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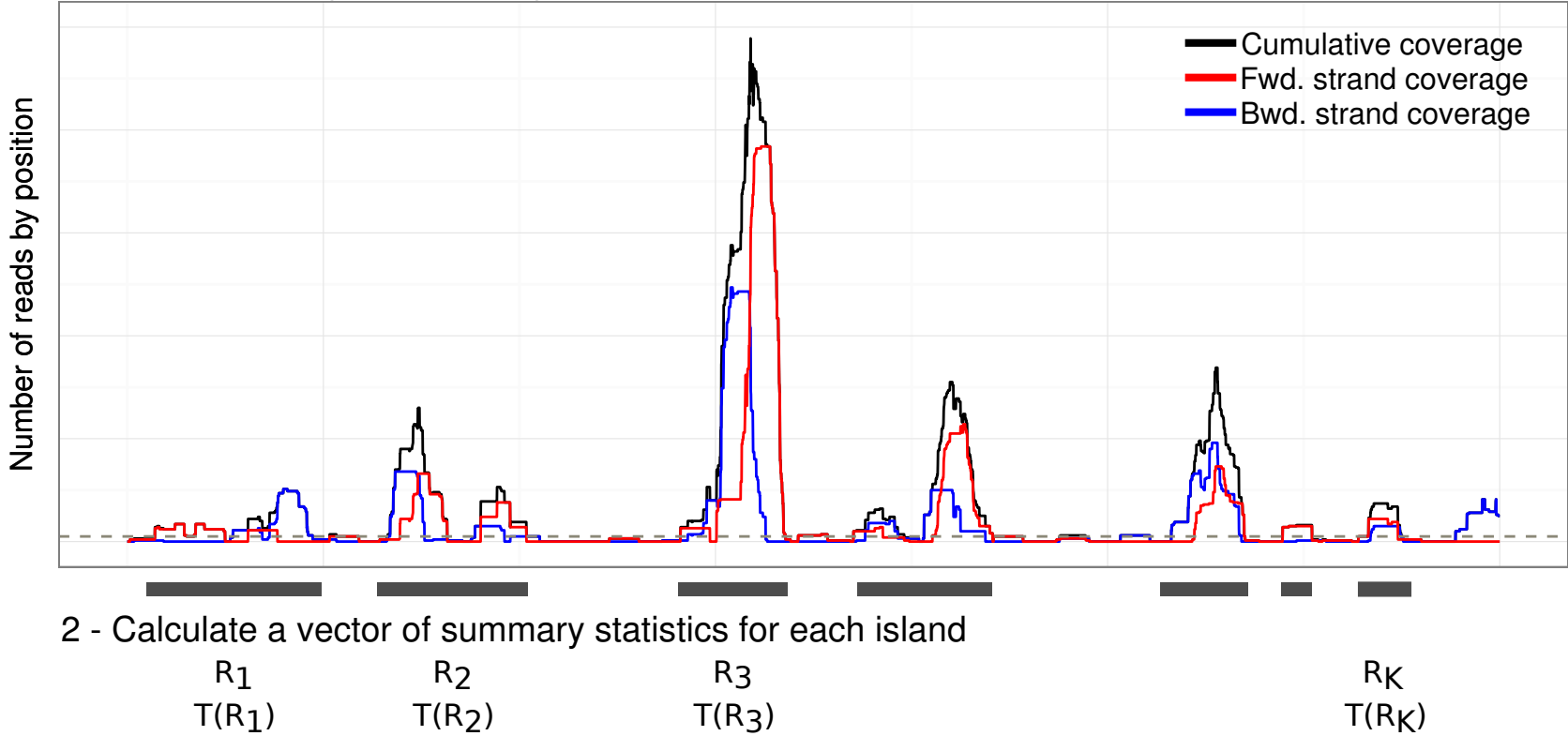

Legend:
- Cumulative coverage (black)
- Fwd. strand coverage (red)
- Bwd. strand coverage (blue)

y-axis: Number of reads by position

2 - Calculate a vector of summary statistics for each island

R_1 R_2 R_3 R_K

$T(R_1)$ $T(R_2)$ $T(R_3)$ $T(R_K)$

3 - Visualize all islands together:

A

count: 10^0 10^2 10^4 10^6

x-axis: Average read coverage (ARC)
y-axis: Unique read coverage rate (URCR)

B

Legend: 0.1, 0.25, 0.5, 0.75, 0.9

x-axis: min number of reads
y-axis: Fwd strand ratio (FSR)

C

Legend: both, fwd, bwd

x-axis: min number of reads
y-axis: Proportion of islands
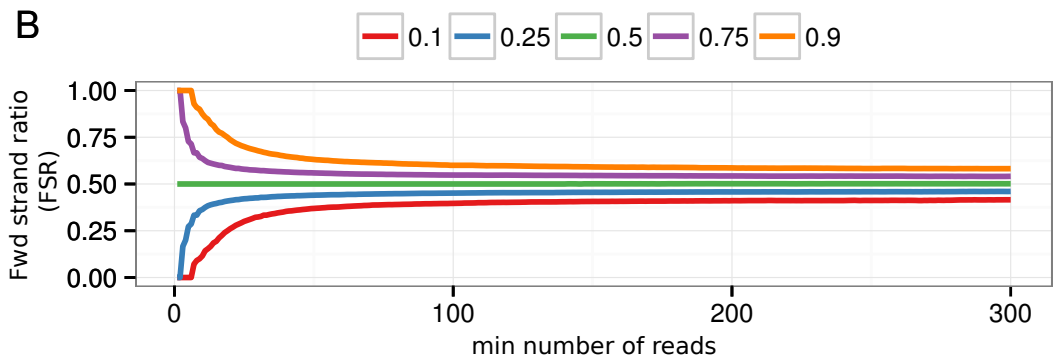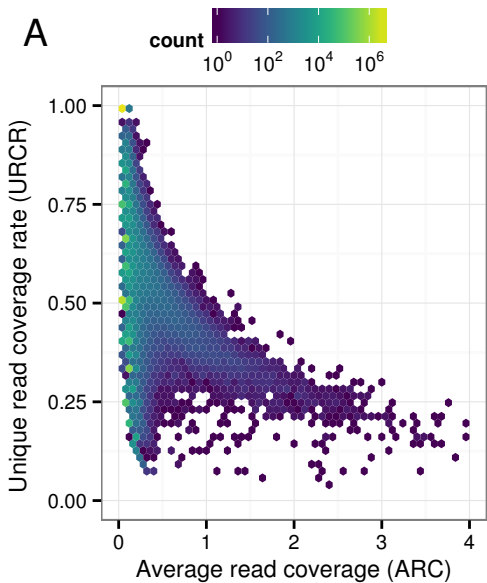
3A - ARC vs. URCR - This plot presents a global view of the balance between library complexity and enrichment. There are two arms, one with low ARC, which corresponds to regions formed by few aligned positions , and the other where the URCR decreases as the ARC increases.

3B - Min depth vs. FSR - This plot depicts how quickly the distribution of the FSR approximates the median. In a high quality dataset sample, the median is around 0.5, and the other quantiles reach that value quickly.
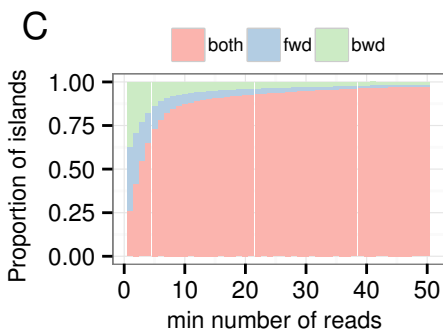
3C - Min depth vs. Proportion of Islands - This plot provides a more detailled view of the FSR. Islands with low depth then to have reads only from one strand. Hence, the plot compares the percentage of islands that contain at least one read of each strand vs. the regions that consist of only reads with only one strand.