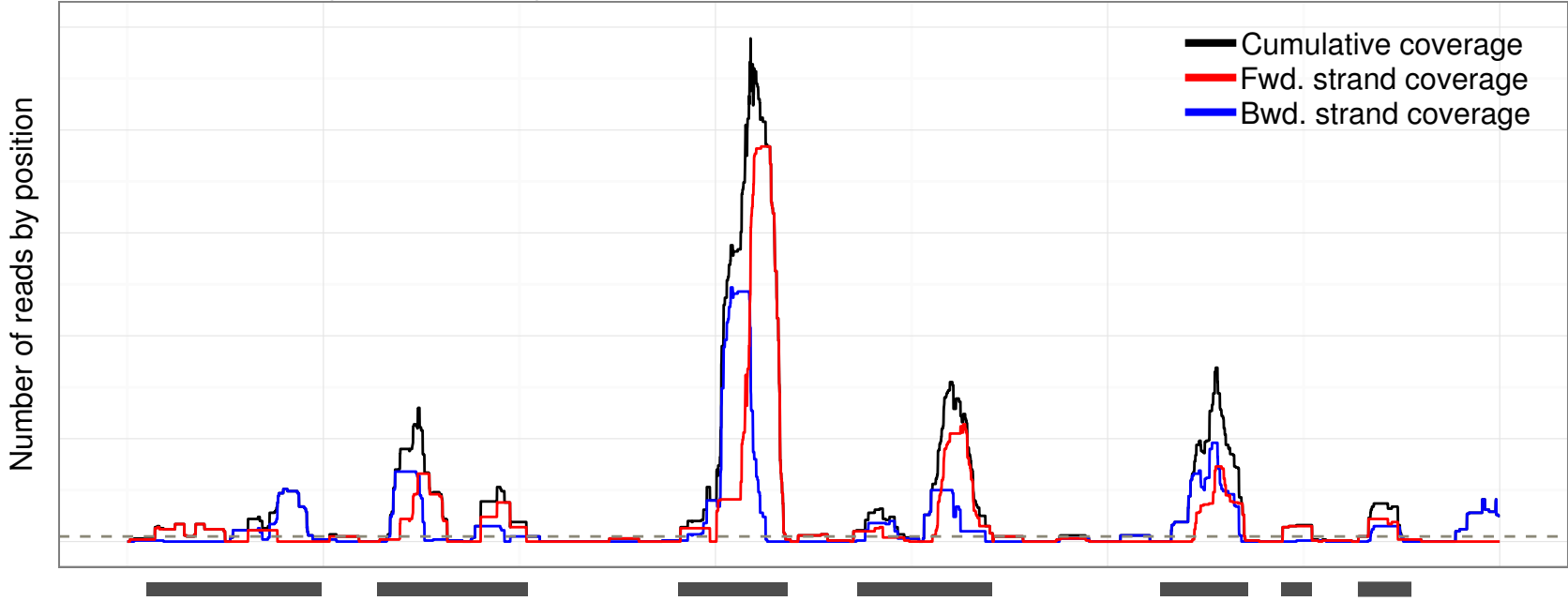# 1 - Partitions the genome and generate ChIP-exo islands.

Number of reads by position

- Cumulative coverage
- Fwd. strand coverage
- Bwd. strand coverage

2 - Calculate a vector of summary statistics for each island

R_1

$T(R_1)$

R_2

$T(R_2)$

R_3

$T(R_3)$

R_K

$T(R_K)$
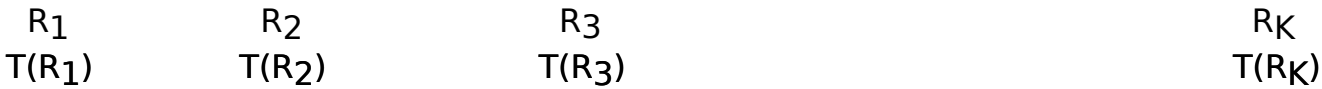
3 - Visualize all islands together:

A

count

10^0 10^2 10^4 10^6

Unique read coverage rate (URCR)

Average read coverage (ARC)

B

0.1 0.25 0.5 0.75 0.9

Fwd strand ratio (FSR)

min number of reads

C

both fwd bwd

Proportion of islands

min number of reads
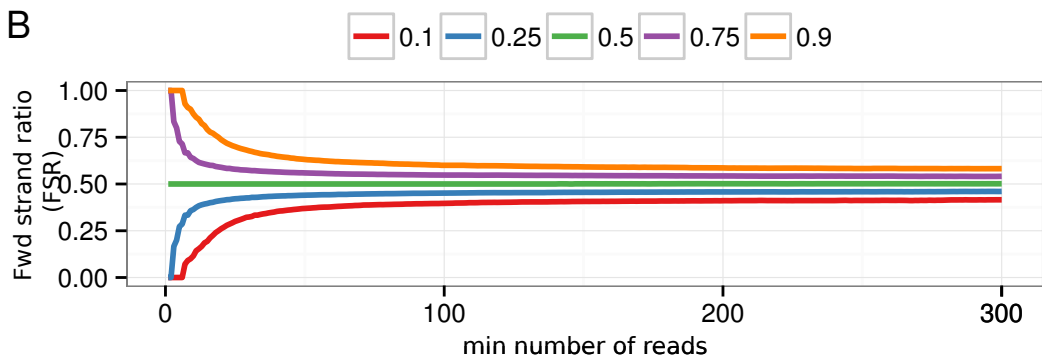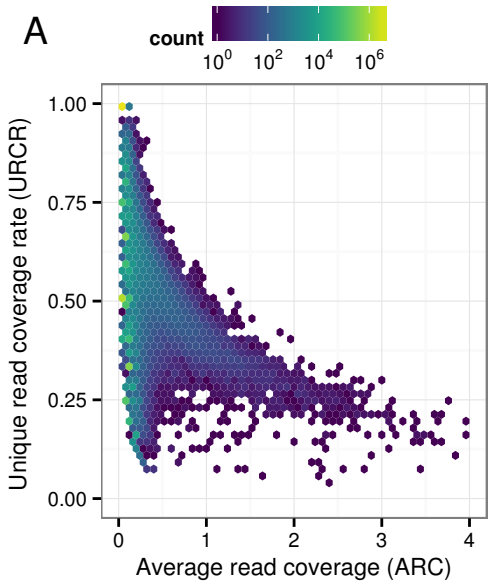
3A - ARC VS URCR - This plot gives a global view of the balance between library complexity and enrichment. There are two arms, one with low ARC, which corresponds to regions formed by few aligned positions, and the other were the URCR decreases as the ARC increases.
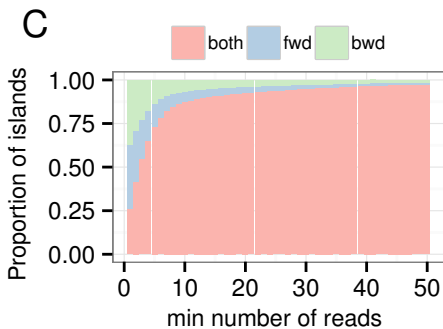
3B - Min depth VS FSR - This plot gives an idea of how quickly the distribution of the FSR approximates the median. In a good sample, the median is approx. 0.5, and the other quantiles reach that value quickly

3C - Min depth VS Proportion of Islands - This plot gives a more detailed view of the FSR. For islands with a low amount of reads is common to contain only reads from one strand. Hence, it compares the percentage of islands that contain at least one read of each strand vs the regions that consist of only reads with only one strand.