TF Cookbook

Zheng Zuo( )

2020

ii

Contents

Chapter 1

Preface

This is a book and online resource about how to study transcription factors (TFs) in principled and logically consistent way. Since the discovery of the first transcription factor lac repressor (lacI) in E. coli, there has been tremendous progress in this field, particularly after the wide adoption of high-throughput sequencing technology. However, so many "experts" and "non-experts" study TFs in so many different ways, both experiemntally and computationally, which cause lots of confusion and inconsistent results in various circumstances. As a nerd who has been working in this field for more than 10 years, I feel the need to write a small book myself to clarify some basic concepts of protein-DNA interactions, and demostrate how to design, perform, and analyze transcription factors in principled and logical consistent way. So this is not a encyclopedia on TFs, but a "How to" reference for everyone.

This book is an evolving project, currently authored and maintained by Zheng. I certainly welcome more TF researchers to contribute to this online book project in the long run. If you are interested to be part of this effort, please do not hesitate to contact me (zzuo@stanford.edu).

1.1 Why study TFs?

To be simple, because life matters.

To be honest, because some nerds like me need find some boring topics to sudy, to survice in academia.

To be specific, there are ~20,000 genes in human genome, and ~1,500 of them are transcription factors, which control how other genes work. When some mutations happen in those TFs, people get diseases. Understanding how those TFs work can help us figure out how biology works, and why we get sick in many cases.

1.2   Organization of this book

1.3   Who am I?

The **bookdown** package can be installed from CRAN or Github:

```r
install.packages("bookdown")
# or the development version
# devtools::install_github("rstudio/bookdown")
```

Remember each Rmd file contains one and only one chapter, and a chapter is defined by the first-level heading `#`.

To compile this example to PDF, you need XeLaTeX. You are recommended to install TinyTeX (which includes XeLaTeX): https://yihui.name/tinytex/.

Chapter 2

Prerequsites

This is a *sample* book written in **Markdown**. You can use anything that Pandoc's Markdown supports, e.g., a math equation $a^2 + b^2 = c^2$.

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter 2. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter **??**.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```r
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 2.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 2.1.

```r
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the **bookdown** package (Xie, 2019) in this sample book, which was built on top of R Markdown and **knitr** (**?**).
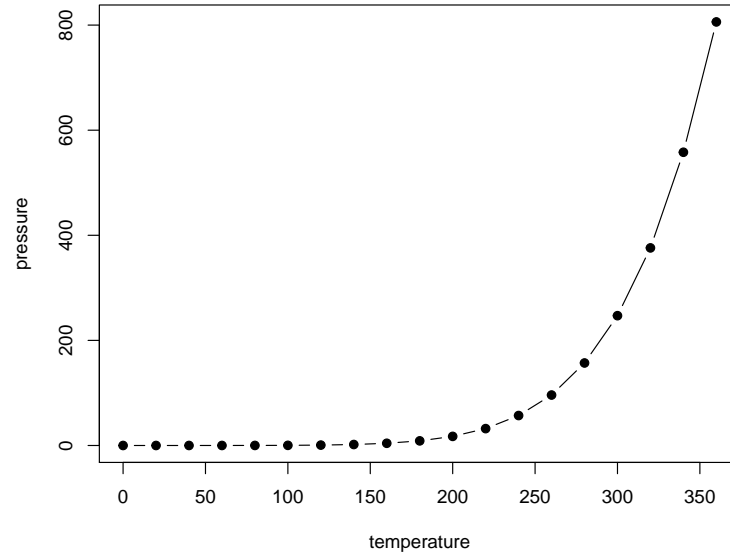
Figure 2.1: Here is a nice figure!

Table 2.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Chapter 3

Experimental Methods

3.1   In-vitro measurements

3.1.1   SELEX

3.1.2   PBM

3.1.3   B1H

3.1.4   Spec-seq

3.1.5   MITOMI

3.2   In-vivo methods

3.2.1   ChIP-seq

Chapter 4

Data Analysis

4.1 Bit-level nucleotide coding

4.2 Position-weighted motif representation

4.2.1 Additivity assumption

4.3 Cooperativity analysis

4.4 Methylation sensitivity analysis

Chapter 5

Applications

5.1 Motif analysis of tandem zinc finger proteins (ZFPs)

Some *significant* applications are demonstrated in this chapter.

5.2 Cooperativity analysis of SOX2/17-Oct4 TF pairs

5.3 Methylation sensitivity analysis of CTCF and ZFP57

Chapter 6

Online resources

Chapter 7

Literatures

Bibliography

Xie, Y. (2019). *bookdown: Authoring Books and Technical Documents with R Markdown.* R package version 0.16.