Distribution of Size of TADs

Spiro Stilianoudakis

August 14, 2018

Loading Libraries

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.4.4
```

```
library(plyr)
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 3.4.4
```

Read in domain data

```
domains <- readRDS("C:/Users/Spiro Stilianoudakis/Documents/TAD_data/RData/GM12878/domains.rds")
```
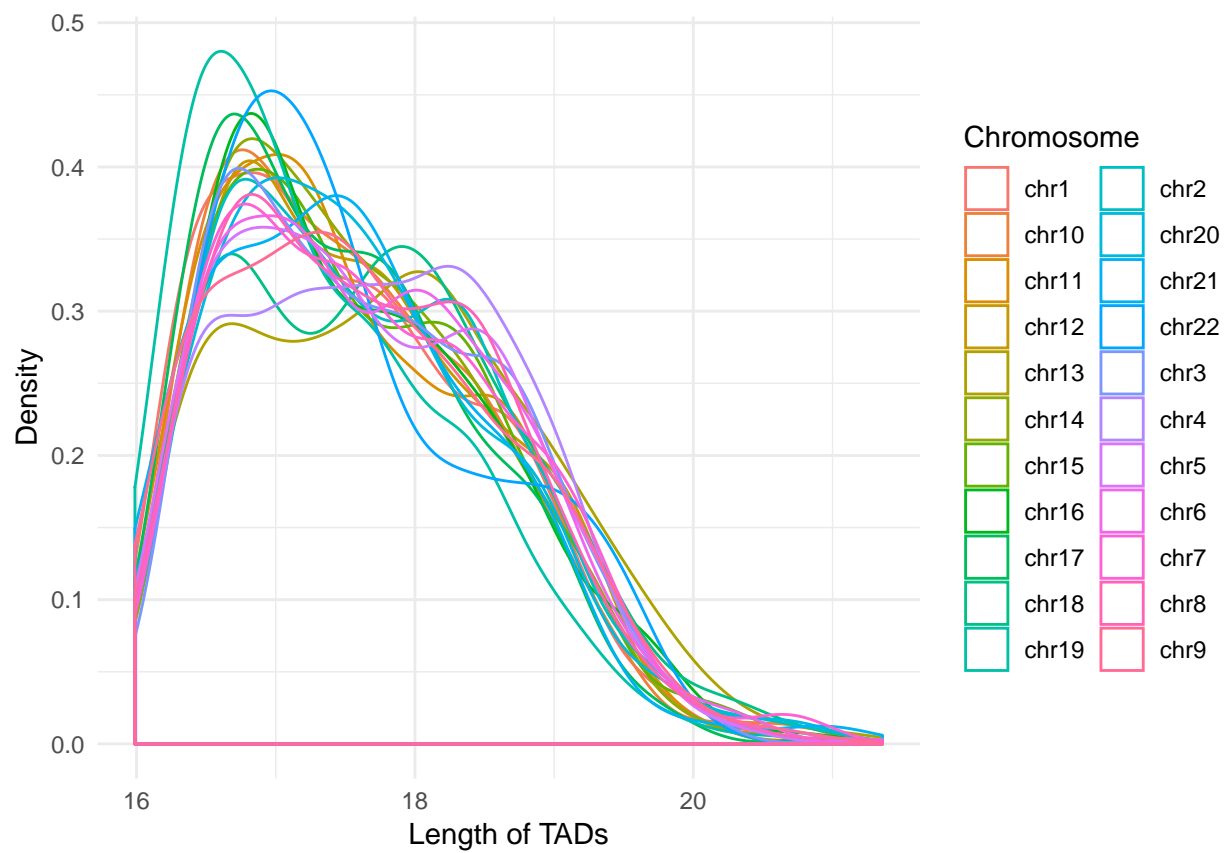
Density plots of TAD distances by chromosome

```
#plotting distances grouped by chromosome
ggplot() +
  geom_density(data=domains, aes(x=distance, group=Chromosome, color=Chromosome)) +
  xlab("Length of TADs") + xlim(0,1000000) +
  ylab("Density")+
  theme_minimal()
```

```
## Warning: Removed 113 rows containing non-finite values (stat_density).
```
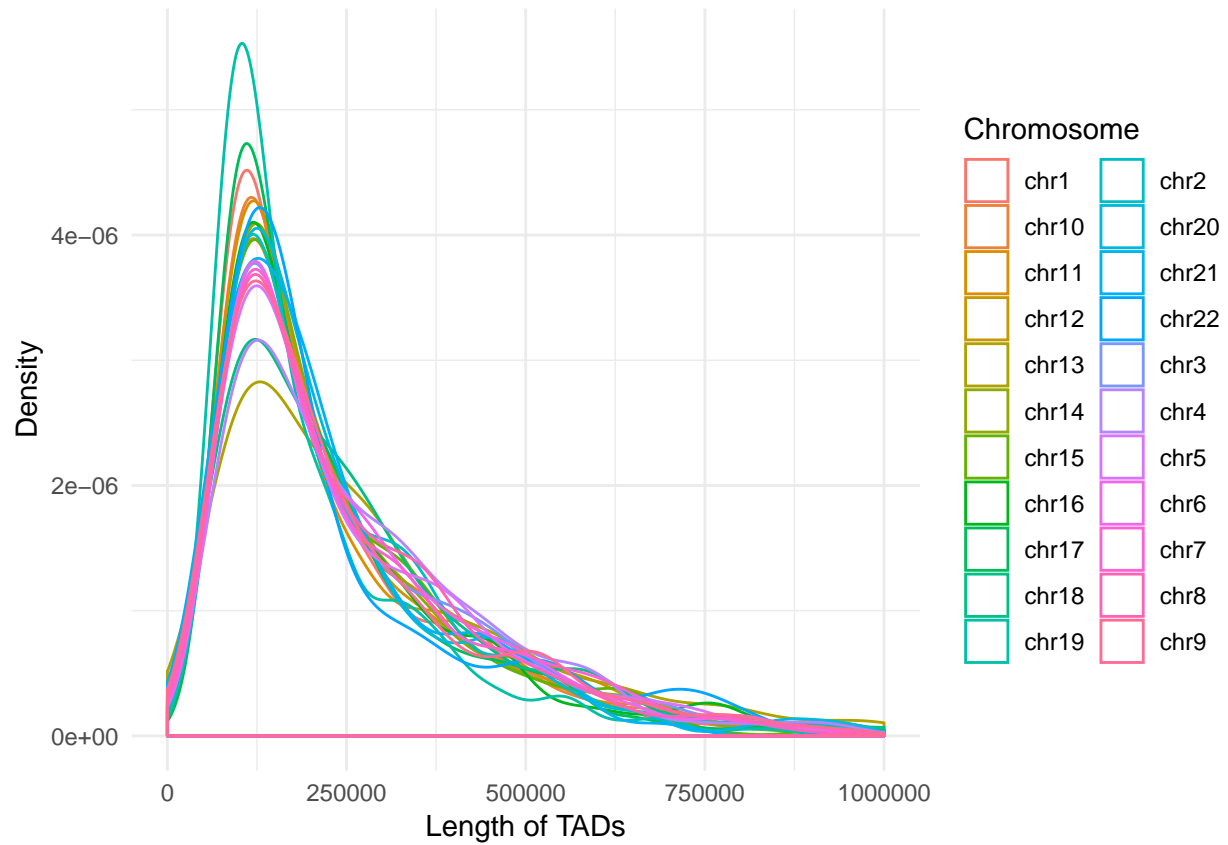
```r
#log distance
ggplot() +
  geom_density(data=domains, aes(x=logdist, group=Chromosome, color=Chromosome)) +
  xlab("Length of TADs") +
  ylab("Density")+
  theme_minimal()
```

Summaries of TAD distances by chromosome

```
dist <- tapply(domains$distance, domains$Chromosome, summary)


dist <- adply(dist, 1)
colnames(dist)[1] <- "Chromosome"
dist[,2:7] <- round(dist[,2:7], 2)
dist <- dist[order(as.numeric(substr(dist$Chromosome, 4, 5))),]


kable(dist)
```

	Chromosome	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1	chr1	65000	110000	170000	240281.2	305000	1535000
12	chr2	65000	115000	190000	269809.8	335000	2130000
16	chr3	65000	115000	190000	263021.2	355000	2345000
17	chr4	65000	125000	215000	279013.5	365000	1650000
18	chr5	65000	115000	190000	266491.0	355000	2690000
19	chr6	65000	115000	195000	266987.5	345000	1515000
20	chr7	70000	115000	192500	279410.5	345000	1990000
21	chr8	65000	115000	190000	262163.5	345000	2220000
22	chr9	65000	115000	185000	263395.7	335000	1885000
2	chr10	65000	115000	177500	248479.7	315000	1855000
3	chr11	65000	115000	170000	243016.0	322500	1520000
4	chr12	65000	115000	185000	250220.0	326250	2025000
5	chr13	65000	125000	225000	307134.4	395000	2280000
6	chr14	65000	117500	175000	240401.3	317500	1125000
7	chr15	65000	115000	175000	253300.6	325000	1455000
8	chr16	65000	115000	175000	257306.3	320000	1910000
9	chr17	65000	110000	170000	227652.1	275000	2000000
10	chr18	65000	115000	210000	285698.7	345000	2010000
11	chr19	65000	95000	140000	206719.2	245000	1545000
13	chr20	65000	120000	185000	259317.3	305000	1920000
14	chr21	65000	105000	185000	246302.1	293750	2025000
15	chr22	65000	115000	165000	248418.1	315000	1060000

Log 2 Distance

```r
logdist <- tapply(domains$logdist, domains$Chromosome, summary)

logdist <- adply(logdist, 1)
colnames(logdist)[1] <- "Chromosome"
logdist[,2:7] <- round(logdist[,2:7], 2)
logdist <- logdist[order(as.numeric(substr(logdist$Chromosome, 4, 5))),]

kable(logdist)
```

	Chromosome	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1	chr1	15.99	16.75	17.38	17.53	18.22	20.55
12	chr2	15.99	16.81	17.54	17.66	18.35	21.02
16	chr3	15.99	16.81	17.54	17.66	18.44	21.16
17	chr4	15.99	16.93	17.71	17.75	18.48	20.65
18	chr5	15.99	16.81	17.54	17.66	18.44	21.36
19	chr6	15.99	16.81	17.57	17.67	18.40	20.53
20	chr7	16.10	16.81	17.55	17.69	18.40	20.92
21	chr8	15.99	16.81	17.54	17.65	18.40	21.08
22	chr9	15.99	16.81	17.50	17.63	18.35	20.85
2	chr10	15.99	16.81	17.44	17.57	18.26	20.82
3	chr11	15.99	16.81	17.38	17.55	18.30	20.54
4	chr12	15.99	16.81	17.50	17.60	18.32	20.95
5	chr13	15.99	16.93	17.78	17.81	18.59	21.12
6	chr14	15.99	16.84	17.42	17.57	18.28	20.10
7	chr15	15.99	16.81	17.42	17.60	18.31	20.47
8	chr16	15.99	16.81	17.42	17.60	18.29	20.87
9	chr17	15.99	16.75	17.38	17.48	18.07	20.93
10	chr18	15.99	16.81	17.68	17.73	18.40	20.94
11	chr19	15.99	16.54	17.10	17.33	17.90	20.56
13	chr20	15.99	16.87	17.50	17.60	18.22	20.87
14	chr21	15.99	16.68	17.50	17.53	18.16	20.95
15	chr22	15.99	16.81	17.33	17.57	18.26	20.02