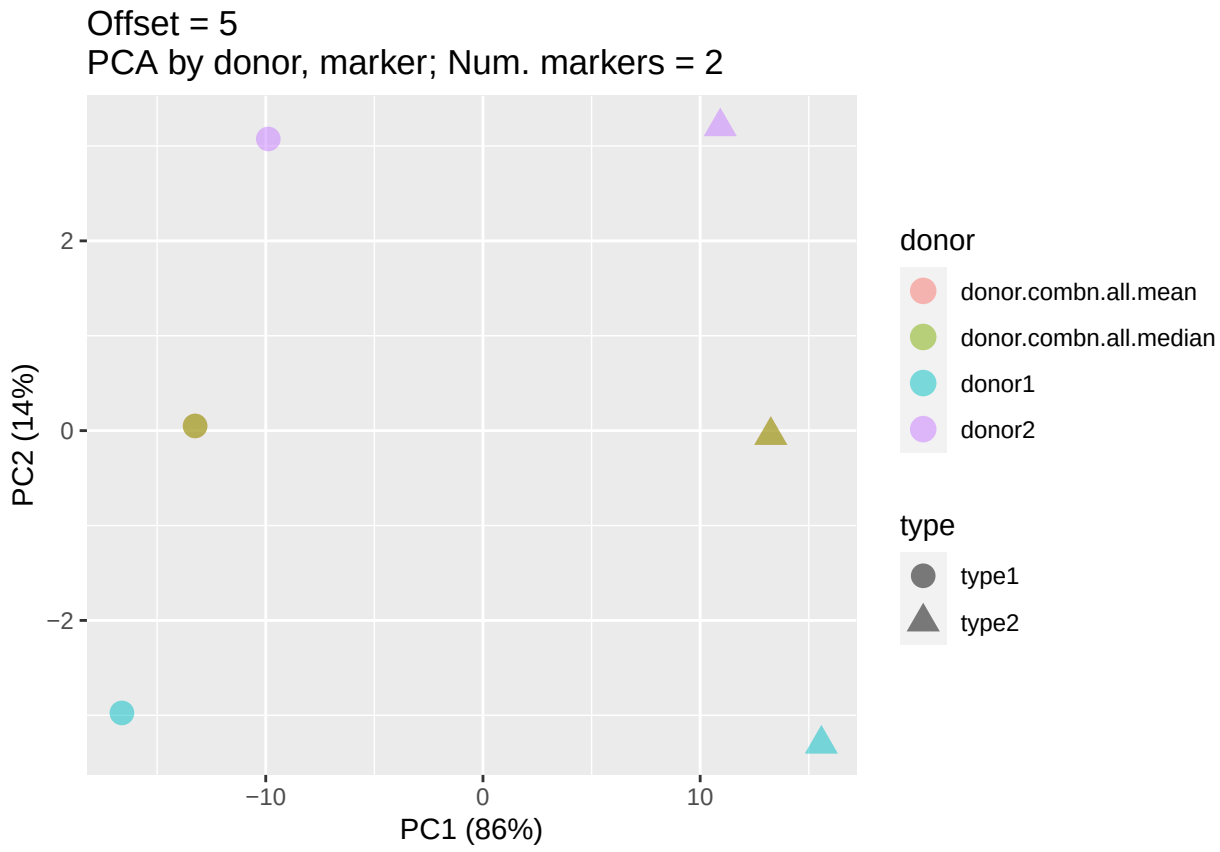


PC1 vs. PC2 of donor, marker variances

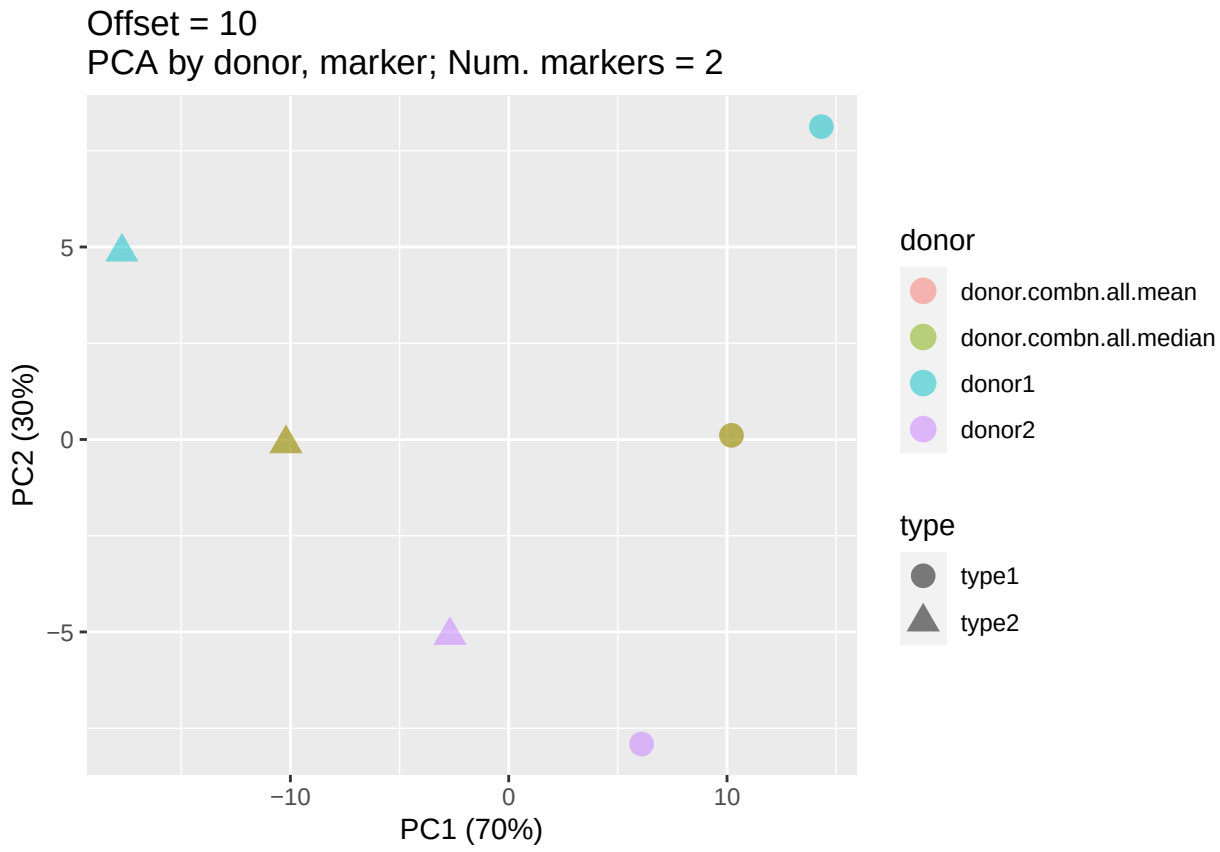
```
lexpt$`offset:2`$lpca.markers$pca.bydonortype$scatterplot.pc1.pc2
```




```
lexpt$`offset:5`$lpca.markers$pca.bydonortype$scatterplot.pc1.pc2
```

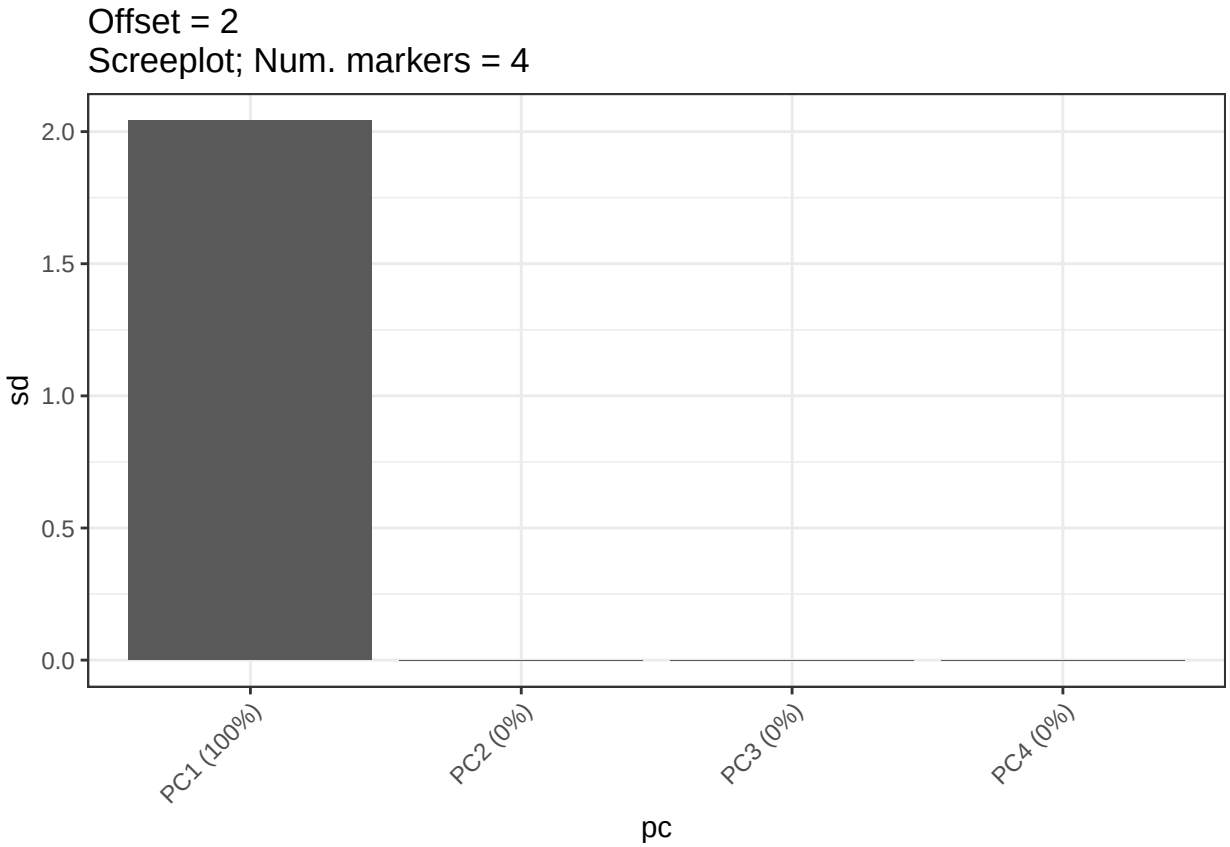



```
lexpt$`offset:10`$lpca.markers$pca.bydonortype$scatterplot.pc1.pc2
```

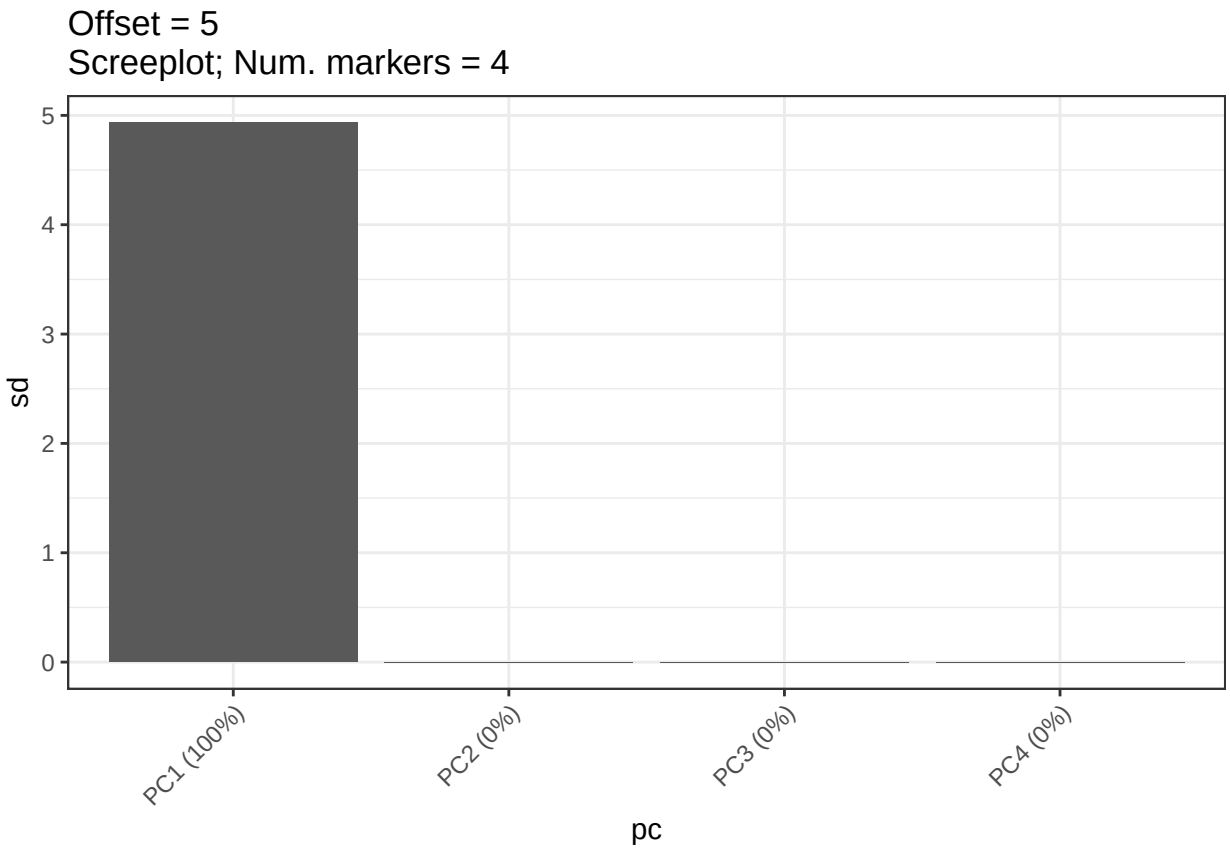


Screeplots, PCA by donor

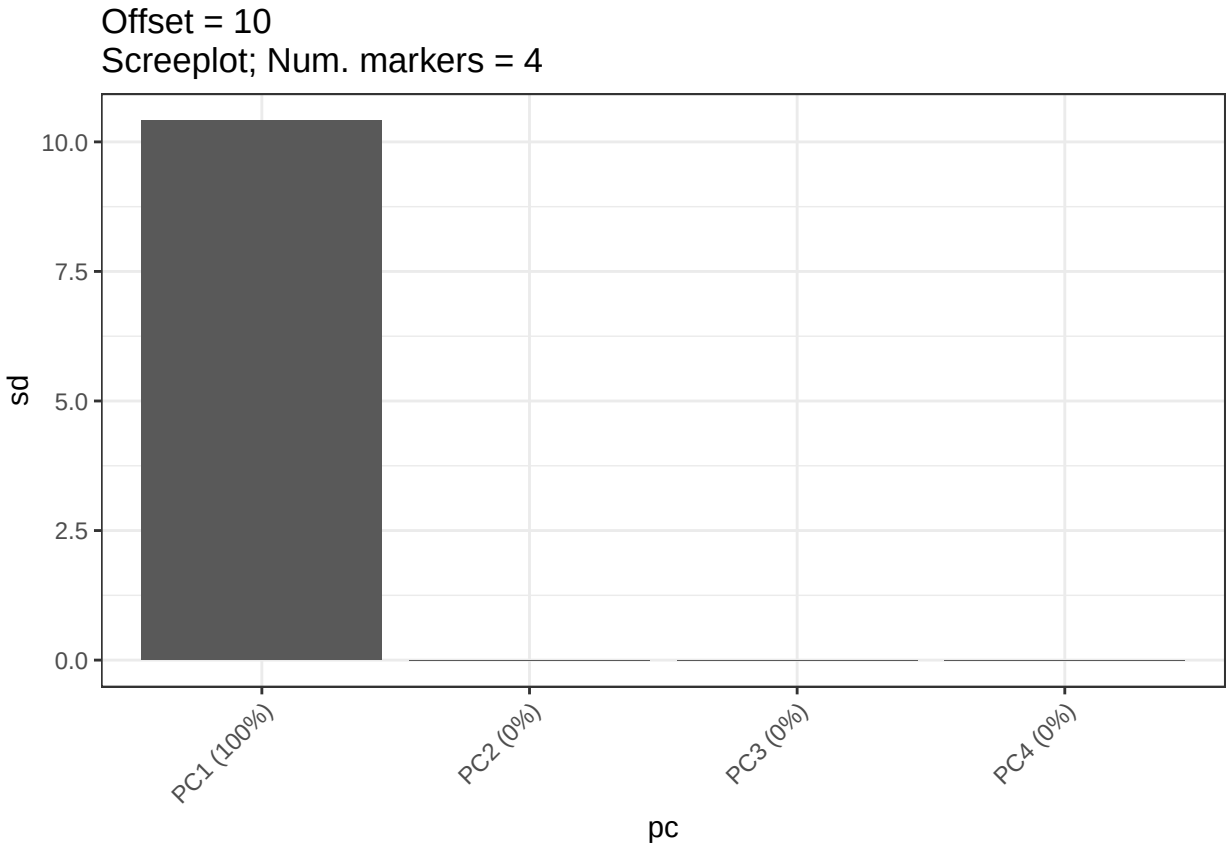
```
lexpt$`offset:2`$lpca.markers$pca.bydonor$screeplot
```




```
lexpt$`offset:5`$lpca.markers$pca.bydonor$screeplot
```

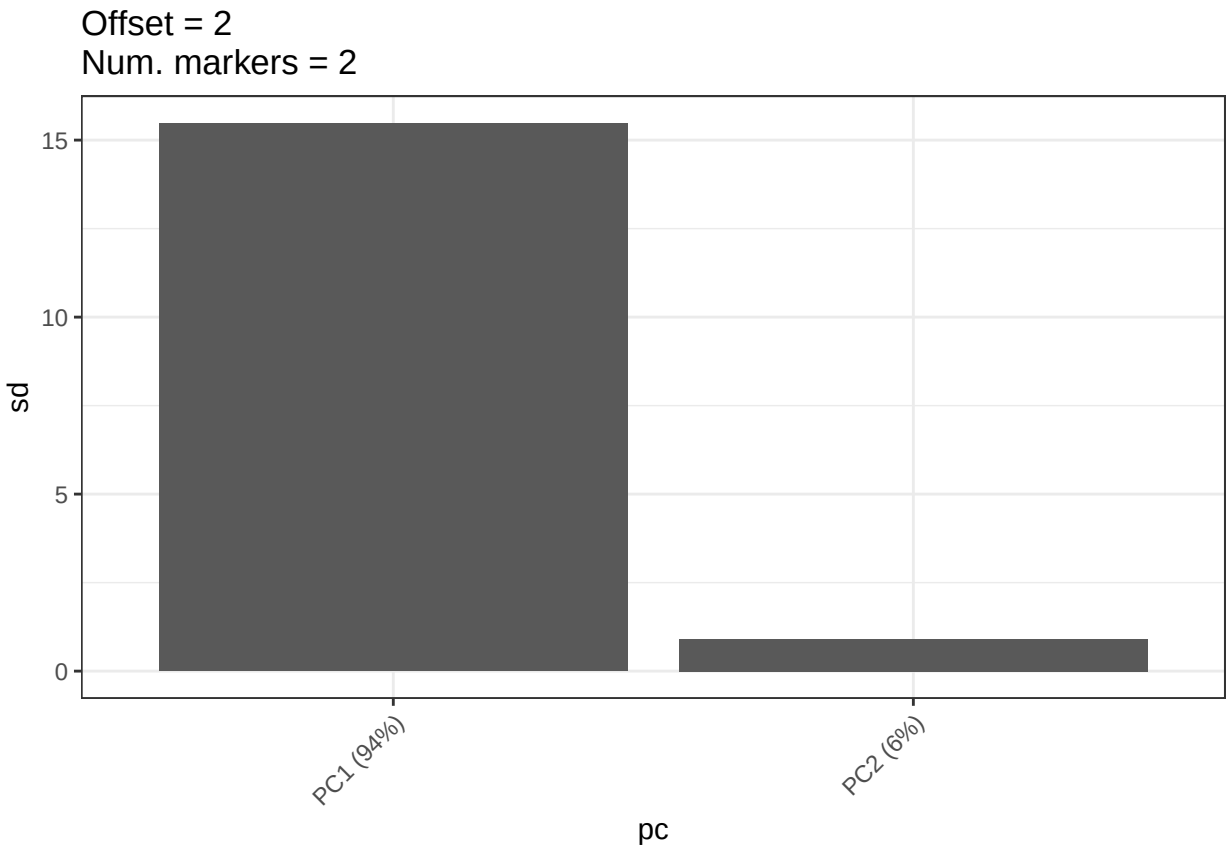



```
lexpt$`offset:10`$lpca.markers$pca.bydonor$screeplot
```

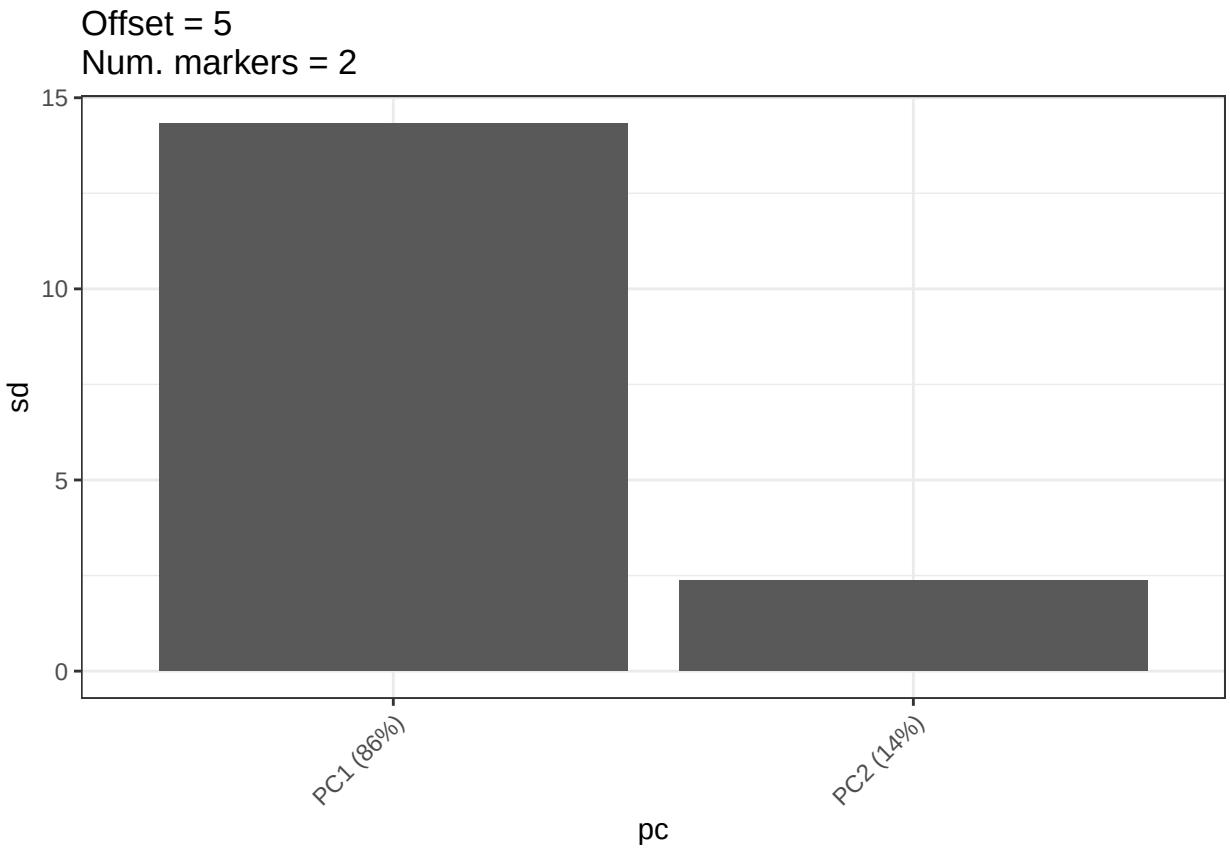


Screeplots, PCA by donor;type

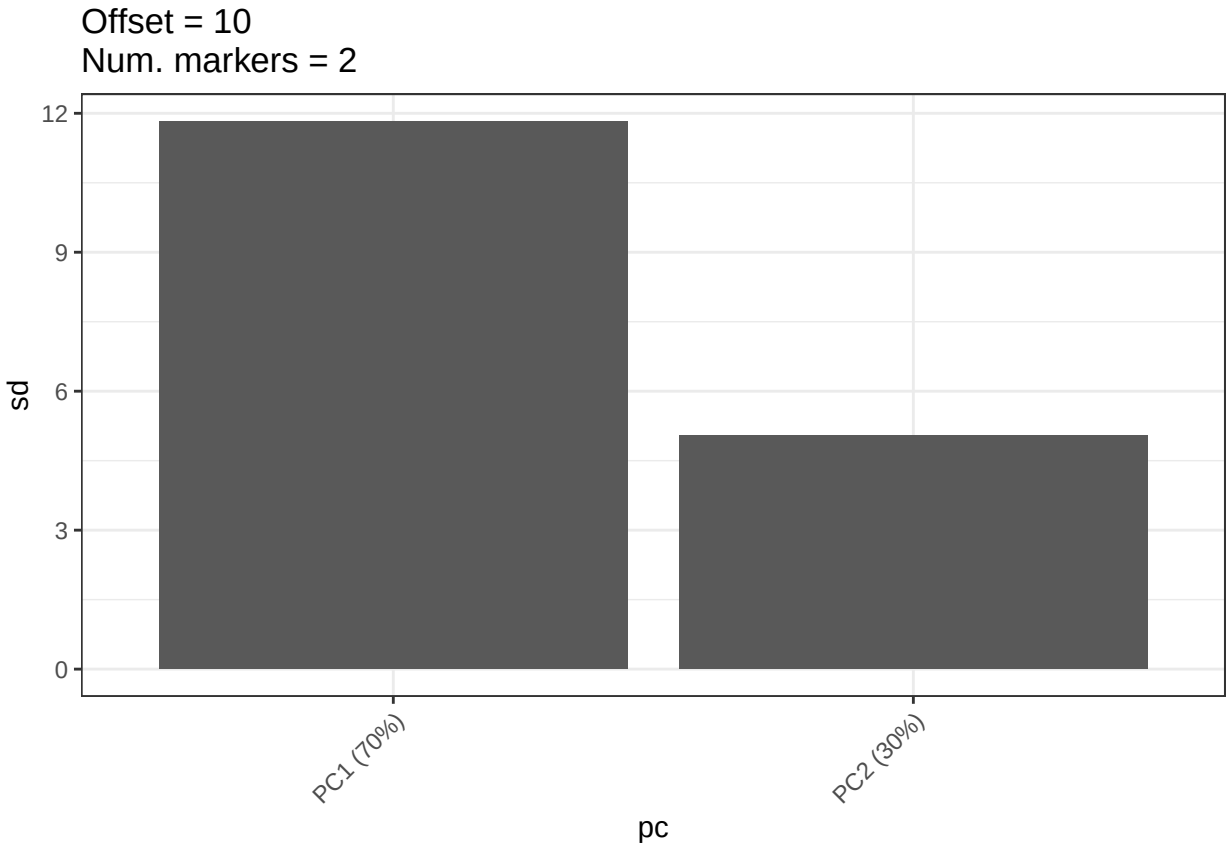
```
lexpt$`offset:2`$lpca.markers$pca.bydonortype$screeplot
```




```
lexpt$`offset:5`$lpca.markers$pca.bydonortype$screeplot
```




```
lexpt$`offset:10`$lpca.markers$pca.bydonortype$screeplot
```



Deconvolution results – all tests

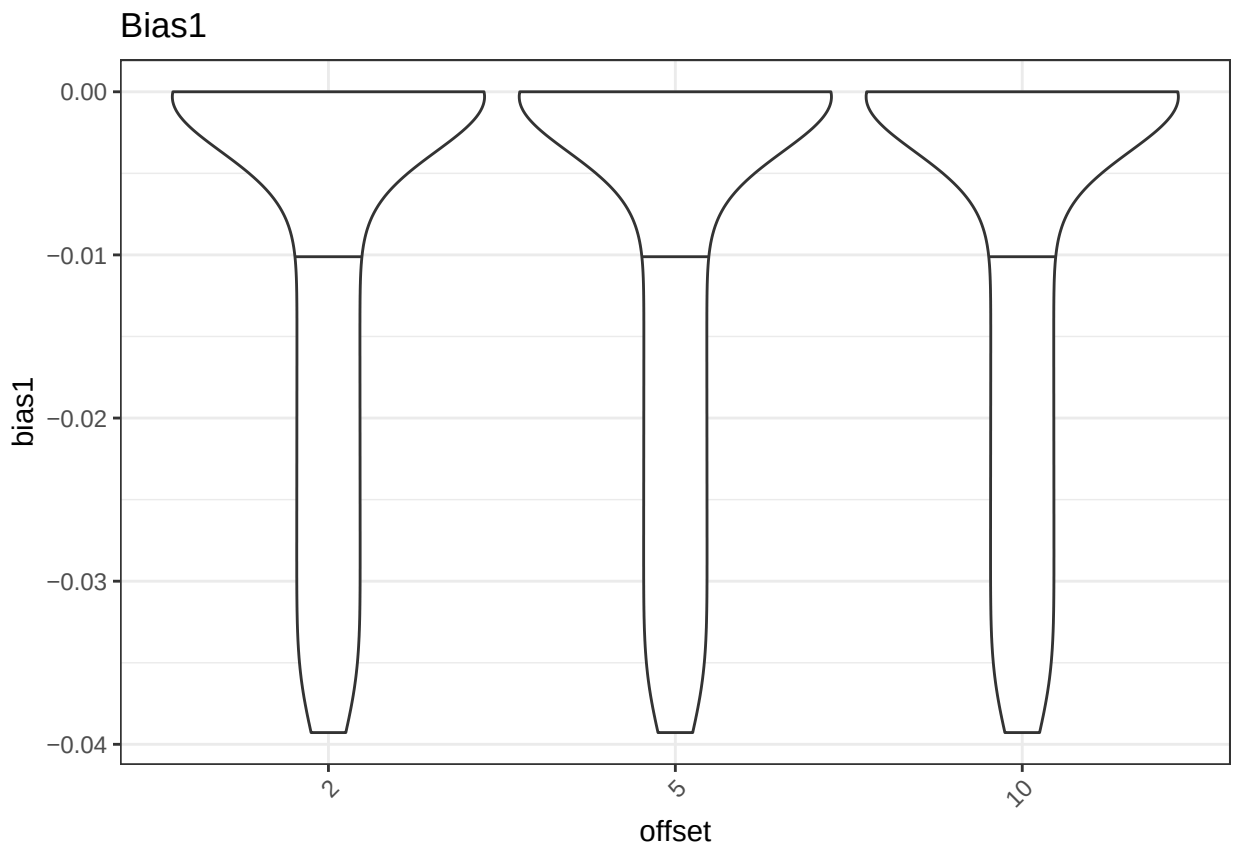
Results table

```
# get all results
dfres <- do.call(rbind, lapply(offsetv, function(offi){
  lres <- lexpt[grepl(paste0("offset:",offi,"$"), names(lexpt))][[1]]$decon.results
  dfresi <- do.call(rbind, lapply(lres, function(resi){resi$dfres}))
  dfresi <- as.data.frame(dfresi)
  dfresi$offset <- offi
  dfresi
}))
```

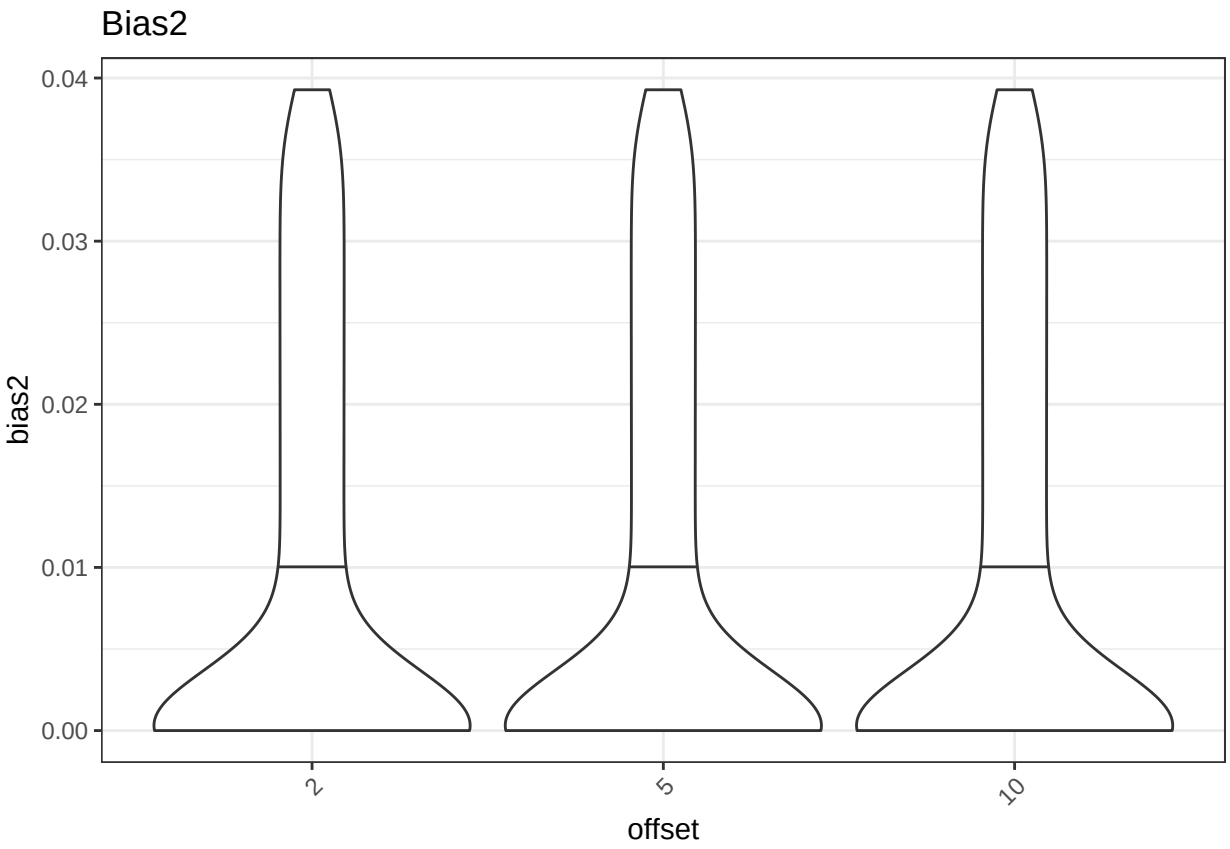
Bias violin plots

```
dfres$offset <- as.factor(dfres$offset)

ggplot(dfres, aes(x = offset, y = bias1)) + geom_violin(draw_quantiles = 0.5) +
  theme_bw() + ggtitle("Bias1") + theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

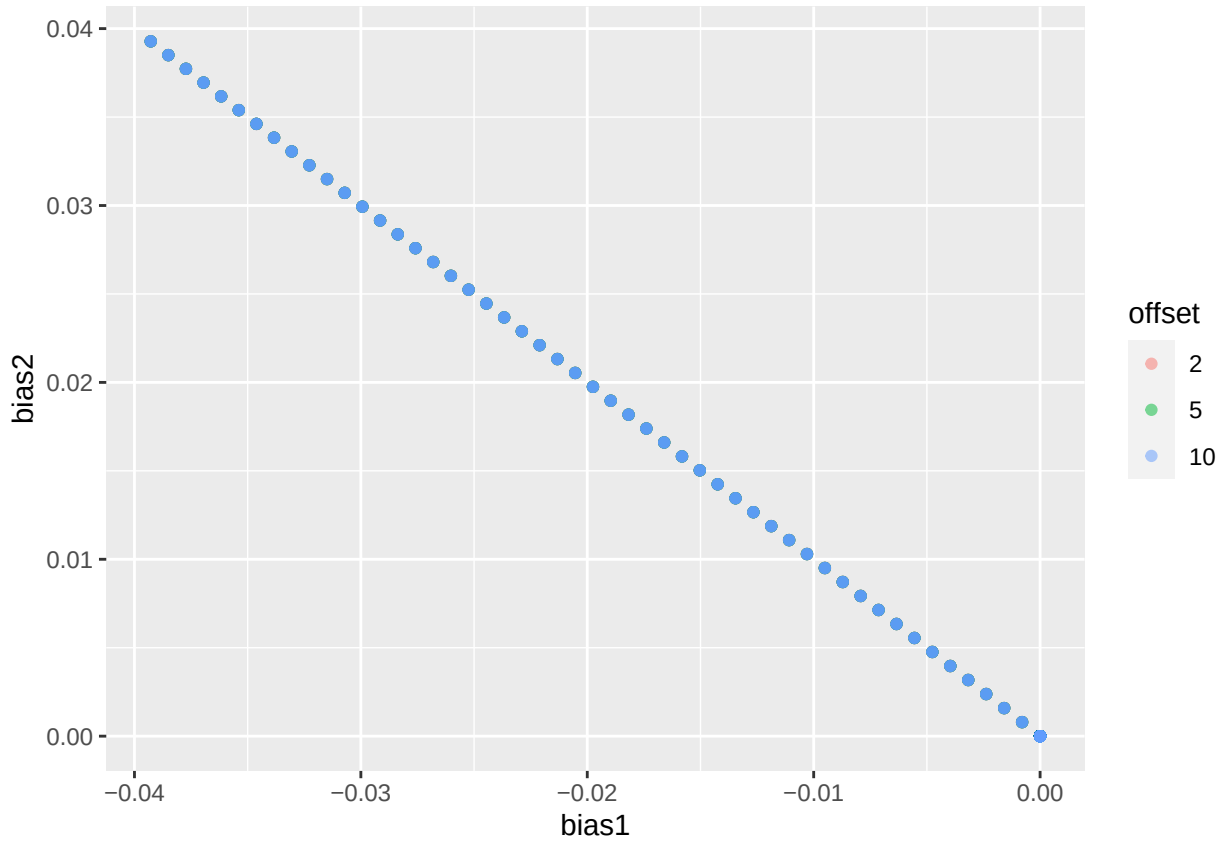



```
ggplot(dfres, aes(x = offset, y = bias2)) + geom_violin(draw_quantiles = 0.5) +
  theme_bw() + ggtitle("Bias2") + theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

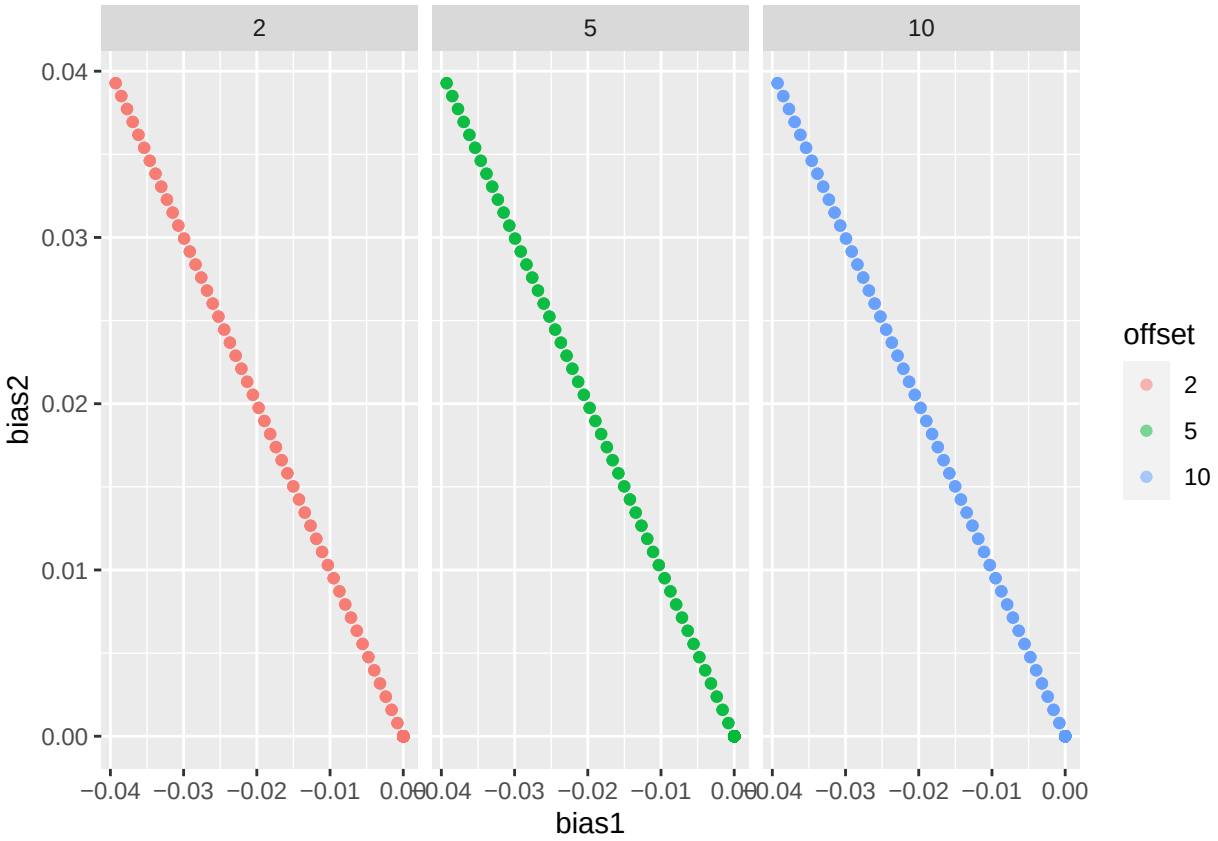


Bias scatterplots

```
ggpt <- ggplot(dfres, aes(x = bias1, y = bias2, color = offset)) + geom_point(alpha = 0.5)
ggpt
```

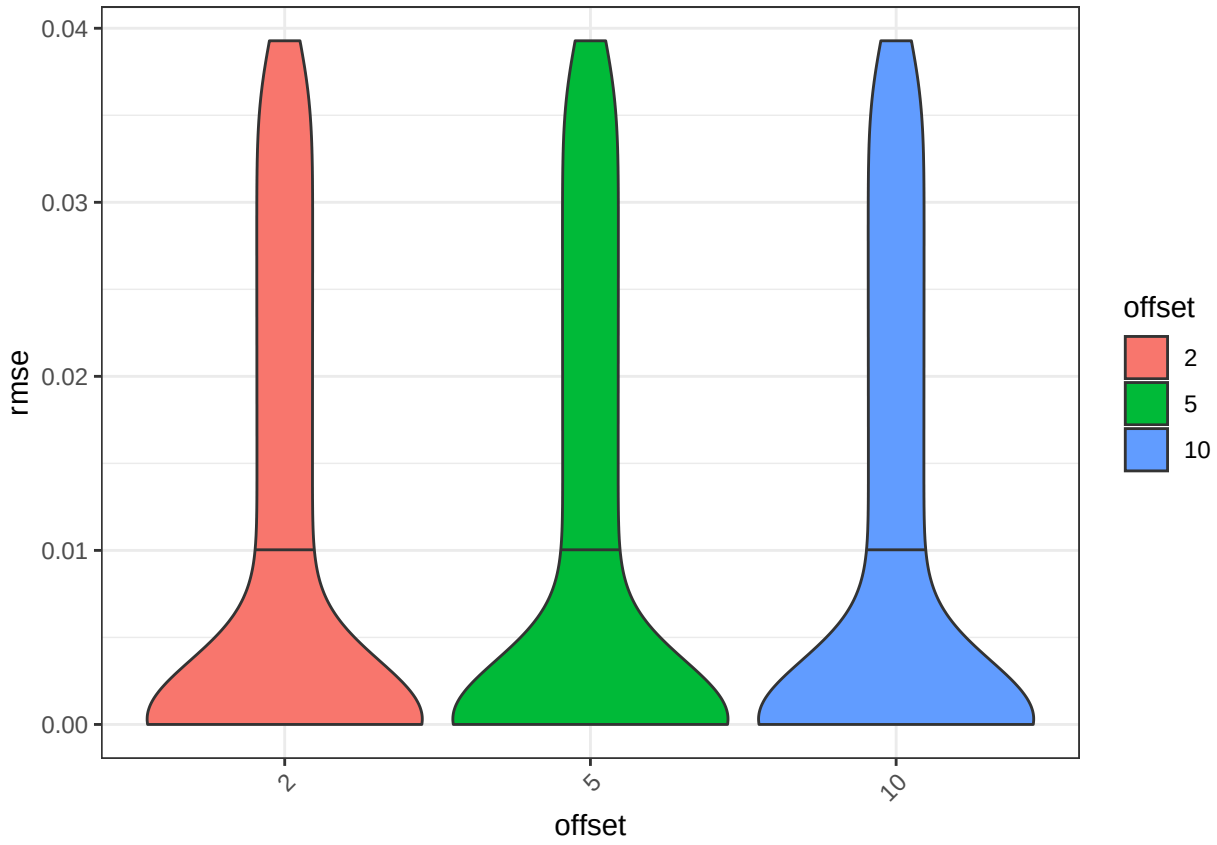



```
ggpt + facet_wrap(~offset)
```



RMSE violin plot

```
ggplot(dfres, aes(x = offset, y = rmse, fill = offset)) +
  geom_violin(draw_quantiles = 0.5) + theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Deconvolution results – without S transformations

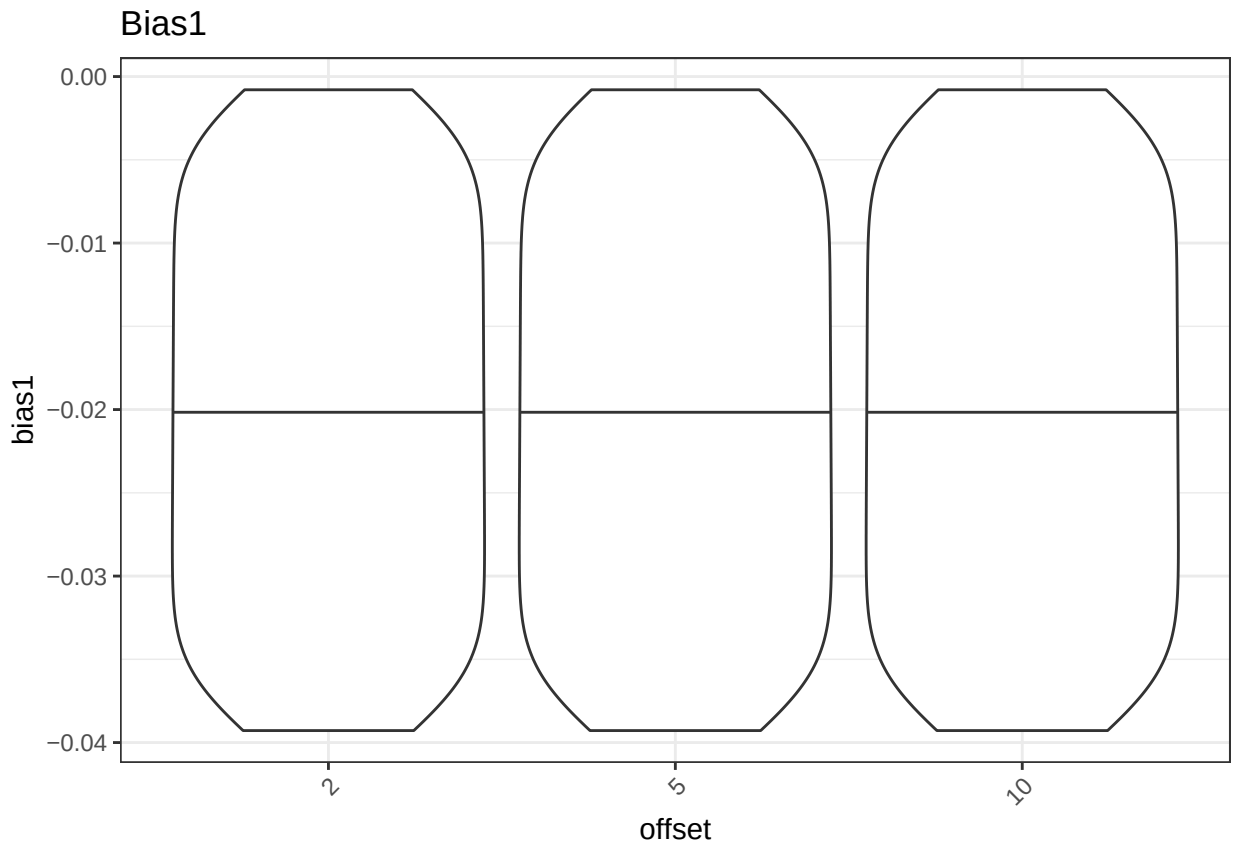
Results table

```
dfres.all <- dfres
dfres <- dfres.all[dfres.all$zs_transform==FALSE,]
```

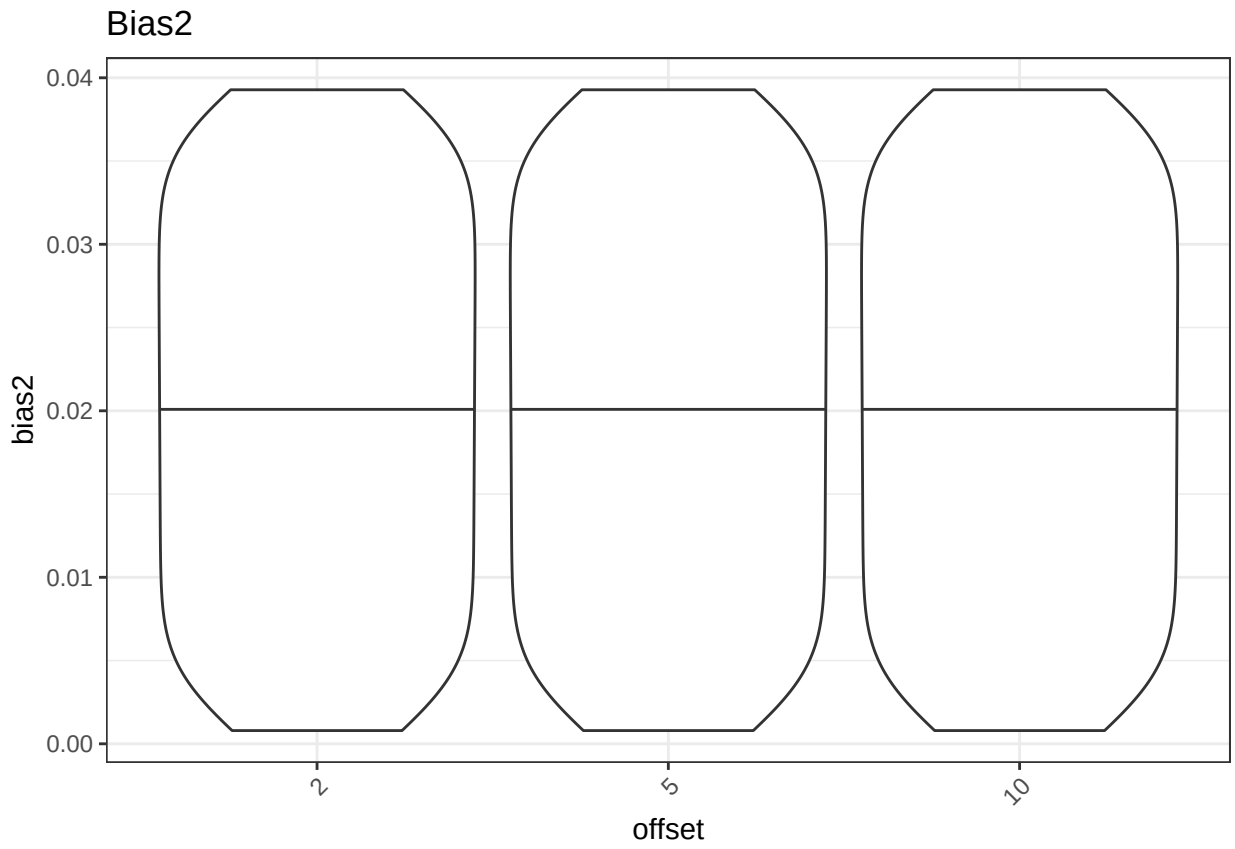
Bias violin plots

```
dfres$offset <- as.factor(dfres$offset)

ggplot(dfres, aes(x = offset, y = bias1)) + geom_violin(draw_quantiles = 0.5) +
  theme_bw() + ggtitle("Bias1") + theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

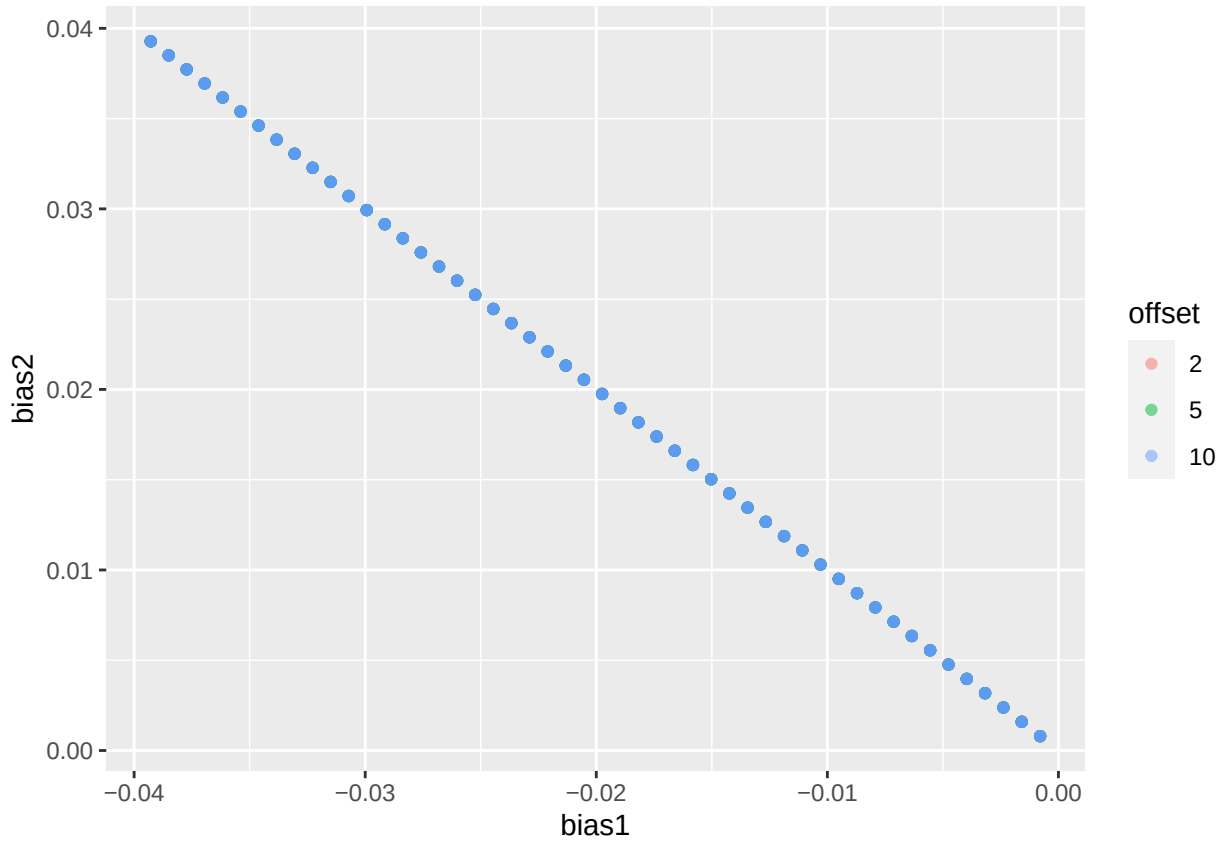



```
ggplot(dfres, aes(x = offset, y = bias2)) + geom_violin(draw_quantiles = 0.5) +
  theme_bw() + ggtitle("Bias2") + theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

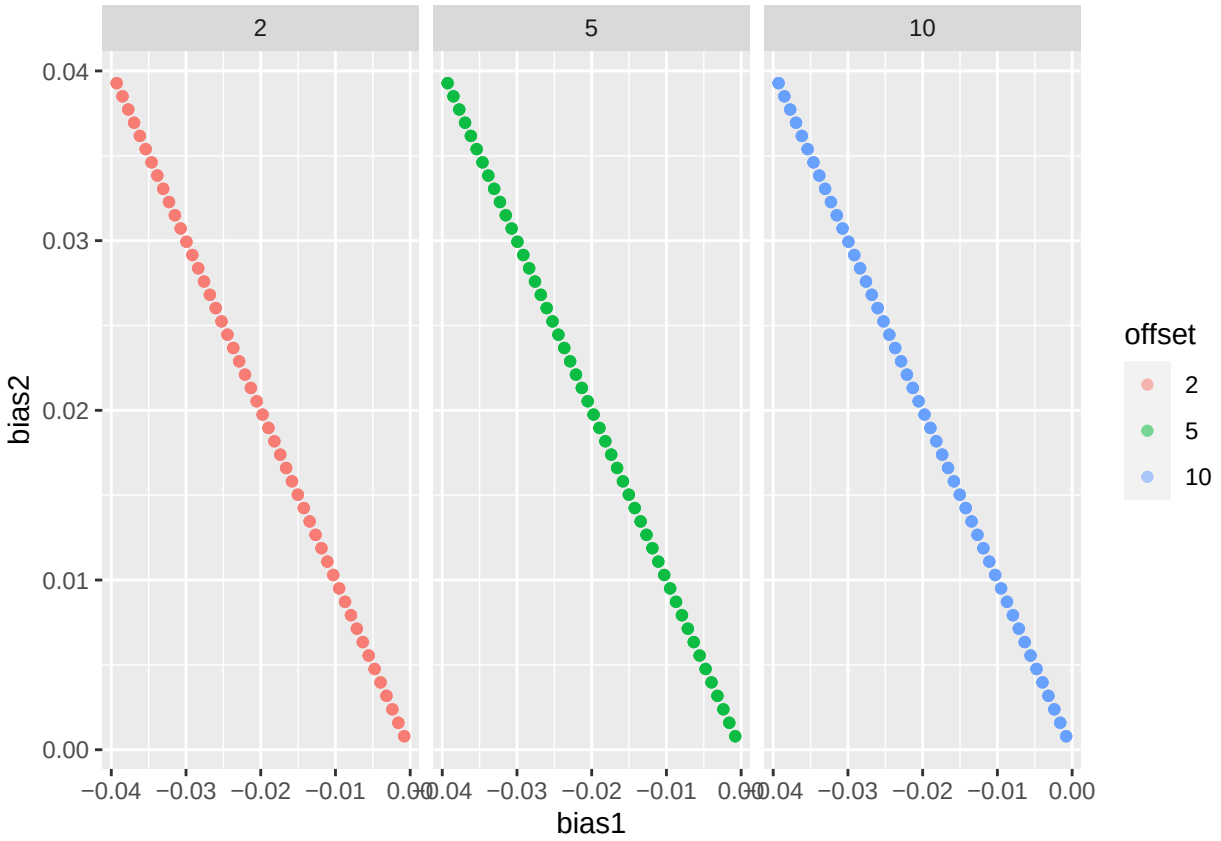


Bias scatterplots

```
ggpt <- ggplot(dfres, aes(x = bias1, y = bias2, color = offset)) + geom_point(alpha = 0.5)
ggpt
```

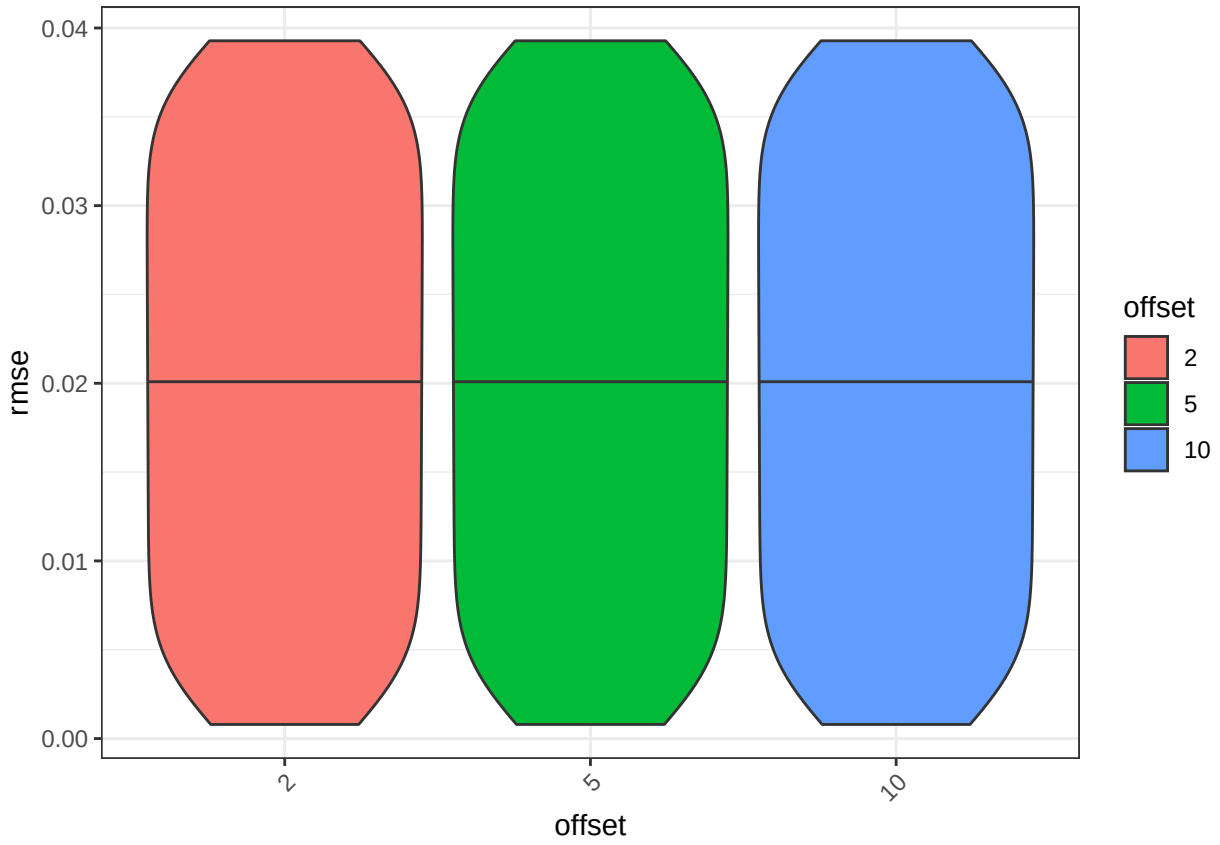



```
ggpt + facet_wrap(~offset)
```



RMSE violin plot

```
ggplot(dfres, aes(x = offset, y = rmse, fill = offset)) +
  geom_violin(draw_quantiles = 0.5) + theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



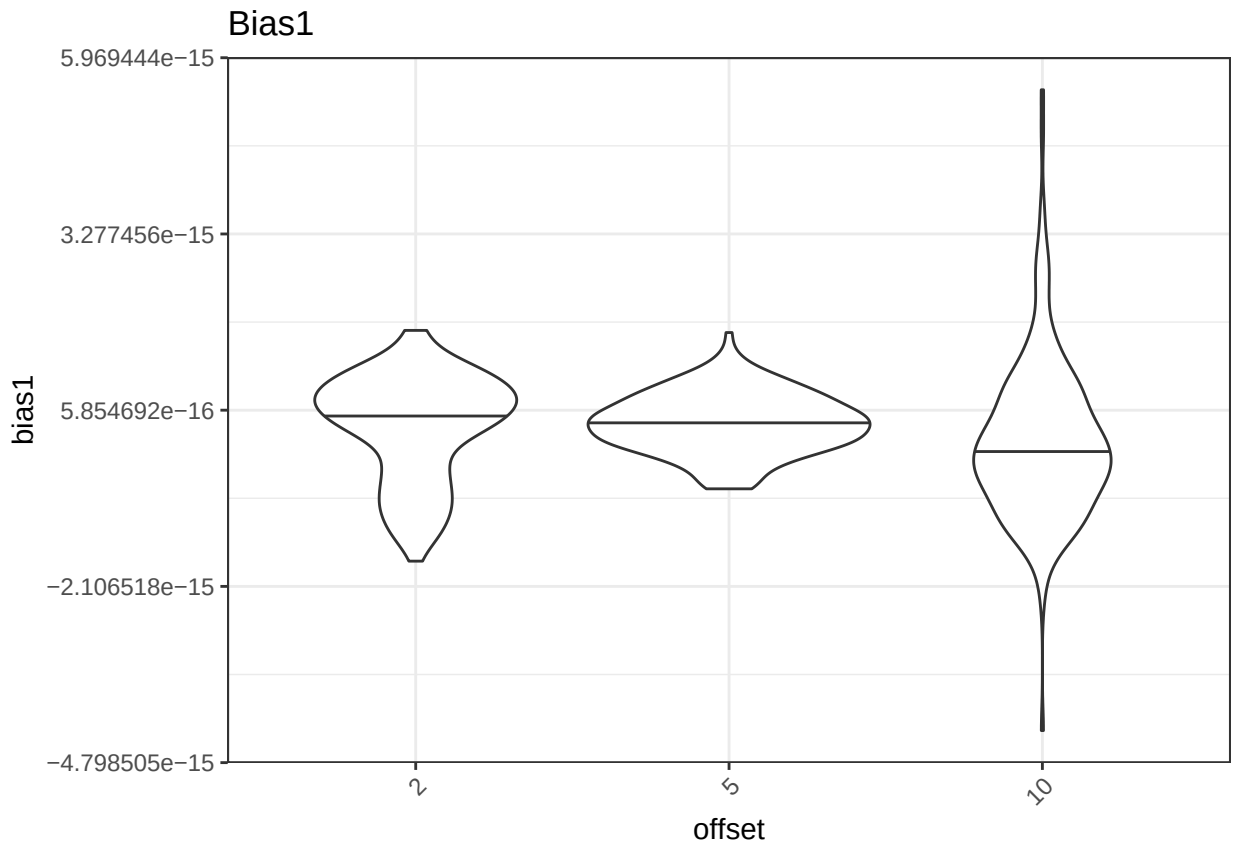
Deconvolution results – with S transformations

```
dfres <- dfres.all[dfres.all$zs_transform==TRUE,]
```

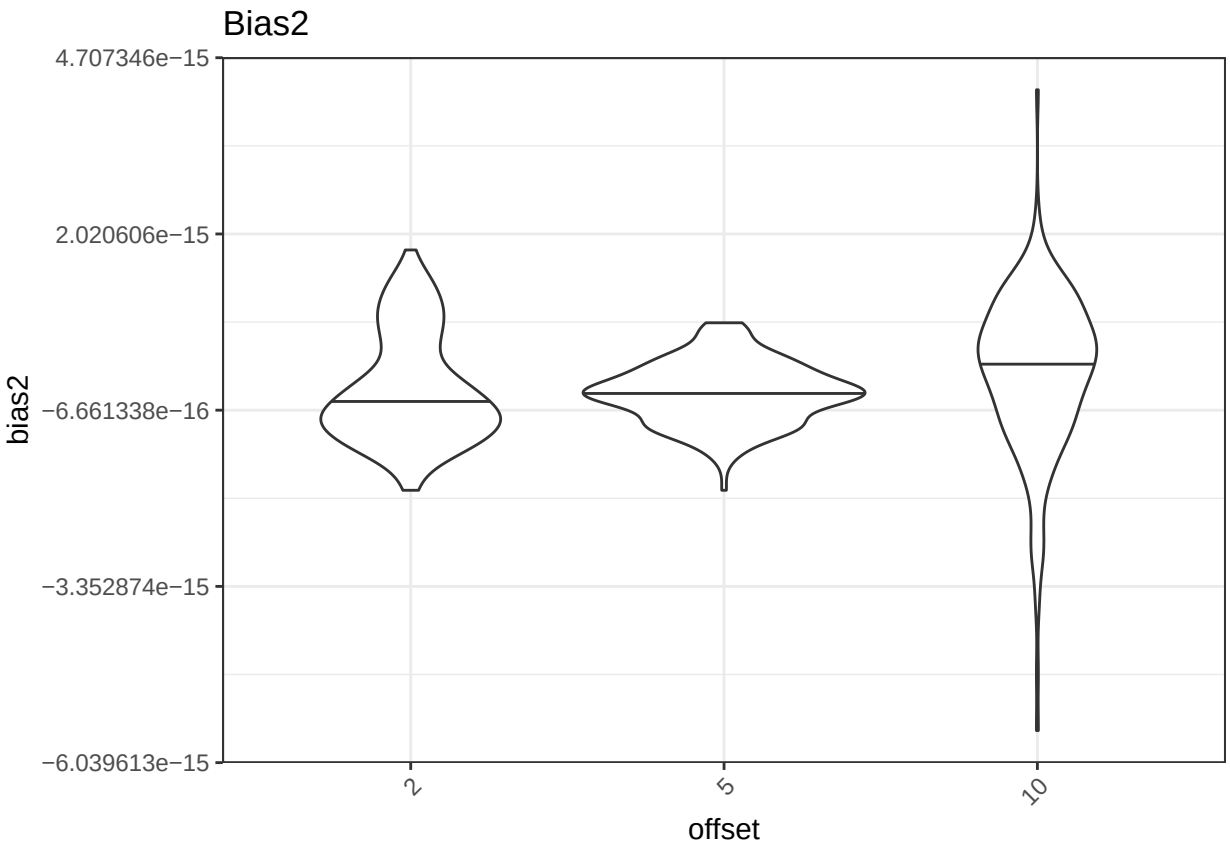
Bias violin plots

```
dfres$offset <- as.factor(dfres$offset)

ggplot(dfres, aes(x = offset, y = bias1)) + geom_violin(draw_quantiles = 0.5) +
  theme_bw() + ggtitle("Bias1") + theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

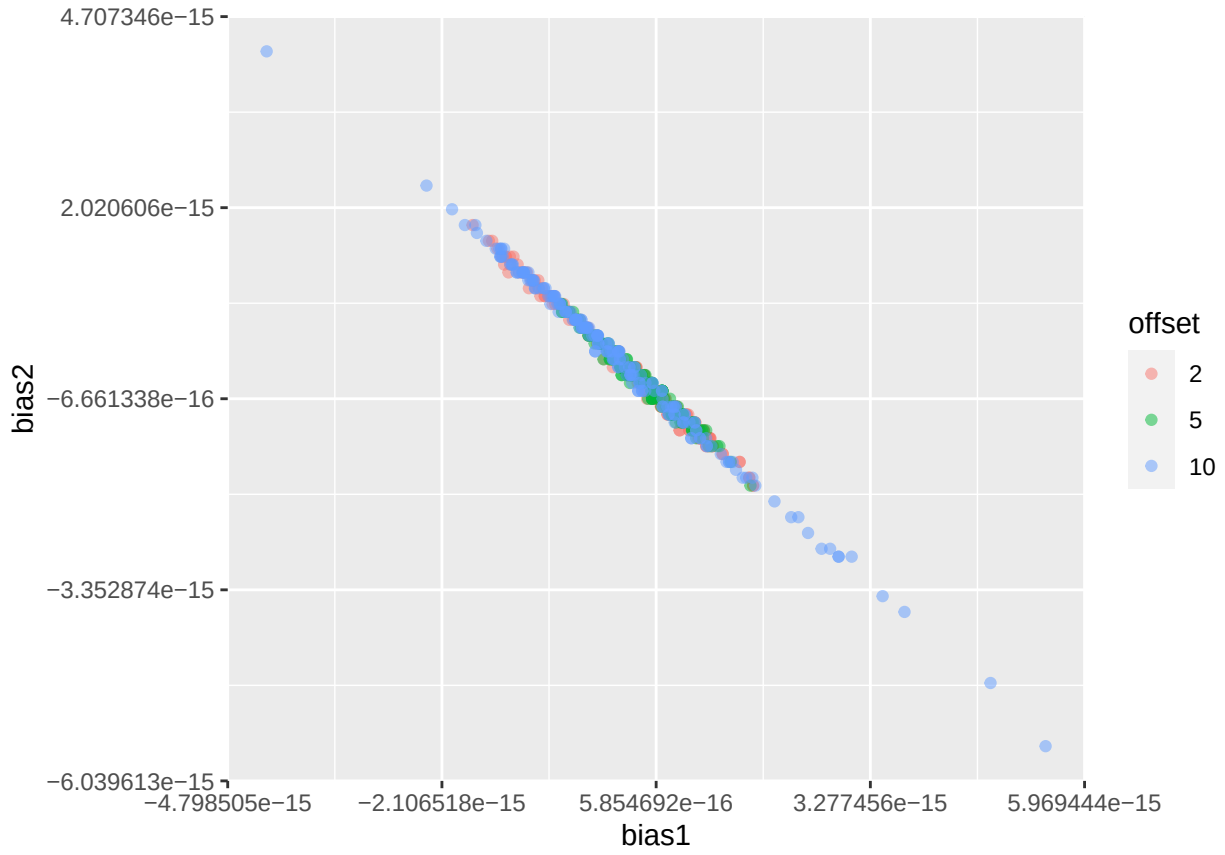



```
ggplot(dfres, aes(x = offset, y = bias2)) + geom_violin(draw_quantiles = 0.5) +
  theme_bw() + ggtitle("Bias2") + theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

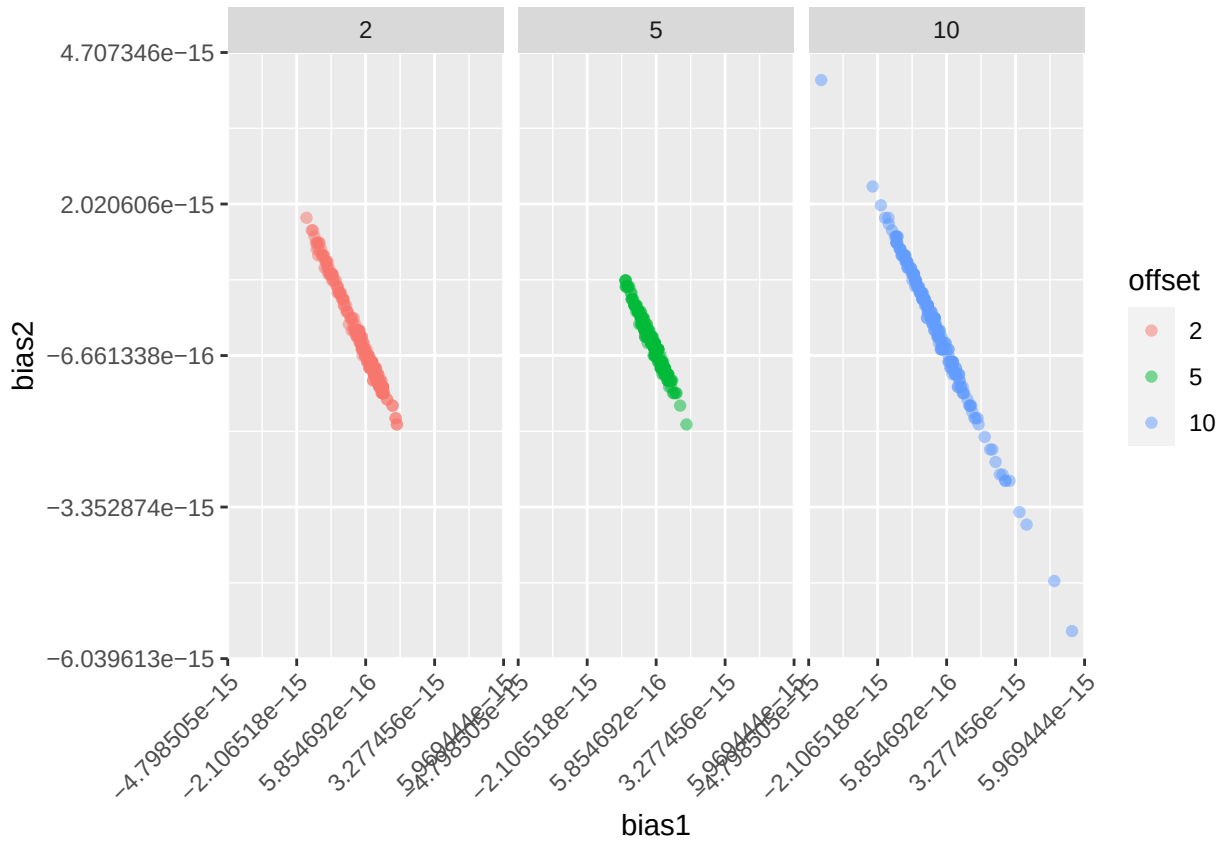


Bias scatterplots

```
ggpt <- ggplot(dfres, aes(x = bias1, y = bias2, color = offset)) + geom_point(alpha = 0.5)
ggpt
```

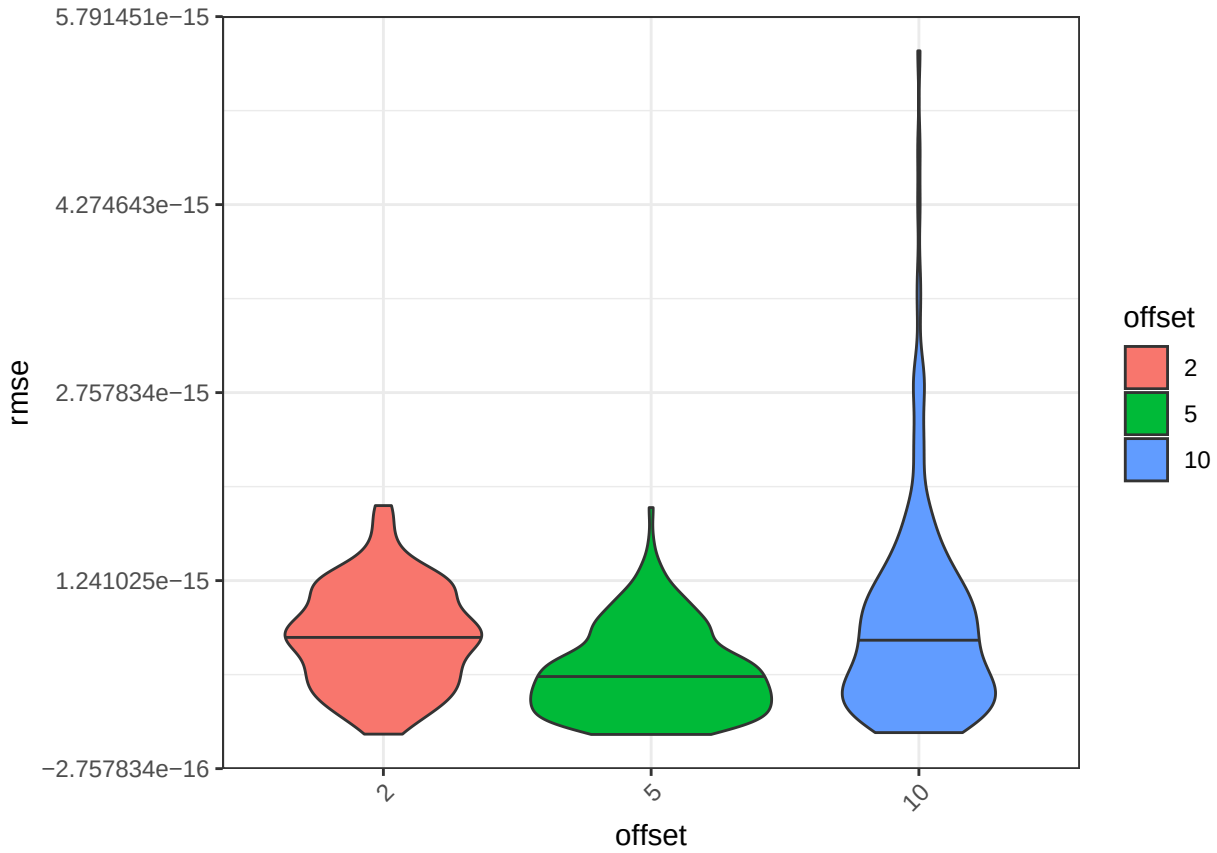



```
ggpt + facet_wrap(~offset) + theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



RMSE violin plot

```
ggplot(dfres, aes(x = offset, y = rmse, fill = offset)) +
  geom_violin(draw_quantiles = 0.5) + theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```


5.791451e−15
4.274643e−15
2.757834e−15
1.241025e−15
−2.757834e−16
rmse
2
5
10
offset
offset
2
5
10