**Written by Simon Zhang**

**Prokaryotic DNA replication overview**

Prokaryotic cells generally have circular, double-stranded DNA. For this DNA to replicate, there needs to be an origin of replication where the double-stranded DNA separates, creating a separation fork. Since the DNA’s two strands are complementary, one runs in the 5’-3’ direction while the other runs in the 3’-5’ direction. As the DNA unwinds in both directions, creating two replication forks, four half-strands are created, with two (one opens up in the 5’-3’ direction, the other opens up in the 3’-5’ direction) on each DNA strand.

**Reverse Half-Strand/Leading Strand**

For each DNA strand, there will be a half-strand opening towards the 5’ end of the DNA strand relative to oriC. This is known as the reverse half-strand due to the 3’-5’ direction. On this reverse half-strand, new DNA is created in the complementary 5’-3’ direction. Since DNA polymerase can only traverse in the reverse 3’-5’ direction (on the original DNA strand), DNA can be replicated in the forward direction for the reverse half-strand, therefore this strand is also known as the leading strand.

**Forward Half-Strand/Lagging Strand**

For each DNA strand, there will also be a half-strand opening towards the 3’ end, relative to oriC. This is known as the forward half-strand due to the 5’-3’ direction. On this forward half-strand, new DNA is created in the complementary 3’-5’ direction. However, as DNA polymerase cannot move in the 5’-3’ direction on the original DNA half-strand, this half-strand has to replicate DNA in the direction opposite to the way this DNA strand is being opened up. More specifically, this forward half-strand has to open up and make space for about 2000 nucleotides, before the DNA polymerase can attach itself towards the end of the nucleotide sequence, and replicate DNA backwards, towards the oriC. As a result, this replication is not done continuously as the strand opens up; these separate replication segments are named Okazaki fragments. Due to these characteristics, the strand is known as the lagging strand.

**Parent-Daughter Strand Relationship**

A parent forward half-strand synthesizes a daughter reverse half-strand; a parent reverse half-strand synthesizes a daughter forward half-strand. This convention makes sense once we consider the next cycle of replication.

**Inferring oriC based on empirical evidence and deamination**

One way of determining the location of oriC is through the GC skew. Empirical evidence shows that reverse half-strands have more C nucleotides than their forward counterparts, while forward half-strands have more G nucleotide than their reverse counterparts. At the same time, we know that due to a process called deamination, C nucleotides would mutate into T. During replication, mismatched T-G pairs are formed which eventually corrects to T-A. This process is 100x more likely to occur on single-stranded DNA. Combining these observations, we can infer that the oriC, which marks the region where the forward strand spends a long time single-stranded, lies in a region with high C content and low G content. This is because the oriC being in this region would account for (daughter) reverse strands not having a lot of G, due to mutation of C (from parent forward strand, which was supposed to pair with a G) to T and the subsequent correction of T-G into T-A. As a result of this process, the next generation’s daughter forward half-strand would have less C, due to its parent reverse half-strand having less G, which is also in correspondence with the empirical evidence. The GC skew can thus be used to find the oriC.