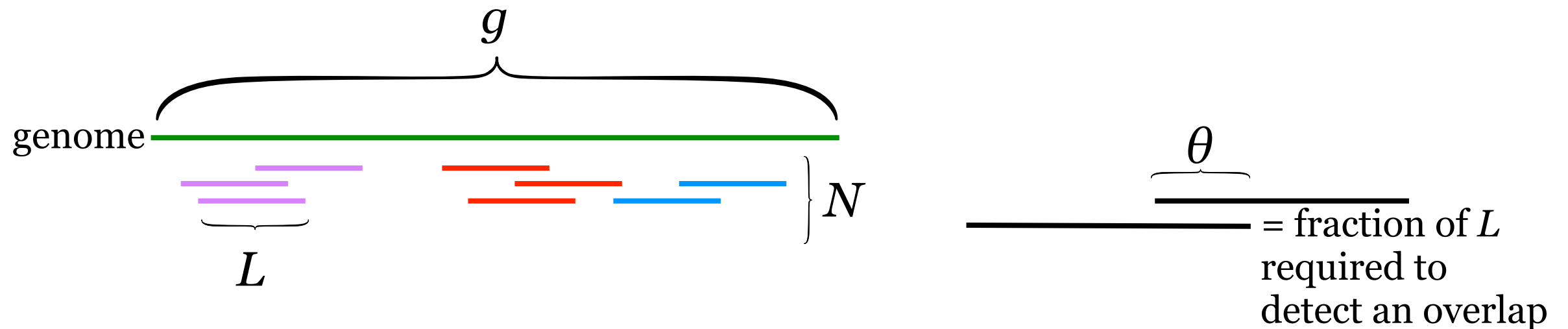


# Lander-Waterman Statistics

How many reads do we need to be sure we cover the whole genome?



An **island** is a contiguous group of reads that are connected by overlaps of length  $\geq \theta L$ .  
(Various colors above)

Want: Expression for expected # of islands given  $N, g, L, \theta$ .