

Statistical analysis of tissue-scale lifetime ratios

Erika Donà, Joseph D. Barry, Guillaume Valentin, Charlotte Quirin,
Anton Khmelinskii, Andreas Kunze, Sevi Durdu, Lionel R. Newton,
Ana Fernandez-Minan, Wolfgang Huber, Michael Knop, Darren Gilmour

October 18, 2014

Contents

1	Introduction	1
2	Load and inspect data	1
3	Statistical tests	2
4	Normality	4
5	Alternative tests	5

1 Introduction

In this vignette we present the statistical analysis that was performed on the tissue-scale lifetime ratios in the main paper.

2 Load and inspect data

The data was compiled into a table containing median whole-tissue ratios for each primordium.

```
> data("statsTable", package="DonaPLLP2013")
> x <- statsTable
> dim(x)
```

```
[1] 216  2
```

```
> head(x)
```

```
      ratio condition
1 0.2923994      WT
2 0.2386834      WT
3 0.1966154      WT
4 0.2129015      WT
5 0.2100342      WT
6 0.1991967      WT
```

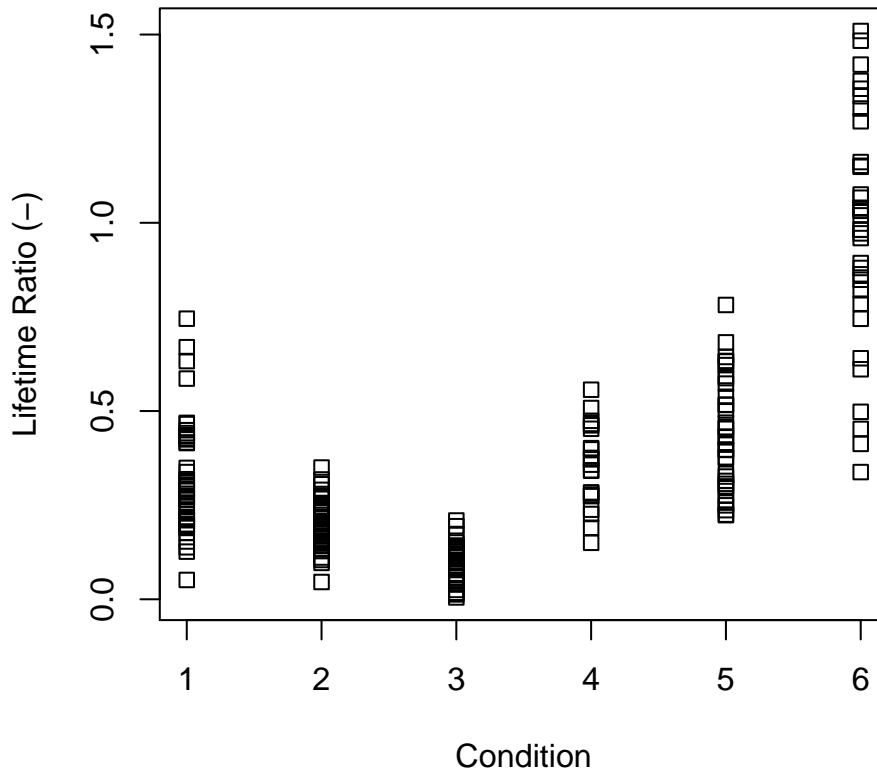
In total we had 6 conditions:

```
> table(x$condition)
```

Cxcl12a-/-	Cxcr4b-/-	Cxcr7-/-	Cxcr7-/-Cxcl12aMo
35	46	35	21
WT	mem-tFT		
45	34		

1. wild-type (WT),
2. a mutant of the tagged receptor *cxcr4b*^{-/-} (*Cxcr4b*^{-/-}),
3. a mutant of the rear ligand-sequestering receptor *cxcr7*^{-/-} (*Cxcr7*^{-/-}),
4. a *cxcr7*^{-/-} mutant with an additional morpholino knockdown of the signalling ligand *cxcl12a* (*Cxcr7*^{-/-}*Cxcl12a*Mo),
5. a mutant of the signalling ligand *cxcl12a*, also known as *sdf1a* (*Cxcl12a*^{-/-}), and
6. a membrane-tethered control protein tagged with the fluorescent timer (mem-tFT).

```
> splitByCond <- split(x$ratio, x$condition)
> plotOrder <- c("WT", "Cxcr4b-/-", "Cxcr7-/-", "Cxcr7-/-Cxcl12aMo", "Cxcl12a-/-",
+               "mem-tFT")
> splitByCond <- splitByCond[plotOrder]
> stripchart(splitByCond, vertical=TRUE, xlab="Condition", ylab="Lifetime Ratio (-)",
+            group.names=1:length(splitByCond))
```



For 1-5, the readout was the lifetime-ratio from a *cxcr4b* receptor tagged with the fluorescent timer, which was expressed from a bacterial artificial chromosome. For 6, the readout was the lifetime-ratio from a different, membrane-tethered control protein.

3 Statistical tests

We performed two-sided *t*-tests for each of the following comparisons of interest.

1. WT to *Cxcr4b*^{-/-}
2. WT to *Cxcr7*^{-/-}
3. WT to *Cxcl12a*^{-/-}
4. WT to mem-tFT
5. *Cxcr7*^{-/-} to *Cxcr7*^{-/-}*Cxcl12a*Mo

6. Cxcr4b-/- to Cxcr7-/-

```
> compareConds <- as.data.frame(
+   matrix(nr=6, data=c("WT", "WT", "WT",
+                       "WT", "Cxcr7-/-", "Cxcr7-/-",
+                       "Cxcr4b-/-", "Cxcr7-/-", "Cxcl12a-/-",
+                       "mem-tFT", "Cxcr7-/-Cxcl12aMo", "Cxcr4b-/-")
+   ), stringsAsFactors=FALSE)
> colnames(compareConds) <- c("condition 1", "condition 2")
```

Results from the *t*-tests were appended to our table.

```
> for (i in seq_len(nrow(compareConds))) {
+   res <- t.test(x$ratio[x$condition == compareConds[i,1]],
+                x$ratio[x$condition == compareConds[i,2]])
+   compareConds[i, "t"] <- res$statistic
+   compareConds[i, "df"] <- res$parameter
+   compareConds[i, "mean 1"] <- res$estimate[1]
+   compareConds[i, "mean 2"] <- res$estimate[2]
+   compareConds[i, "difference in means"] <- res$estimate[2]-res$estimate[1]
+   compareConds[i, "p.value"] <- res$p.value
+   compareConds[i, "method"] <- res$method
+ }
> compareConds
```

	condition 1	condition 2	t	df	mean 1	mean 2
1	WT	Cxcr4b-/-	4.907150	58.85822	0.3182417	0.2005986
2	WT	Cxcr7-/-	9.079875	56.46167	0.3182417	0.1028506
3	WT	Cxcl12a-/-	-3.599910	73.09063	0.3182417	0.4389546
4	WT	mem-tFT	-11.643242	44.59746	0.3182417	0.9844275
5	Cxcr7-/-	Cxcr7-/-Cxcl12aMo	-9.901493	25.21075	0.1028506	0.3537685
6	Cxcr7-/-	Cxcr4b-/-	-7.778590	78.68026	0.1028506	0.2005986

	difference in means	p.value	method
1	-0.11764313	7.661584e-06	Welch Two Sample t-test
2	-0.21539114	1.244956e-12	Welch Two Sample t-test
3	0.12071291	5.765433e-04	Welch Two Sample t-test
4	0.66618575	4.098828e-15	Welch Two Sample t-test
5	0.25091794	3.588092e-10	Welch Two Sample t-test
6	0.09774801	2.404200e-11	Welch Two Sample t-test

Multiple testing correction was performed using the method of Bonferroni. We noted that since the p-values are so small, this was not a critical step.

```
> compareConds[, "p.adjusted"] <- p.adjust(compareConds[, "p.value"],
+   method="bonferroni")
```

We preferred to view the table in decreasing order of the change in stability.

```
> compareConds[order(compareConds[, "condition 1"],
+   compareConds[, "difference in means"], decreasing=TRUE), ]
```

	condition 1	condition 2	t	df	mean 1	mean 2
4	WT	mem-tFT	-11.643242	44.59746	0.3182417	0.9844275
3	WT	Cxcl12a-/-	-3.599910	73.09063	0.3182417	0.4389546
1	WT	Cxcr4b-/-	4.907150	58.85822	0.3182417	0.2005986
2	WT	Cxcr7-/-	9.079875	56.46167	0.3182417	0.1028506
5	Cxcr7-/-	Cxcr7-/-Cxcl12aMo	-9.901493	25.21075	0.1028506	0.3537685
6	Cxcr7-/-	Cxcr4b-/-	-7.778590	78.68026	0.1028506	0.2005986

	difference in means	p.value	method	p.adjusted
4	0.66618575	4.098828e-15	Welch Two Sample t-test	
3	0.12071291	5.765433e-04	Welch Two Sample t-test	
1	-0.11764313	7.661584e-06	Welch Two Sample t-test	
2	-0.21539114	1.244956e-12	Welch Two Sample t-test	
5	0.25091794	3.588092e-10	Welch Two Sample t-test	
6	0.09774801	2.404200e-11	Welch Two Sample t-test	

```

4      0.66618575 4.098828e-15 Welch Two Sample t-test 2.459297e-14
3      0.12071291 5.765433e-04 Welch Two Sample t-test 3.459260e-03
1     -0.11764313 7.661584e-06 Welch Two Sample t-test 4.596950e-05
2     -0.21539114 1.244956e-12 Welch Two Sample t-test 7.469737e-12
5      0.25091794 3.588092e-10 Welch Two Sample t-test 2.152855e-09
6      0.09774801 2.404200e-11 Welch Two Sample t-test 1.442520e-10

```

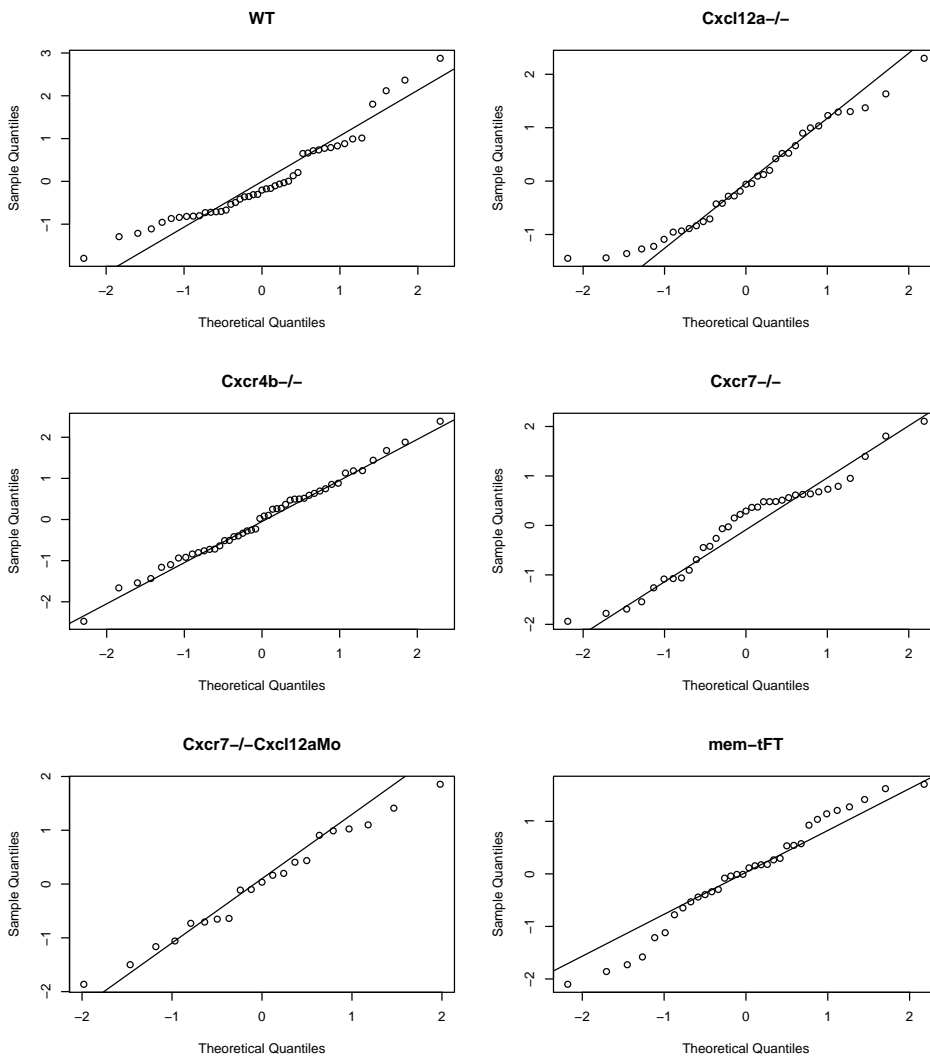
4 Normality

To assess whether the data were consistent with assumptions of normal distribution, we generated QQ-plots for each condition individually.

```

> myPlotQQ <- function(residuals, main) {
+   qqnorm(residuals, main=main)
+   qqline(residuals)
+ }
> standardize <- function(x) {(x-mean(x, na.rm=TRUE))/sd(x, na.rm=TRUE)}
> par(mfrow=c(3, 2))
> for (c in unique(x$condition)) {
+   dataPts <- standardize(x[x$condition == c, "ratio"])
+   myPlotQQ(dataPts, c)
+ }

```



The QQ plots indicated that the data was sufficiently close to being normally distributed.

5 Alternative tests

We also verified that an alternative, non-parametric test, the two-sided Mann-Whitney test (a two-sample Wilcoxon test), returned equivalent results.

```
> compareCondsMW <- compareConds[, c("condition 1", "condition 2")]
> for (i in seq_len(nrow(compareCondsMW))) {
+   res <- wilcox.test(x$ratio[x$condition == compareCondsMW[i, 1]],
+                     x$ratio[x$condition == compareCondsMW[i, 2]])
+   compareCondsMW[i, "W"] <- res$statistic
+   compareCondsMW[i, "p.value"] <- res$p.value
+   compareCondsMW[i, "method"] <- res$method
+ }
> compareCondsMW
```

	condition 1	condition 2	W	p.value	method
1	WT	Cxcr4b-/-	1583	7.594851e-06	Wilcoxon rank sum test
2	WT	Cxcr7-/-	1515	2.281662e-16	Wilcoxon rank sum test
3	WT	Cxcl12a-/-	419	2.695266e-04	Wilcoxon rank sum test
4	WT	mem-tFT	45	4.265137e-17	Wilcoxon rank sum test
5	Cxcr7-/-	Cxcr7-/-Cxcl12aMo	6	4.455117e-14	Wilcoxon rank sum test
6	Cxcr7-/-	Cxcr4b-/-	163	2.184994e-11	Wilcoxon rank sum test

We saw that the p-values were extremely similar to those generated by *t*-tests. Therefore the biological interpretation of our results was identical in both cases.