

Human Genome Analysis Lab 4 : Publication Graphics

- Learning objectives
- Fine tuning ggplots
 - Controlling graph size in RMarkdown
 - Graphic Output
 - RMarkdown loading images
 - Interactive graphs and tables in RMarkdown reports
- Exercises
 - Exercise 1
 - Exercise 2
 - Exercise 3
 - Exercise 4
 - Exercise 5
 - Exercise 6

Learning objectives

- Be able to add title, axis labels, legends, colors to ggplot graphs
- Resize graphs in RMarkdown
- Print graphics to a file (e.g. jpeg, pdf)
- Loading images into a RMarkdown file
- Making interactive graphs and tables in RMarkdown

Fine tuning ggplots

Today we will use the Cookbook for R by Winston Chang (<http://www.cookbook-r.com/Graphs/>) for tidying up our graphs. Please take your time and go through the following web pages. Copy and paste the example code into R Studio.

- Cookbook for R - Title ([http://www.cookbook-r.com/Graphs/Titles_\(ggplot2\)/](http://www.cookbook-r.com/Graphs/Titles_(ggplot2)/))
- Cookbook for R - Axes ([http://www.cookbook-r.com/Graphs/Axes_\(ggplot2\)/](http://www.cookbook-r.com/Graphs/Axes_(ggplot2)/))
- Cookbook for R - Legends ([http://www.cookbook-r.com/Graphs/Legends_\(ggplot2\)/](http://www.cookbook-r.com/Graphs/Legends_(ggplot2)/))
- Cookbook for R - Colors ([http://www.cookbook-r.com/Graphs/Colors_\(ggplot2\)/](http://www.cookbook-r.com/Graphs/Colors_(ggplot2)/))

Here are a couple of cheatsheets that can be useful

- R Studio ggplot2 cheatsheet (<http://www.rstudio.com/wp-content/uploads/2015/12/ggplot2-cheatsheet-2.0.pdf>)
- RMarkdown cheatsheet (<https://rstudio.com/wp-content/uploads/2016/03/rmarkdown-cheatsheet-2.0.pdf>)

Controlling graph size in RMarkdown

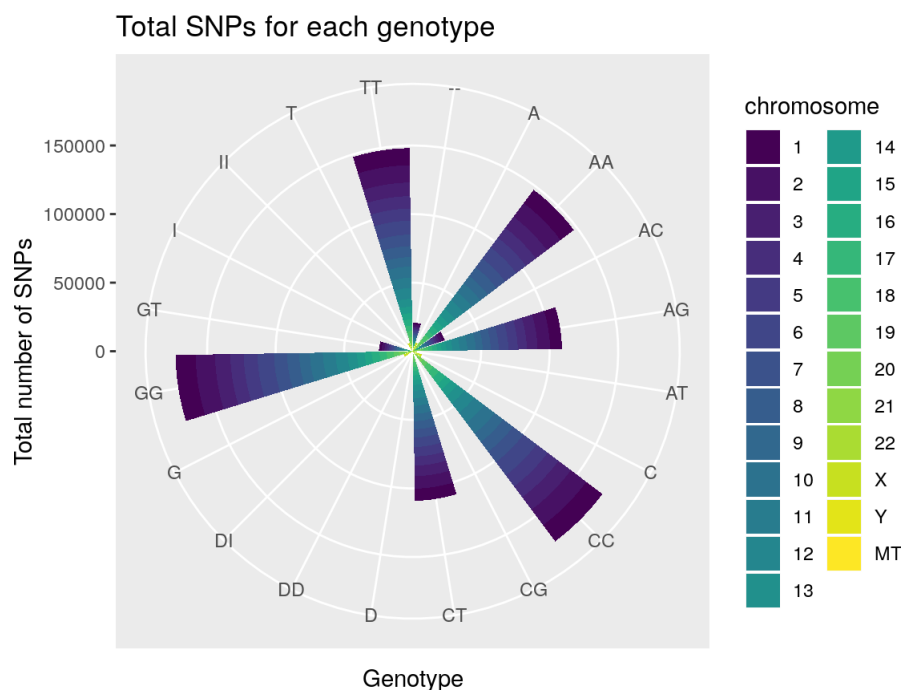
In the opening line of the RMarkdown code chunk {r} you can control the output of the code, graphs, tables using knitr syntax. For example if {r, eval = FALSE} the code will not be run, but will be shown. If {r, code = FALSE} the code will not be shown, but will be run and the output will be shown (useful in reports where the reader is only interested in the results/graphs, but not the code). You can also suppress error messages and warnings so that the reader isn't bothered by them (but you should take notice).

The dimensions of an individual graph in the RMarkdown document be adjusted by specifying the graph dimensions as below.

```
# Load Libraries
library(tidyverse)
```

```
# Load Files
SNPs<- read.table("23andMe_complete.txt", header = TRUE, sep = "\t")
```

```
# to adjust figure size {r, fig.width = 6, fig.height = 6}
SNPs$chromosome = ordered(SNPs$chromosome, levels=c(seq(1, 22), "X", "Y", "MT"))
ggplot(data = SNPs) +
  geom_bar(mapping = aes(x = genotype, fill = chromosome)) +
  coord_polar() +
  ggtitle("Total SNPs for each genotype") +
  ylab("Total number of SNPs") +
  xlab("Genotype")
```



Graphic Output

You may have realized that you can export plots in R Studio by clicking on Export in the Plots window that appears after you make a graph. You can save as a pdf, svg, tiff, png, bmp, jpeg and eps. You can also write the output directly to a file. This is particularly useful for controlling the final dimensions in a reproducible way and for manuscripts.

- Cookbook for R - Output to a file - PDF, PNG, TIFF, SVG (http://www.cookbook-r.com/Graphs/Output_to_a_file/)

```
# Plot graph to a pdf outputfile
pdf("SNP_example_plot.pdf", width=6, height=3)
ggplot(data = SNPs) +
  geom_bar(mapping = aes(x = chromosome, fill = genotype))
dev.off()
```

```
## png
## 2
```

```
# Plot graph to a png outputfile
ppi <- 300
png("SNP_example_plot.png", width=6*ppi, height=6*ppi, res=ppi)
ggplot(data = SNPs) +
  geom_bar(mapping = aes(x = chromosome, fill = genotype))
dev.off()
```

```
## png
## 2
```

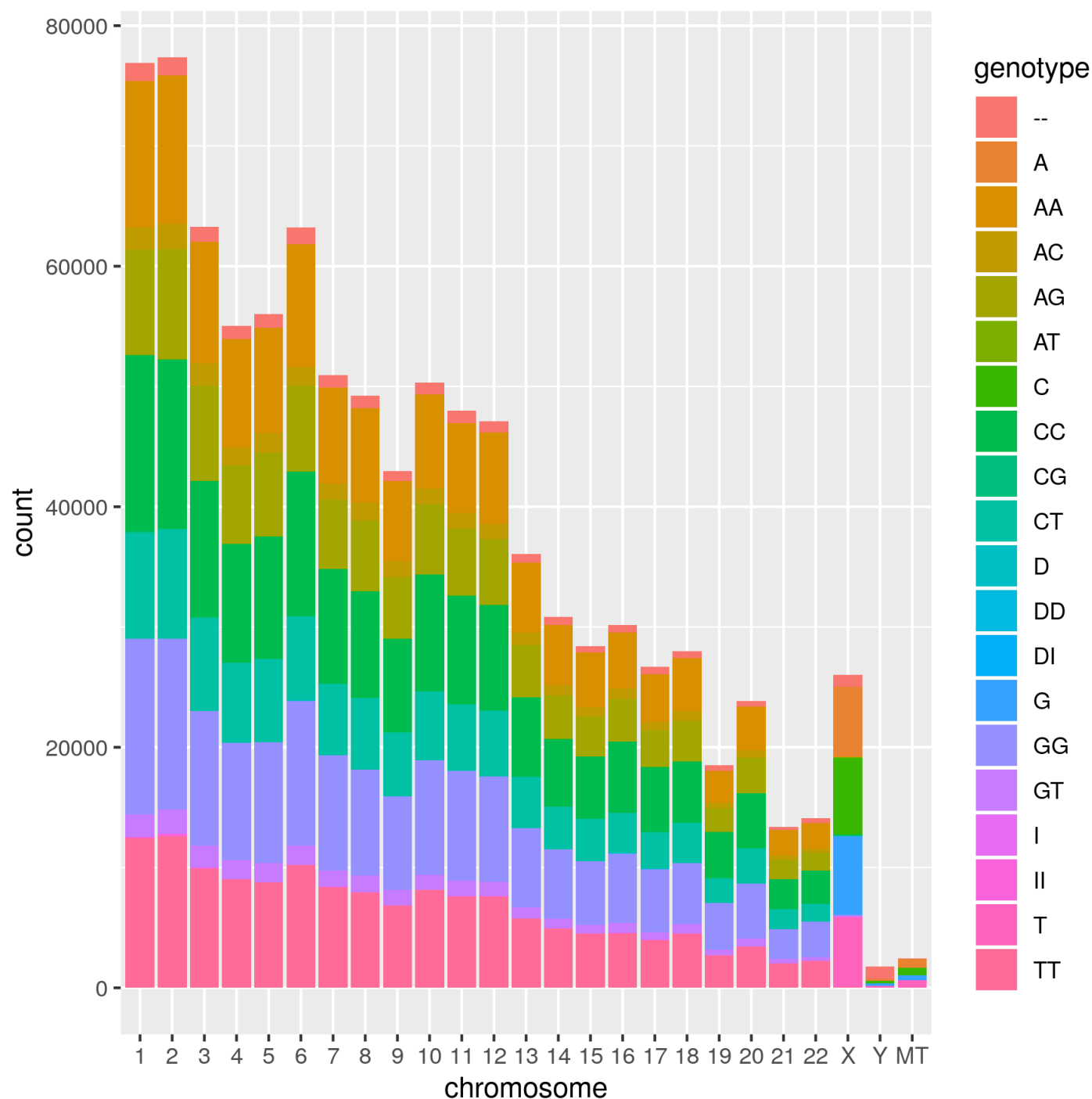
For more details on sizing output Cookbook for R - Output to a file - PDF, PNG, TIFF, SVG (http://www.cookbook-r.com/Graphs/Output_to_a_file/)

RMarkdown loading images

Sometimes it is useful in controlling the image layout for a report to file with the graph and then subsequently load it into the .Rmd file. This works with png files, but not pdfs. You can also upload images made with other bioinformatic tools into your RMarkdown report.

```
# This is the RMarkdown style for inserting images
# Your image must be in your working directory
# This command is put OUTSIDE the r code chunk

![Genotype counts per chromosome](SNP_example_plot.png)
```

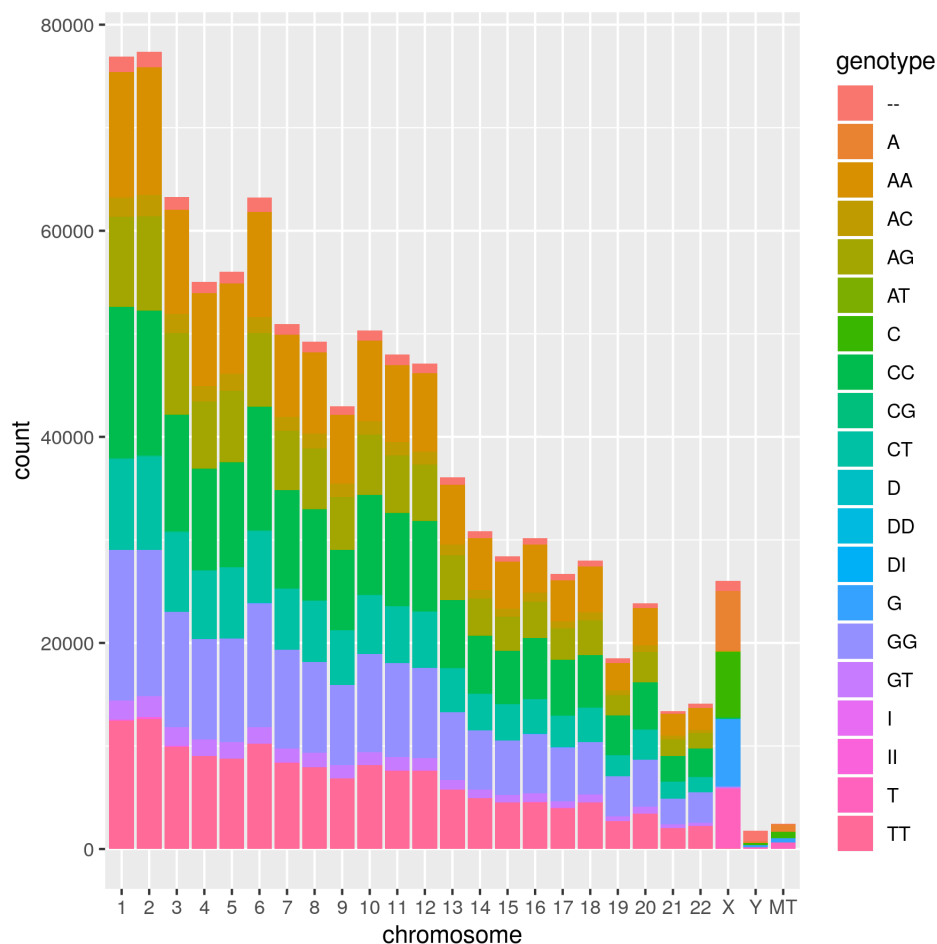


Genotype counts per chromosome

```
# This is an alternative way using html.
# Remember that it must be in your working directory or you will need to specify the full path.
# The html is put OUTSIDE the r code chunk.
```

```

```



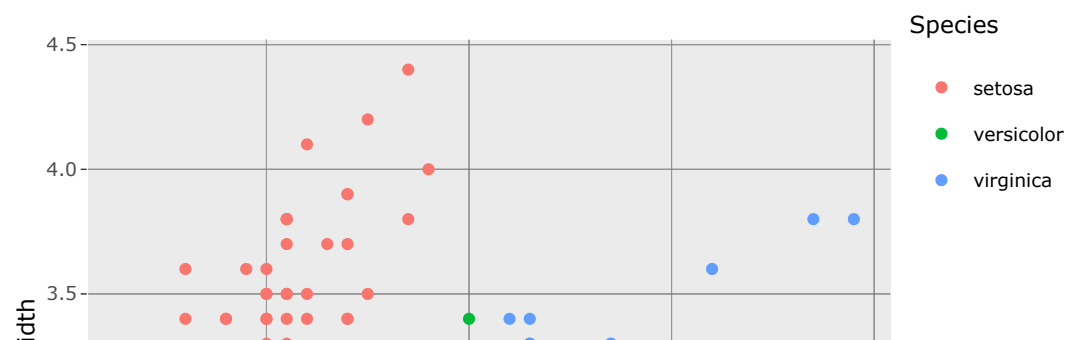
Another way to present a graph without the code is adding `echo = FALSE` within the `r{}` chunk - `{r echo = FALSE}`. This prevents code, but not the results from appearing in the knitr file.

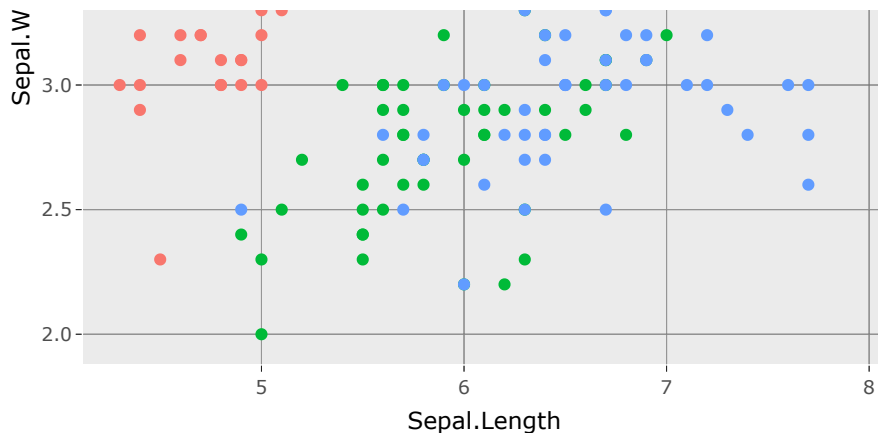
Interactive graphs and tables in RMarkdown reports

With `plotly/ggplotly` (<https://plot.ly/ggplot2/>) you can make interactive graphs in your lab report.

```
# Version 1 1
library(plotly)
p <- ggplot(data = iris, aes(x = Sepal.Length, y = Sepal.Width, color = Species)) +
  geom_point()
ggplotly(p)
```

```
# Version 2
library(plotly)
ggplotly(
  ggplot(data = iris, aes(x = Sepal.Length, y = Sepal.Width, color = Species)) +
    geom_point()
)
```





You can also make interactive data tables with the DT package (<https://rstudio.github.io/DT/>) *** Don't do this with tables of hundreds of thousands of rows (as in your complete SNP table)

```
library(DT)
datatable(iris)
```

Show entries

Search:

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa

Showing 1 to 10 of 150 entries

Previous 2 3 4 5 ... 15 Next

Exercises

Exercise 1

Add title and labels for the x and y axis to Lab3 ex1. Color the bars blue

Exercise 2

To Lab3 ex3 add more defined x and y axis labels, add a title, Change the colors of the genotypes, so that the dinucleotides (e.g. AA) are one color, the mononucleotides (A) are another and D's and I's a third color. One way to do this is to specify the color of each genotype.

Exercise 3

From Lab3 ex5, make an output png file, then load the file into report using the RMarkdown or html format.

Exercise 4

For Lab3 ex6 add more descriptive x and y axis labels, add a title, make the x-axis for each graph readable in your final report file.

Exercise 5

Turn Lab3 ex6 into an interactive graph using plotly

Exercise 6

Make an interactive table of the SNPS found in on the Y chromosome from the 23andMe_complete data set