**Answers to Questions,
Chapter 5
DNA Replication I**

Answer 5.1. The production of LL shows that replication is *not* random.

Answer 5.2. In contrast to the replication eyes, the two new strands are not synthesized simultaneously at the replication fork in D loop replication.

Answer 5.3. In an neutral sucrose gradient, the two strands of the DNA duplex should stay together. Because the short Okazaki fragments should still be in duplex with the large parental DNA strands, the duplex would not separate from the bulk of the DNA. Therefore, by the model of semidiscontinuous synthesis shown in Fig. 5.7 you would not expect to see a slow-sedimenting peak of nascent DNA.

> [Surprisingly, when Okazaki et al. did this analysis, they still saw a slow-sedimenting peak. They also showed that this peak had single-stranded DNA in it, and proposed that this DNA was in "an unusual secondary structure." Perhaps this contained Okazaki fragments that were so short that they melted from the parental strands during isolation or centrifugation. They did this experiment to test a model in which both strands, parental and new, are in short pieces at the replication fork. Interested students may wish to read the original paper. Many subsequent papers have shown that the Okazaki fragments are made as intermediates in replication and are ligated together to form the lagging strand.]

Answer 5.4. A hypothetical head-growth mechanism for DNA synthesis would have the 5' end of the primer at the active site; this 5' end would have a triphosphate on the last nucletide added. The 3' hydroxyl on an incoming nucleotide could react with the α-phosphate of the 5' nucleotide by a nucleophilic attack. The β- and γ-phosphates would be liberated as pyrophosphate. All these steps are similar to those in the tail-growth mechanism at the 3' end, except that the nonactivated end of the incoming nucleotide initiates the reaction with the activated end of the growing chain. Chain synthesis would occur in a 3' to 5' direction. Note that if an incorrectly incorporated nucleotide were removed by a proofreading exonuclease (a 5' to 3' exonuclease in this hypothetical example), then the activated end of the chain would be removed, and synthesis would stop.

Answer 5.5. Removal of a nucleotide from the 3' end of the growing chain by a 3' to 5' exonuclease catalyzes the hydrolysis of the 3' nucleotide (adding a molecule of water across the bond that is broken), generating a nucleoside **mono**phosphate and a DNA chain shorter by one nucleotide. The reverse of the polymerization reaction is pyrophosphorolysis (adding a molecule of pyrophosphate across the bond that is broken), resulting in a nucleoside **tri**phosphate and a DNA chain shorter by one nucleotide.

Answer 5.6. As the replication fork moves 60,000 nucleotides per min, it produces both daughter strands at the same rate. Thus in 40 min, one replication fork replicates 60,000 bp

per min × 40 min = 2.4×10^6 bp. Dividing the size of the chromosome by this amount synthesized per fork gives 4.64×10^6 bp / 2.4×10^6 bp, or 1.93. Hence two replication forks are sufficient. For bidirectional replication, this requires only one origin, and indeed this is the case. The *E. coli* chromosome is replicated from one origin, called *oriC*.

Answer 5.7. PriA tracks alongthe single-stranded DNA in a 3' to 5' direction (relative to the single-stranded DNA). By moving in this direction after initially binding to the single-stranded DNA, it will encounter the duplex including molecule B, and then displace it by its helicase activity.

Answer 5.8

This bacterium replicates conservatively. Since no hybrid density DNA is formed, replication is not semi-conservative or distributive. The parental DNA remains heavy (HH) but is diluted out by the progeny DNA, which is light (LL).

Answer 5.9

The number of base pairs per helical turn for B-DNA is about 10. During DNA replication, the complementary strands of DNA must unwind completely to allow the synthesis of a new strand on each template.

The number of helical turns = number of base pairs/number of base pairs per helical turn.

Thus, $4.64 \times 10^6/10 = 4.64 \times 10^5$ turns must be unwound.

Note that this would generate an equal number of positive superhelical turns, if topoisomerases were not acting as a swivel during replication.

Answer 5.10

a) True
b) False, they are formed during synthesis of the lagging strand of DNA.
c) True
d) True

Answer 5.11 (a) At short pulse times (5 sec), the labeled thymidine or thymine appears exlusively in small DNA chains sedimenting at about 8 to 10S. As the pulse time increases, more of the labeled thymidine or thymine appears in large DNA, sedimenting at greated than 60S. This is consistent with discontinuous DNA synthesis of the lagging strand. The 8 to 10S DNA consists of Okazaki fragments, and the fast sedimenting, large DNA contains the growing chains after the Okazaki fragments have been joined to them.

Answer 5. 11 (b) The nascent DNA chain grows in a 5' to 3' direction. Because completed Okazaki fragments (short nascent chains) were isolated before the analysis, the labeled nucleotides incorporated at the earliest times had to be added as part of the process of completing the molecule. That is, the earliest-incorporated nucleotides are added to the part of the DNA synthesized last. The experimental results show that the 3' end is the portion synthesized as the molecule is completed. During longer labeling

periods, labeled nucleotides can be incorporated during initiation of the short nascent chain as well as the during the elongation and termination. Since the 5' end was labeled only during longer pulses, it must be the part synthesized first. Thus the direction of chain growth is 5' to 3.

Answer 5.12   In a pulse-chase experiment, the initial pulse labeling is stopped by adding a large excess of unlabeled precursor molecules, in this case unlabeled thymidine. Synthesis continues during the chase, but only a small portion of the new molecule being made (in this case DNA) was labeled during the pulse. To examine the fate of the Okazaki fragments, one could label DNA in growing cells for about 10 sec with [3H] thymidine, then dilute the culture into media with a large excess of unlabeled thymidine, which begins the chase. Samples of the culture are removed at a series of times during the chase, DNA is isolated from the bacteria, denatured and separated on a denaturing sucrose gradient. At the beginning of the chase, some of the labeled DNA should be slowly sedimenting; these are the new Okazaki fragments. Although additional Okazaki fragments are made during the chase, they will not be labeled (after the unlabeled thymidine swamps out the labeled thymidine). As the chase progresses, the labeled Okazaki fragments should be joined and added to previously synthesized lagging strand DNA. Hence the labeled DNA should become progressively larger sediment faster over the course of the chase.

Answer 5.13

The leading strand of newly replicated DNA is produced by continuous replication of the DNA template strand in the 5' to 3' direction at the replication fork. The lagging strand is synthesized in the form of Okazaki fragments, which are then spliced together. Common requirements for synthesis of both strands include the precursors (dATP, dGTP, dCTP, and dTTP are the source of nucleotides in the new DNA strand) a template DNA strand and a priming DNA strand. Enzymes and cofactors required for synthesis of both strands are:

• DNA *helicase,* which unwinds short segments of the DNA helix just ahead of the replicating fork; it requires ATP.
•*Single-strand DNA-binding proteins*, which bind tightly to the separated strands to prevent base pairing while the templates are being replicated.
•*DNA gyrase,* a topoisomerase, which permits swiveling of the DNA, thereby relieving the superhelical tension that would otherwise accumulate from the unwinding of the strands at the replication fork; it requires ATP.
•*DNA polymerase III,* which carries out the elongation of the leading strand by addition of nucleotide units; the cofactors Mg^{2+} and Zn^{2+} are required.
•*Pyrophosphatase,* which hydrolyzes the pyrophosphate released as each new nucleotide unit is added, thereby "pulling" the reaction in the forward direction.

Discontinuous synthesis at the replication fork has several additional requirements, all involved in synthesizing and then joining the short Okazaki fragments. Additional

precursors needed are UTP, ATP, CTP, and GTP, which are required for formation of the RNA primer that starts of each Okazaki fragment. Additional enzymes needed are:

- *Primase,* which constructs a short RNA primer, complementary to the DNA template, to initiate the Okazaki fragment. It functions with a complex primosome. Assembly of the primosome <u>requires ATP</u>, and movement of the primosome <u>requires ATP</u>.
- *DNA polymerase I,* which removes the RNA primer (exonuclease activity), replacing each NMP unit with a dNMP unit (polymerase activity); <u>cofactors required are Zn^{2+} and Mg^{2+}</u>.
- *DNA ligase,* which carries out the final step of splicing the new fragment to the lagging strand (the *E. coli* enzyme uses <u>NAD^+</u> as energy source).

Answer 5.14

a) The α subunit catalyzes 5'to 3' polymerization of new DNA, but it is most active within the catalytic core (αεθ).
b) The ε subunit as the proofreading function, but it is most active within the catalytic core (αεθ)
c) The τ subunit dimerizes the two catalytic cores.
d) The β subunit (as a dimer) forms the clamp that is thought to account for its high processivity.
e) The γ complex loads and unloads the sliding clamp.

Answer 5.15

The primosome contains PriA, PriB, PriC, DnaB, DnaT and primase (DnaG). It does not contain DnaC, but this protein is needed to form the primosome. A hexamer of DnaC forms a complex with a hexamer of DnaB, which is the complex needed, with the help of DnaT, to deliver DnaB to the pre-priming complex. The primosome synthesizes a short oligoribonucleotide, and it can include some deoxyribonucleotides in this primer. The primosome contains two different helicases (DnaB and PriA), each of which can move along single-stranded DNA in different directions. Movement in both directions has been observed *in vitro*. A model to accommodate this posits that the replication machinery is stationary, and the helicases with opposite polarity of movement serve to pull the template for lagging strand synthesis into a loop at the replication fork (Fig. 5.26)
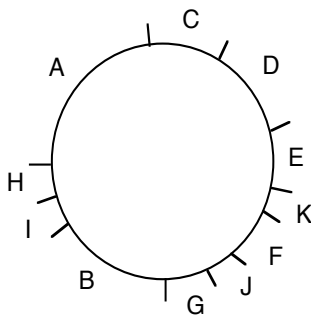
Answer 5.16

Polymerase δ is required for both leading and lagging strand synthesis when assayed in cell-free systems that reconstitute complete replication of templates containing viral origins of replication. Polymerase ε may be used for lagging strand synthesis *in vivo*. Polymerase α has been implicated in primer formation.

Answers to Questions
CHAPTER 6
DNA REPLICATION II:

Answer 6.1. 2,500 origins would be required for the haploid genome. Each bidirectional origin generates two replication forks which move 2,000 bp per min. Thus in 5 hr (which is 300 min), each fork moves 2,000 bp min^{-1} x 300 min = 600,000 bp. For the two forks per origin, this is 1.2 x 10^6 bp. In order to replicate the haploid genome, one needs 3 x 10^9 bp/ 1.2 x 10^6 bp ori^{-1}, or 2,500 origins.

Answer 6.2. A unidirectional mode of replication would show a monophasic gradient of label, highest at the terminus and lowest at the origin, and decreasing continously around the circle between these two sites.

Answer 6.3. If the bidirectional origins were in fragments E and H, these fragments would be labeled last in the pulse-labeleing experiment. Assuming equal elongation rates for all four replication forks, the termini would be half-way between E and H on both halves of the SV40 molecule, i.e. fragment C for the "top" half in the figure below and roughly the junction between B and G. These would label first in the pulse-labeleing experiment, and fragments between the termini and origin would have progressivly less label. For instance, at early times, the amount of labeling would go in the order C>D>E for the upper-right quadrant in the map below, and G>J>F>K>E for the lower-right quadrant in the map below.

Physical map of SV40 DNA fragments produced by cleavage with *H. influenza* restriction endonucleases

Answer 6.4. Fragment 3, with a bubble arc, has an origin, and fragment 5, with a double-Y arc, has a terminus, as diagrammed below. The other fragments have Y arcs, indicative of replication forks moving through them. Fork movement in fragment 4 is from left to right, moving from an origin to a terminus. Fork movement in fragment 6 is from right to left, moving into a terminus from an origin not on the map. If you knew that these were bidirectional origins, then one could conclude that fork movement in fragments 1 and 2 are from right to left.

Restriction fragments:

origin   termination

Fork movement
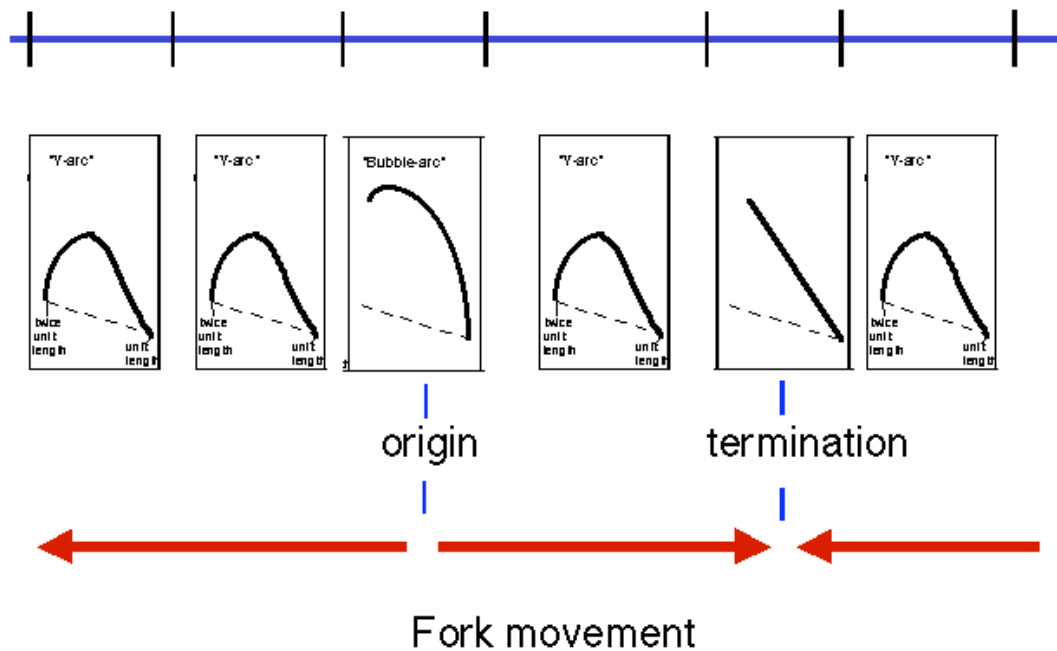
Answer 6.5 For bidirectional replication, the linear-equivalent length (determined from the first dimension) at the transition point from the bubble arc to the Y arc reflects the position of one replication fork at the time the other fork extends beyond a restriction site at the end of the fragment closest to the origin. For example, if the linear-equivalent length at the transition point is 1600 base pairs, and the unit length were 1000 bp, then the two forks traversed 600 bp before one reached the restriction site. Assuming and equal fork movement, then the origin is 600/2 = 300 bp from the end of the fragment.

Answer 6.6. The DNA has melted locally and the strands separated in the open complex. Thus a DNA fragment in the initial, closed complex will be cleaved by *Bgl*II but resistant to nuclease P1. DNA in the open complex will show the opposite effect, i.e. cleaved by nuclease P1 but resistant to *Bgl*II.

Answer 6.7 (a) Unwinding of 20 base pairs is a change in the twist (ΔT) of –2. If not counteracted, this is a change in the writhing of +2. Gyrase is a member of the Topoisomerase II family, which uses the energy of ATP to introduce negative superhelical turns, changing the linking number in steps of 2 (recall this from the section on supercoiling in the chapter on DNA structure). Thus one cycle of the reaction catalyzed by gyrase will counteract the unwinding of 20 bp.

Answer 6.7 (b) Helicases consume 2 ATP molecules for every base pair broken, so the unwinding of 20 bp (see 6.6.(a)), will require 40 ATPs. One cycle of gyrase action will be needed, using one ATP. Thus the total if 40+1= 41 ATP molecules.

Answer 6.8. We discussed a model for lagging strand synthesis in which the β subunit (sliding clamp) dissociates from Pol III core when it encounters the 5' end of an Okazaki fragment. If we apply that to this situation, each leading strand polymerase would stop synthesis and dissociate as soon as it encounters the 5' end of an Okazaki fragment; for a circular chromosome this would be the last Okazaki fragment synthesized by the fork moving in the opposite direction. Action by a 5' to 3' exonuclease and polymerase (e.g. DNA Pol I) to replace the RNA primer at the 5' end of the Okazaki fragment, followed by ligase, would join the products from the two replication forks.

Answer 6.9 (a), (b) DNA with *oriC* that was completely unmethylated at GATC motifs would not be competent for initiation if either of these hypotheses were correct.

Answer 6.10. Telomerase catalyzes the synthesis of one hexanucleotide repeating unit (GGGGTT in the case of *Tetrahymena*) and then shifts over to synthesize another repeating unit. If the enzyme dissociates from one telomere after each repeating unit, then its processivity is very low, i.e. 6 nucleotides. If it shifts over on the same telomere, then its processivity is higher. Note that the template RNA has at least two copies of the complement of the telomere repeating unit, so that there is still some overlap with the extending DNA strand when the enzyme shifts over to make a new repeating unit.

Answer 6.11. There is an average of six replication forks per chromosome. New replication must initiate every 20 min to sustain this rate of growth. If you picture a replicating DNA molecule that will complete synthesis in 20 min, it is already half-replicated, and each of the nascent daughter molecules has also initiated synthesis, for a total of 3 origins fired, and 6 replication forks for bidirectional replication.
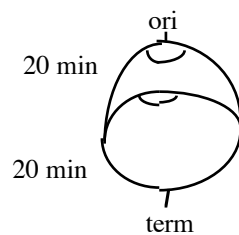
In the molecule illustrated above, the two older replication forks will meet and terminate in 20 min. This will leave 2 molecules in the cell (until the cell divides 20 min later). However, replication re-initiates every 20 min as well, so each molecule will still have 6 replication forks.

Answer 6.12 Replication is regulated primarily at initiation.

Answer 6.13 a)   Fragment B has the origin, and E has the terminus.
       b)   Replication is bi-directional, from an origin in fragment B.

Answer 6.14 Fragment Q has an origin and fragment P is replicated from that origin.
Note that the pattern for Q is a "bubble arc" and that for P is a "Y arc".

Answer 6.15 a. Fragments K and/or L contain the origin.

       b. Bi-directional.

       c. The leading strand extends away from the origin, and in the case of the bi-directional replication, the leading strands will extend divergently from the origin. Since only the leading strand is being synthesized and labeled, the hybridization pattern indicates that the bottom strand is made continuously beginning with fragment K (hybridizes to the top strand), and the top strand is made continuously beginning with fragment L (hybridizes to bottom strand).   Thus a bi-directional origin must exist around the junction between K and L. You cannot map unidirectional origins by this technique - can you see why?

       d. The replication fork moves from right to left through fragments A through K. The top, or nontemplate, strand hybridizes, which tells you that the leading strand is the bottom strand, whose 5' to 3' orientation is right to left.

       e. The replication fork moves from left to right through fragments L and M. The bottom, or template, strand hybridizes, which tells you that the leading strand is the top strand, whose 5' to 3' orientation is left to right.

       f. Any enzyme that is specifically involved in lagging strand synthesis is a candidate, e.g. primase or any component of the pre-priming complex (homologs to DnaG, DnaB, DnaC, DnaT, PriA, PriB, and PriC). Perhaps ligase or DNA polymerase I "homolog" could also be considered. In fact, emetine is an inhibitor of protein synthesis. The fact that it also blocks lagging strand synthesis indicates that some component of the machinery that synthesizes the lagging strand requires constant protein synthesis, suggesting that some component is very unstable.
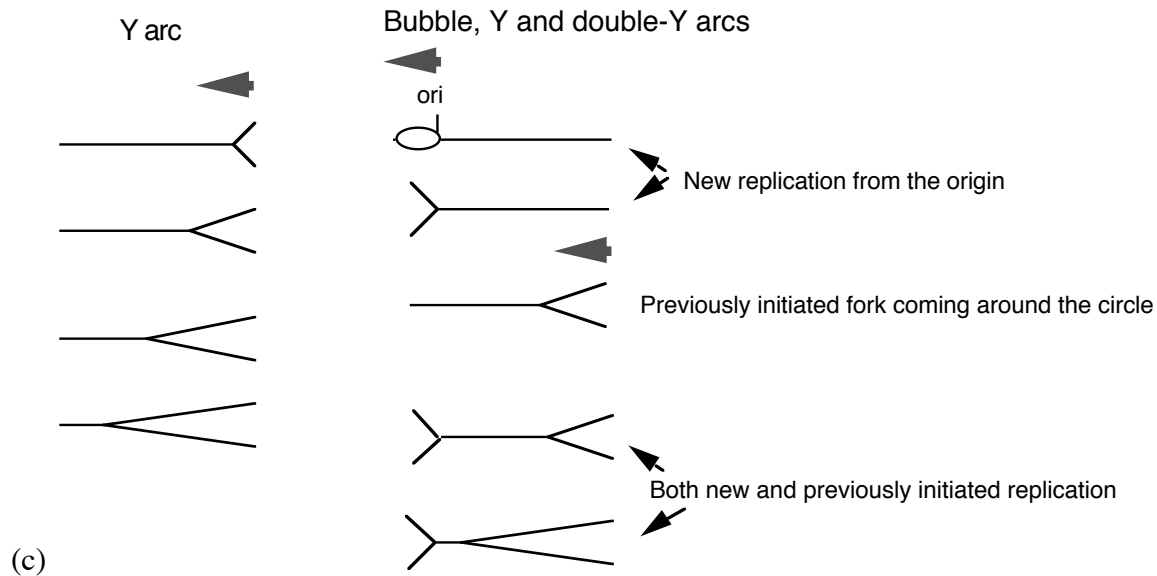
       e. M13 vectors. Placing the restriction fragment in one orientation will produce one strand in the viral progeny, whereas placement in the other orientation will produce the other strand.

Another good choice (and the one used in this paper) is a vector like pBluescript, which has promoters for RNA polymerases from bacteriophage T3 and T7 on the two sides of the insert, so transcription from one promoter generates an RNA with the "top strand" sequence, and transcription from the other promoter generates the "bottom strand" sequence.

Answer 6.16   (a) The origin is close to the C/D boundary, and the terminus is adjacent to it, in a clockwise direction.

(b) Replication goes in one direction, and that direction is counterclockwise

Y arc   Bubble, Y and double-Y arcs

ori

New replication from the origin

Previously initiated fork coming around the circle

(c)

Both new and previously initiated replication

The simple Y for fragment A is expected, since the replication fork should be elongating through this fragment. The fact that fragment C is also a simple Y-arc tells you that the origin is not in C, but the fact that it labels last in the pulse-labeling experiment tells you that it is very close to the origin. That suggests that the origin is in D, very close to the C/D boundary. Thus one can explain the rather complicated pattern in the 2-D gels for fragment D. Initiation at the origin will generate a bubble, but that will turn into a Y as soon as the replication fork passes the C/D boundary. However, previously initiated forks will enter fragment D from the other end (the D/E border). These can also generate Y-arcs, but there should be a lot of molecules that are being replicated both by new-initiated and previously initiated forks. Thus the double-Y pattern (the straight line on 2-D gels) will be generated.

(d)

Hhal fragment

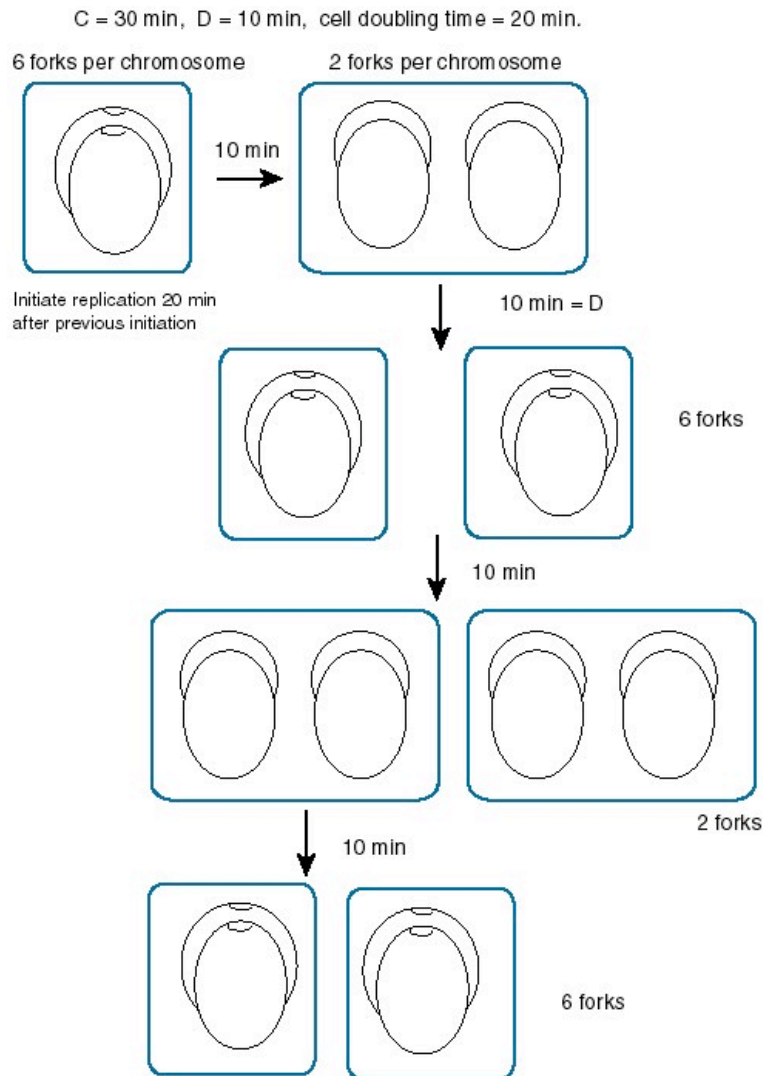|	A	B	C	D	E	F
Outer strand	▬	▬	▬	▬	▬	▬
Inner strand	▭	▭	▭	▭	▭	▭

Only the outer strands will hybridize to the leading strands, and no information will be gleaned about the origin or terminus. This assay only gives information when there is a transition of leading strand synthesis from one strand to the other, such as at a bi-directional origin.

Answer 6.17
    a)   True
    b)   False, primase (DnaG) catalyzes the synthesis of primers.
    c)   True
    d)   True

Answer 6.18. Since the cells are dividing every 20 min, replication initiates every 20 min. The time to replicate the chromosome is 30 min. Consider initiation event n. After 20 min, initiation n+1 occurs, giving 6 forks per chromosome: 4 from event n+1 (2 for each origin) and 2 from event n. Then after 10 min, the forks from event n terminate, and there are 2 DNA molecules in the cell, each with 2 replication forks (from event n+1). After 10 more min, the cell will divide, and initiation event n+2 occurs. So immediately after cell division, the chromosome has 6 forks again: 4 from event n+2 (2 for each origin) and 2 from event n+1. Thus the cells cycle between 6 forks and 2 forks, giving an average of 4 forks per chromosome.

See following diagram.

C = 30 min, D = 10 min, cell doubling time = 20 min.

6 forks per chromosome   2 forks per chromosome

10 min

Initiate replication 20 min
after previous initiation

10 min = D

6 forks

10 min

2 forks

10 min

6 forks

Answer 6.19 The fraction of cells showing the indicated gene with 4 dots (two doublets) is 0.5 for *GENEA* and 0.25 for *GENEB*. This is calculated by adding all the time in the cell cycle after replication and dividing by the total time of the cell cycle (24 hr in this case, i.e. $11 + 8 + 4 + 1 = 24$ hr). *GENEA* replicates 1 hr into S phase, so it will show as 2 doublets for 7 hr of S phase. The late replicating *GENEB* will show as 2 doublets for only 1 hr of S phase. Both will show as 2 doublets (4 dots) for all of G2 (4 hr), and for simplicity in arithmetic, we are assuming they will be 4 dots for all of M (1 hr). Thus *GENEA* will show as 4 dots for $7 + 4 + 1 = 12$ hr, or 0.5 of the 24 hr cell cycle. *GENEB* will show as 4 dots for $1 + 4 + 1 = 6$ hr, or 0.25 of the cell cycle. The fraction of cells will be same as the fraction of the cell cycle occupied, for an asynchronous population.

Consideration of the mitotic cells is an interesting complication. For much of mitosis, the nuclear envelope is disassembled, so technically there is no nucleus. Of course one will still get hybridization to the condensed mitotic chromosomes. As the chromosomes align during metaphase and separate during anaphase and telophase, the fluorescent dots will move close together and then move further apart, which will affect one's ability to distinguish 2 versus 4 dots. For this problem, we lumped all the mitotic cells into the "4 dot" category.

**Chapter 7
Mutation and Repair
Answers**

Answer 7.1. We will use arbitrary colors to help in following the fates of the incorporated dCTP and the A in the template strand. An A:T base pair is changed to an A:C in the initial product of replication. Upon another round of replication, an A:T will be at this position in one daughter molecule (the wildtype) and a G:C mutation will be at this position in the other daughter molecule. The A:T to G:C substitution (T to C on one strand, A to G on the other) is a *transition*.

Answer 7.2. A T:A base pair is changed to an T:C in the initial product of replication. Upon another round of replication, a T:A will be at this position in one daughter molecule (the wildtype) and a G:C mutation will be at this position in the other daughter molecule. The T:A to G:C substitution (A to C on one strand, T to G on the other) is a *transversion*.

Answer 7.3. First draw the base paired structure with the nucleoside deoxyguanidine in the *enol* tautomer and the nucleoside thymidine in the *keto* tautomer.

enol dG   dT
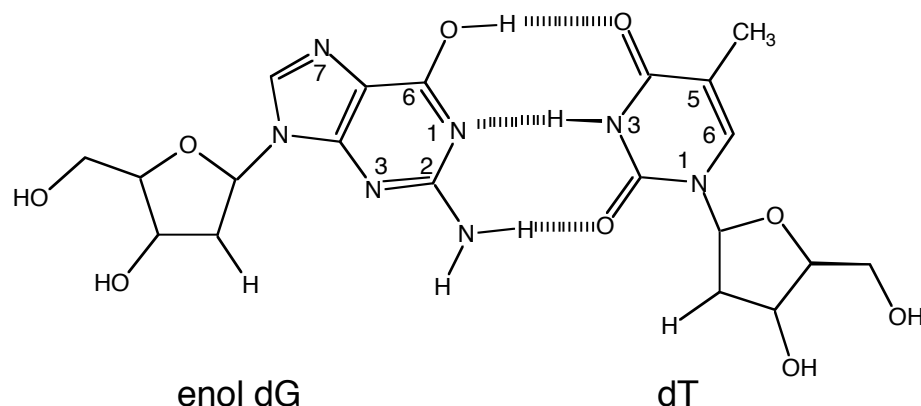
The oxygen attached to position 6 of guanine (O^6) is a hydroxyl in the *enol* tautomer, and the nitrogen at position 1 (N1) is fully bonded to the carbons on either side, so it has no hydrogen. Thus the O^6 hydroxyl is a hydrogen bond donor and the N1 imine is a hydrogen bond acceptor. The amino group bonded to position 2 is unchanged by the tautomerization, and it continues to serve as a hydrogen bond donor. The two keto groups on *keto* thymidine are hydrogen bond acceptors, one from the O^6 hydroxyl of *enol* guanidine and one from the N3 amino group of *enol* guanidine. The N3 amino group of thymidine is a hydrogen bond donor to the N1 imino group of *enol* guanidine.

Answer 7.4. The distance between phosphodiester backbones of the complementary strands of B form DNA is sufficient to accommodate a purine base and a pyrimidine base in the *anti* conformations. The purine base is larger than the pyrimidine base, and two of them in the *anti* conformation cannot be accomodated (without changing the distance between phosphodiester

backbones). By swinging the purine base back over the deoxyribose (i.e. the *syn* conformation), there is more room for the second purine base. Since pyrimidine bases are smaller than purine bases, two pyrimidine bases can fit between the two phosphodiester backbones without a shift from *anti* to *syn*.

Answer 7.5. Both hypoxanthine and xanthine have a keto group at the number 6 position (the "top" of the 6-membered ring), hence they have hydrogen bond acceptors at this position. The nitrogen at position 1 is in the amino form for both, i.e. it is a hydrogen-bond donor. This is the configuration needed for base pairing with the pyrimidine base cytosine, with a hydrogen bond donor at the "top" of the ring (the amino group on position 4) and a hydrogen bond acceptor (an imino group at position 3) .

Answer 7.6. The 5-methyl CpG sites would be oxidized to TpG's as a result of the spontaneous oxidative deamination of C's and failure to repair them. Thus the CpG dinucleotides would be replaced by TpG. Note that if you are looking only at the sequence of one strand of the DNA, a former CpG can become either TpG (if the C on the strand you are considering is methylated) or CpA (if the C on the complementary strand is methylated). This has been observed after the inactivation of a pseudogene for alpha-globin and as repetitive elements such as *Alu* repeats in humans evolve after transposition.

Answer 7.7. To make the 10 base pairs per turn in B form DNA, each base is rotated 36° relative to the adjacent base (note that $10 \times 36^\circ = 360^\circ$, or one full turn). When adjacent pyrimidines are covalently linked by the cyclobutane or the 6-4 bond between the bases, the bases are not able to make the 36° rotation, resulting in a change in the helical structure.

Answer 7.8. An exonuclease requires a free end on linear DNA to cut, whereas an endonuclease cuts within a DNA molecule (and hence can use circular DNA as a substrate, whereas an exonuclease cannot). The excinuclease is an excision nuclease used to cut out a segment of single stranded DNA. It is an type of endonuclease, but it makes two nicks (i.e. a break in the phosphodiester backbone on one strand) on the same strand of DNA, and in a precise location, i.e. one on either side of the damage. Thus a helicase can unwind the DNA between the nicks and remove the damged segment.

Answer 7.9. The glycosylases are specific for particular kinds of damage, .e.g uracil-N-glycosylas acts only on uracils, methylpurine glycosylase acts only on methylated purines. All the glycosylases leave sugar without a base, i.e. an AP site. Hence the AP endonuclease, DNA polymerase I and DNA ligaseare used generally for all repair by this pathway.

Answer 7.10. The mismatch recognition (MutS) and activation of endonuclease (MutL) functions are conserved from bacteria to mammals. However, the enzyme that recognizes the

sequence distinctive for newly synthesized DNA (MutH endonuclease) is not conserved. Humans do not methylate DNA at GATC, and presumably some other modification is used to mark parental versus progeny DNA. One possibility is the methylation of CpG dinucleotides.

Answer 7.11. When translesion synthesis occurs, mutations are generated opposite lesions, so defects in translesion synthesis will reduce the number of mutations (i.e. nonmutable phenotype) but also decrease the ability of the cell to survive damage.

Answer 7.12 GGTTGTT, from deamination of the C to form U, which base pairs with A. During replication, a T will incorporate opposite the A to replace the original C. This answer is restricted to the strand as written, but C's on the complementary strand are also subject to deamination.
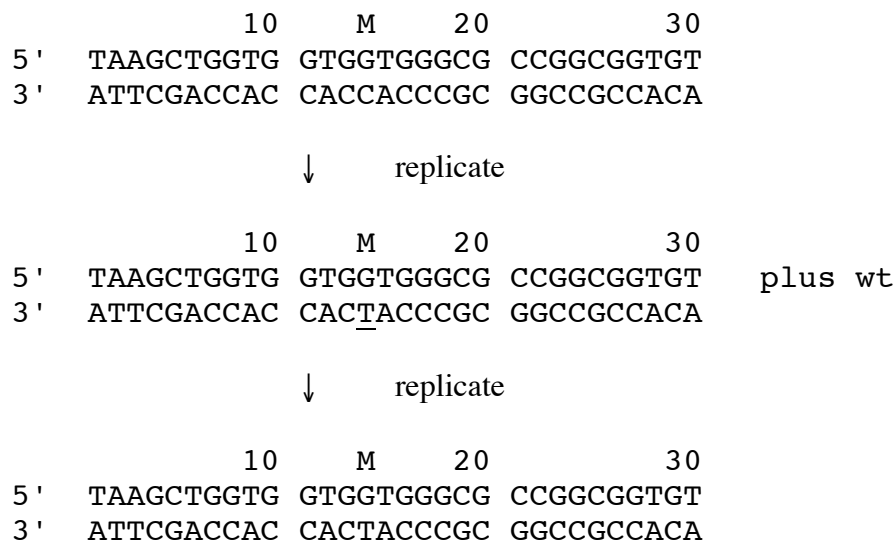
Answer 7.13
         a)   True
         b)   True
         c)   True

Answer 7.14
         a)   True
         b)   True
         c)   False, cleavage is on the non-methylated strand. If the methylated strand were cleaved and degraded, the information in the parental strand would be lost.
         d)   True

Answer 7.15 The G could be converted to O^6 methyl-G (denoted by the M below), which can pair with T. That leads to a GC (original base pair) to AT (mutant base pair) transition.

```
            10      M     20              30
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'   ATTCGACCAC CACCACCCGC GGCCGCCACA

                ↓        replicate

            10      M     20              30
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT     plus wt
3'   ATTCGACCAC CACTACCCGC GGCCGCCACA

                ↓        replicate

            10      M     20              30
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'   ATTCGACCAC CACTACCCGC GGCCGCCACA
```

                        plus

```
5'   TAAGCTGGTG GTGATGGGCG CCGGCGGTGT
3'   ATTCGACCAC CACTACCCGC GGCCGCCACA
```

Answer 7.16   The nitrous acid leads to oxidative deamination of the C, thus making a U. If this altered base is replicated, you get a CG (original base pair) to TA (mutant base pair) transition.

```
              10            20            30
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'   ATTCGACCAC CACCACCCGC GGCUGCCACA
```

                    ↓   replicate

```
5'   TAAGCTGGTG GTGGTGGGCG CCGACGGTGT        plus wt
3'   ATTCGACCAC CACCACCCGC GGCUGCCACA
```

                    ↓   replicate

```
5'   TAAGCTGGTG GTGGTGGGCG CCGACGGTGT
3'   ATTCGACCAC CACCACCCGC GGCTGCCACA
```

        plus

```
5'   TAAGCTGGTG GTGGTGGGCG CCGACGGTGT
3'   ATTCGACCAC CACCACCCGC GGCUGCCACA
```

Answer 7.17   The TT dinucleotide at positions 2 and 3 (bottom strand) would form dimers. Other pyrimidine dinucleotides can also form dimers, such as the CT at positions 5 and 6 (top strand) and any of the several CC dinucleotides.

```
              10            20            30
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'   ATTCGACCAC CACCACCCGC GGCCGCCACA
       \/
```

Answer 7.18   The initial damage, before replication, is O^6 methyl-G (denoted by the M below).

```
            10      M      20              30
5'    TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'    ATTCGACCAC CACCACCCGC GGCCGCCACA
```

This can be repaired by three different pathways:

(1) Direct reversal by O^6 methyl-guanine methyltransferase

(2) Correction by the UvrABC exinuclease in nucleotide excision repair.

Action of UvrABC: (UvrA)$_2$UvrB recognizes the methylated base, and UvrBC cleaves 8 nucleotides away on the 5' side and about 4 away nucleotides on the 3' side.

```
         cut           M   cut            30
5'    TAAGC TGGTG GTGGTGGG CG CCGGCGGTGT
3'    ATTCG-ACCAC CACCACCC-GC GGCCGCCACA
```

UvrD unwinds and liberates the damaged fragment:

```
5'    TAAGC                  CG CCGGCGGTGT
3'    ATTCGACCAC CACCACCCGC GGCCGCCACA
```

plus

```
         M
TGGTGGTGGTGGG  and the displaced UvrBC exinuclease.
```

Now DNA polymerase I can can fill in the gap and DNA ligase can seal the remaining nick to generate the wild-type sequence.

```
            10              20              30
5'    TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'    ATTCGACCAC CACCACCCGC GGCCGCCACA
```

(3) Excision of the methylguanine by methylpurine-N-glycosylase, leaving an apurinic site. AP endonuclease will then nick on the 5' side of the AP site, and DNA polymerase can fill in the sequence directed by the opposite strand as the template, thereby repairing the damage.

If the damage were corrected after replication, then

```
            10      M      20              30
5'    TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'    ATTCGACCAC CAC_TACCCGC GGCCGCCACA
```

would be "repaired" to make

```
5'   TAAGCTGGTG GTGATGGGCG CCGGCGGTGT   ; i.e. the mutation would remain.
3'   ATTCGACCAC CACTACCCGC GGCCGCCACA
```

Answer 7.19   (a) The C was oxidatively deaminated to U:

```
              10         20         30
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'   ATTCGACCAC CACCACCCGC GGCUGCCACA
```

The U will be removed by uracil-N-glycosylase, to leave an apyrimidinic site. In order to show the cleavage of the N-glycosidic bond more clearly, a layer of |||||| has been added to denote the sugar-phosphate backbone. The vertical lines are N-glycosidic bonds between the bases and the deoxyribose.

```
     |||||||||| |||||||||| ||||||||||
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'   ATTCGACCAC CACCACCCGC GGC GCCACA
     |||||||||| |||||||||| ||||||||||
```

Now the AP endonuclease sees the hole in the helix and cleaves the phosphodiester bond just to the 5' side, leaving a nick with a 3'-OH and a 5' phosphate:

```
     |||||||||| |||||||||| ||||||||||
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'   ATTCGACCAC CACCACCCGC GGC GCCACA
     |||||||||| |||||||||| ||||||||||
                               POH
```

DNA polymerase I can "nick translate" through the AP site and beyond, adding nucleotides as directed by the template (top) strand (indicated by the underlining.

```
     |||||||||| |||||||||| ||||||||||
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'   ATTCGACCAC CACCACCCGC GGCCGCCACA
     |||||||||| |||||||||| ||||||||||
                    POH
```

Ligase can now seal the nick.

(b) The thymine dimers could be reversed by photolyase or they could be excised by the
UvrABC exinuclease followed by the action of the UvrD helicase. Polymerase could fill in the
resultant gap with the correct sequence, followed by sealing with ligase. To illustrate that, lets
add another 10 bp to the left of the sequence (3' to the damage on the bottom strand).

```
                        10          20          30
5'  TGACGGAATA TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'  ACTGCCTTAT ATTCGACCAC CACCACCCGC GGCCGCCACA
                 \/
```

(UvrA)₂UvrB recognizes the damage, (UvrA)₂ dissociates and UvrC binds to UvrB at the
damaged site and cleaves 8 bp on the 5' side and about 4 bp on the 3' side. UvrD catalyzes the
breaking of the base pairs to "lift out" the damaged segment.

```
                        10          20          30
5'  TGACGGAATA TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'  ACTGCCT                ACCACCCGC GGCCGCCACA

  plus   TATATTCGACCACC
              \/
```

Polymerase (new DNA is underlined) plus ligase gives:

```
                        10          20          30
5'  TGACGGAATA TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT
3'  ACTGCCTTAT ATTCGACCAC CACCACCCGC GGCCGCCACA
```

Answer 7.20 The mismatch repair system uses the information about the methylation status of
GATC strings to determine which strand is parental and which was made by the most recent
round of replication. Thus in this situation, the top strand is the parental (presumably correct)
strand, since it has the methyl on the A in the GATC. The bottom strand presumably
incorporated an A erroneously at position 24.

```
            10          20          30 ...                    m
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT ...          GGACGGATCC
3'   ATTCGACCAC CACCACCCGC GGCAGCCACA ...          CCTGCCTAGG
```

MutS will recognize the GA mismatch, (MutL)₂ will bind to MutS and activate the endonuclease
MutH to cleave 5' to the G in the unmethylated GATC (bottom strand in this case). The
mismatch will be excised in a patch that starts at the GATC and extends past the mismatch.

```
            10          20          30 ...                    m
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT ...          GGACGGATCC
3'   ATTCGA                                                 G
```

DNA polymerase plus ligase will restore the wild-type sequence. New DNA is underlined.

```
          10         20         30 ...                 m
5'   TAAGCTGGTG GTGGTGGGCG CCGGCGGTGT ...        GGACGGATCC
3'   ATTCGACCAC CACCACCCGC GGCCGCCACA ...        CCTGCCTAGG
```

Note that if the A is not corrected, it will direct the incorporation of the tumorigenic T in the next round of replication.

Answer 7.21
The extracts of cells from FA complementation groups A and D can complement *in vitro* to restore the ability to join DNA ends.
The following are **not** supported by the data:
>   FA is a disease resulting from deficiencies in mismatch repair.
>   The extracts of cells from FA complementation groups A and D contain an inhibitor of normal DNA end-joining.

One specific possibility is that the genes defined by FA complementation groups encode a DNA ligase. However, direct experimental tests (not presented here) show no change in ligase activity compared to wildtype cells.

Answer 7.22. Specific sites on plasmid DNAs can be mutated by denaturing the plasmid and annealing with an oligonucleotide that has the desired nucleotide substitution. After this mutagenic oligonucleotide has annealed to its complementary segment on one strand of the parental plasmid (let's call it plus), it can serve as a primer for synthesis *in vitro* of the other DNA strand (minus), followed by ligation. If this heteroduplex plasmid (with the plus strand parental and the minus strand newly synthesized and containing the mutation) is transformed into *E. coli*, replication will make new plasmids containing either the wild type parental sequence or the mutated sequence.
        Increasing the frequency of plasmids containing the mutated sequence is desirable, and a mutant strain defective in *dut* and *ung* can be used to decrease the frequency of plasmids with the parental sequence. If the parental plasmid is grown in a *dut* -, *ung*- strain, it will incorporate U's in the DNA. The new DNA synthesized *in vitro* with the mutant oligonucleotide will have only T's, no U's. After transformation of the heteroduplex plasmid into a *ung*+ strain, the U's will be removed from the parental plus strand, leaving many AP sites, whereas the mutated minus strand will be intact. Replication of the plasmid is more efficient than repair of the AP sites, thus the replicative polymerase will preferentially use the mutated minus strand as the template, thereby increasing the frequency of the mutated plasmid.
        This strategy was developed by T.A. Kunkel (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. **Proc Natl Acad Sci U S A** 82:488-492.

Chapter 8
Recombination of DNA
Answers

Answer 8.1. If $A^-B^+C^-$ were the result of a double crossover, then one should also find the reciprocal partner $A^+B^-C^+$.

Answer 8.2. The chromosomes would simply be assorted with two to each haploid cell. At the end of meiosis I, chromosome A could assort together with A', B or B', and chromosome A' could end up with A, B, or B', and so forth. Since there is only one of each chromosome in the diploid cell, each chromosome cannot assort with itself. This leaves 12 combinations of two chromosomes, i.e. $2^4-4=16-4=12$. The pairs are A with A', A with B, A with B', A' with A, A' with B, A' with B', B with A, B with A', B with B', B' with A, B' with A' and B' with B. Of these 12 combinations, 8 of them have both an A-type and a B-type chromosome. This is the situation at the end of meiosis I. During meiosis II, the sister chromatids separate to individual cells, resulting in the haploid state. If this occurs normally, each sister chromatid goes to a different cell. In this case, both types of chromosomes are together in 2/3 of the haploid cells. Of course, it is the other 1/3 of the cells that are the problem, since they are missing one of the chromosomes.

Answer 8.3.
The phenotypes indicate the following genotypes.

Spore	*leu* allele	*Sm* allele	*ade8* allele
1	*leu+*	*Sm*R	*ade8+*
2	*leu+*	*Sm*R	*ade8+*
3	*leu+*	*Sm*R	*ade8-*
4	*leu+*	*Sm*S	*ade8-*
5	*leu-*	*Sm*S	*ade8+*
6	*leu-*	*Sm*S	*ade8+*
7	*leu-*	*Sm*S	*ade8-*
8	*leu-*	*Sm*S	*ade8-*

Note that spores 1 and 2 correspond to one parental chromosome, *leu+ Sm*R *ade8+*, and spores 7 and 8 correspond to the other parental chromosome *leu- Sm*S *ade8-*. Spores 3 and 4 are crossovers between genes *leu* and *ade*8, and spores 5 and 6 show the reciprocal arrangement. Thus spores 3, 4, 5, and 6 show that a recombination has occurred with exchange of flanking markers (the *leu* and *ade*8 genes in this case). The ratio of parental alleles is 4:4 for both these genes. In contrast, the *Sm* gene shows a 3:5 ratio of parental alleles, i.e. 3 *Sm*R alleles and 5 *Sm*S alleles. Spores 3 and 4 show different alleles for *Sm*. These spores are formed by the replication of one product of meiosis. The fact that they show different phenotypes for resistance and sensitivity to spectinomycin argues that the duplex DNA in the meiotic product was a heteroduplex of the two alleles. The *leu* and *ade*8 markers also showed recombination in these spores. Thus these data can be interpreted as a recombination between sister chromatids that left

a patch of heteroduplex in the *Sm* gene. The fact that the *Sm*S allele is seen for the other recombinant chromosomes in spores 5 and 6 can be interpreted as an instance of a heteroduplex being repaired, or the *Sm*R allele on these chromosomes was converted to the *Sm*S allele.

Answer 8.4. In the Holliday model, both duplexes exchange DNA strands equally, so the duplex with the initiating chi site would be equally likely to donate to or receive DNA from the other duplex. However, in the double-strand-break model, the initiating strand with the chi site would be the aggressor duplex, which receives DNA from the other duplex during gap repair. Thus preferential use of the duplex with a chi site for initiation for receiving genetic information would support the double-strand-break model.

Answer 8.5. The single-strand with the 3' end is the equivalent of the free 3' end generated by exonucleases in the double-strand-break model (second step in Fig. 8.9). The strand assimilation mediated by RecA-ATP is the strand invasion step. Further steps require DNA synthesis using the strand complementary to the invading single strand as a template, gap repair by DNA polymerase on the invading duplex and ligation to form Holliday junctions.

Answer 8.6. The extent of branch migration after formation of the Holliday junction is the principle determinant of the length of the heteroduplex formed.

Answer 8.7. Cleavage in the strands that were not involved in the original crossover will lead to recombination of flanking markers, whereas cleavage in the <u>same</u> strands that were involved in the original crossover will lead to <u>no</u> recombination of flanking markers.

Answer 8.8. Examination of interallelic recombination during spore formation in heterozygous *Ascocmycetes* (and other fungi) occasionally show an 3:5 ratio of the spores from each allele of the heterozygote, instead of the expected 4:4 ratio.
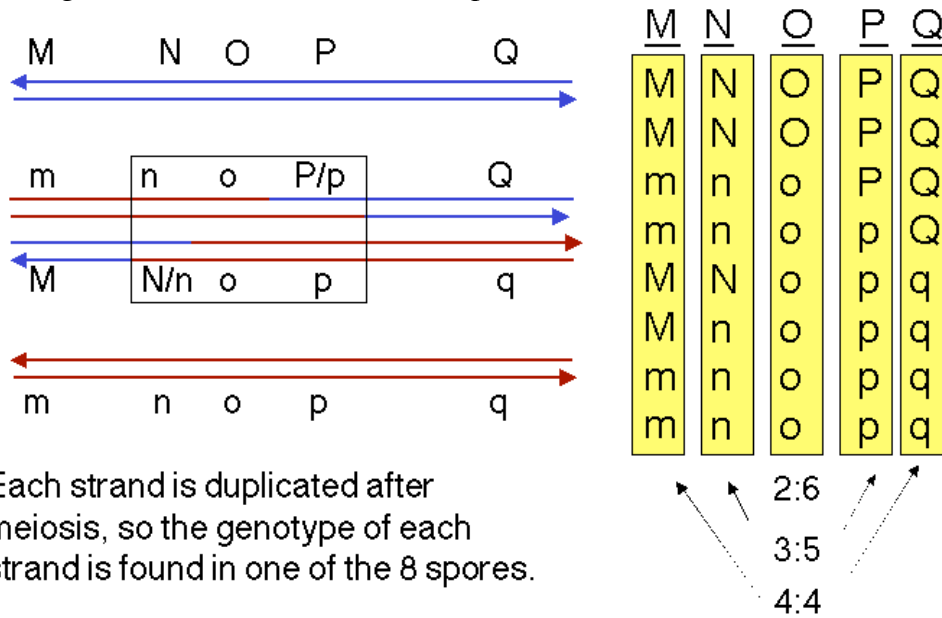
Answer 8.9. a. The products of the stated resolution are shown below.
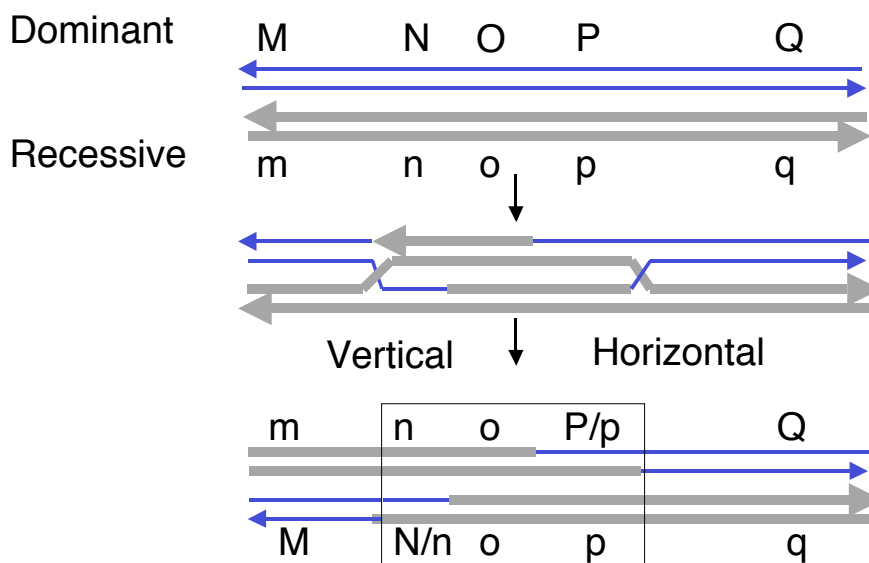
b. The genotype of the recombination products will be M__q / m__Q. In general, resolution of the two Holliday junctions in opposite senses (i.e. vertical-horizontal or horizontal-vertical) will result in recombination of flanking markers.

c. One should see 6:2 o:O, indicating that gene conversion has occurred. Examination of the resolved products in part a) shows that all are white (o) at the O locus, which will give 4 o spores. The sister chromatids that did not undergo recombination retain the parental genotypes, so one is o and the other is O, each of which contributes two daughters to the spores. Thus there will be 4 + 2 = 6 o spores and 2 O spores.

See the figure below, which shows the expectations for each locus.

**Answer 8.10.** The intermediate with 2 Holliday junctions is shown below.
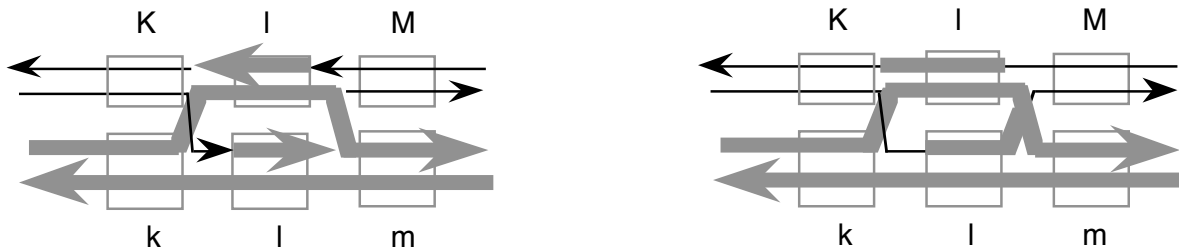
Answer 8.11. The L gene will have the l allele.

Answer 8.12. The K gene will be a heteroduplex on both duplexes, with one strand having the sequence of the K allele and the other strand having the sequence of the k allele.

Answer 8.13. The Holliday model predicts formation of a break in the linear chromosome, whereas the double-strand break model predicts that the gapped will be repaired, leading to insertion of the XY circle flanked by ABc and aBC.

Holliday model with reciprocal invasion of single strands:

Consider recombination initiated at a nick in white C and using the 5' end of the black strand in c.
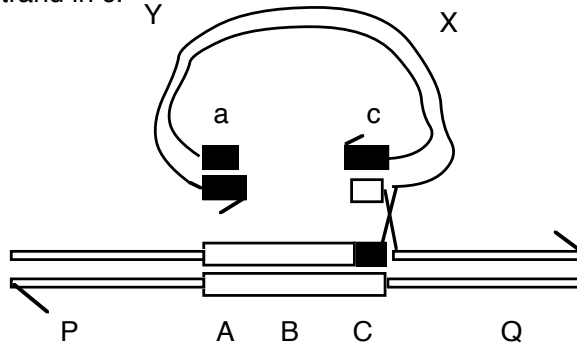
Branch migration cannot proceed into the region of non-homology (X vs. Q), so let's resolve the Holliday intermediate. Horizontal resolution leads to patches of heteroduplex with no repair of the gap, and vertical resolution leads to a break in the white chromosome.
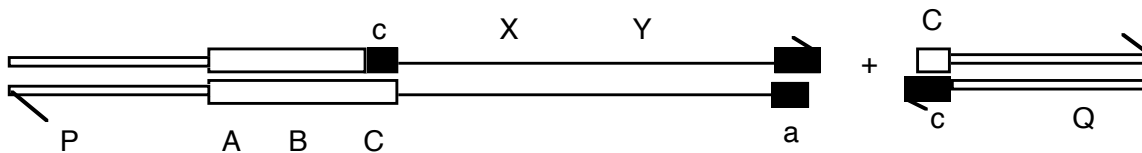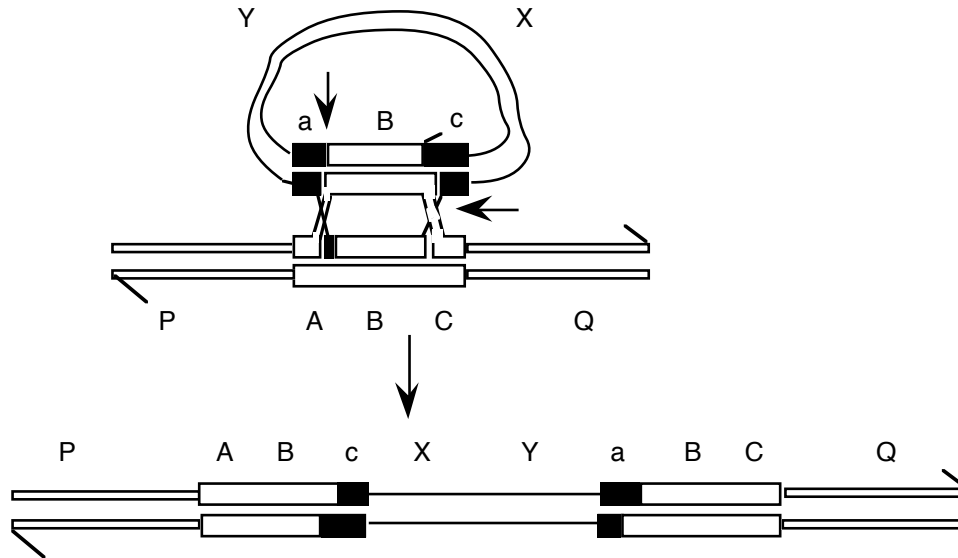

The double strand break model will repair the gap in the agressor (recipient) duplex. That was the observed result, giving strong support to this model.
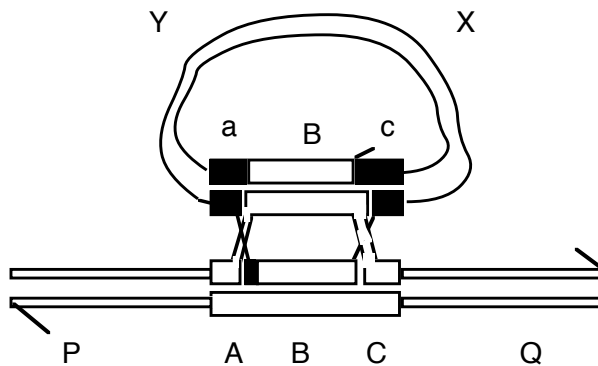
Resolution vertically at the left Holliday junction and horizontally at the right junction
leads to integration of XY flanked by intact aBc and ABC segments.

Answer 8.14. All these substrates fulfill the three requirements for strand exchange; there is a
single-stranded region, homology between the substrates and a free end in the region of
homology. The expected results from complete reaction are diagrammed and explained below.
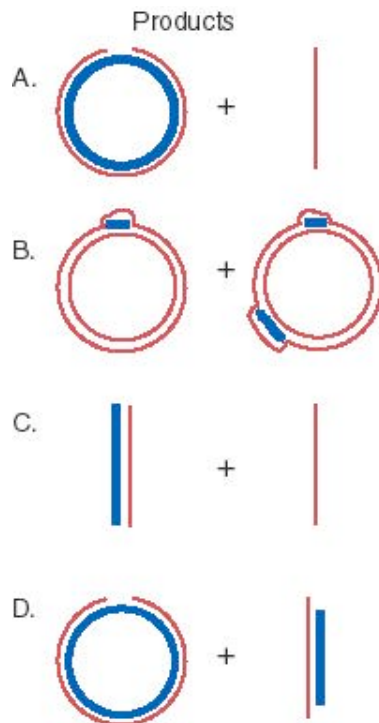
A. A heteroduplex circle with a nick and a linear fragment. The intact strand of the circle is from the invading single strand, and the other strand is its complement in the original linear duplex. The nick is shown exaggerated in the diagram so the break is obvious. However, it is only broken phsophodiester link; no nucleotides are missing since the circle and linear are the same lengths. This is identical to the assay explained in Fig. 8.16.

B. The short linear fragments are assimilated into the circle forming D-loops, with the segments of the circle identical to the invading fragments now displaced from the duplex.

C. A heteroduplex linear, with one strand from the invading single strand and its complement from the original duplex, and a displaced single strand.

D. A nicked heteroduplex circle (as in A) and a linear heteroduplex with one strand longer (from the initial linear duplex) and one strand shorter (the gapped strand from the original gapped circle).

CHAPTER 9
TRANSPOSITION
Answers

Answer 9.1. A classic example of partial dominance is the result of crossing a homozygous red-flowered plant with a homozygous white-flowered plant and obtaining pink-flowered plants in the progeny. In this case, the "white" allele contributes no pigment and the single "red" allele contributes half as much the color as two "red" alleles. This is observed in **all** the pigmented cells of the petal. In the sectored petals shown in Fig. 9.5, some of the cells are just as purple as those in the pure purple petals (panel A), but other cells are white. If the petal color in the wild flox were determined by a single gene with two alleles, it appears that the "purple" allele is expressed in some cells, whereas the "white" allele is expressed in others. Since all the cells start with the same genotype, something is affecting the expression or both alleles. This effect is seen randomly around the flower petal, but the sectoring suggests that once a change is made in one cell, it is inherited in progeny of that cell. One mechanism that can account for the results is the loss of a dominant allele, allowing the phenotype of the recessive allele to be seen, as was discussed for the effects of *Ds*-mediated chromosome breaks in maize. Other possibilities are epigenetic effects (i.e. not affecting the sequence of the DNA, but rather its ability to be expressed), such as silencing of an allele in some cells by methylation or formation of heterochromatin.

Answer 9.2. The phenotype would be variegating. The cells with an unbroken chromosome 9 would be colorless, nonshrunken and nonwaxy (*I Sh Bz Wx*), whereas those with a broken chromosome would reveal the recessive alleles *C* and *sh* from the other chromosome, making it colorless and shrunken. However, the *Wx* allele is still present, and these cells would be nonwaxy.

Answer 9.3. The composite transposon can move as either the original transposon or as the two IS elements with the other plasmid DNA between them. The inverted repeats will appear as inverted repeats for either part of the plasmid. The original transposon may move more frequently, since the characterized transposases have a preference for one end of the inverted repeat. However, the mobilization of the other part of the plasmid can be seen if it contains a distinctive selectable marker. These results have been obtained experimentally.

Answer 9.4. The new copy of the *Ty-1* element would have the same structure as the original copy, with the intron intact. DNA polymerase would copy the entire *Ty-1* element. If part of the *Ty-1* were deleted by errors in replication, there is no reason why they would occur exactly at splice junctions. The frequency of DNA-mediated transposition could be increased by increased transcription of the transposable element, since this frequency of transposition should be related to the amount and efficiency of transposase and resolvase. If these enzymes were encoded on the transposable element, the amount of these enzymes would be increased as transcription (and subsequent translation) increased.

Answer 9.5. The mutation is most likely the result of insertion of a transposon. The increase in size of the *EcoRI* fragment indicates that an insertion of 3 kb has occurred in the gene. Since it generates flanking direct repeats, it is likely to be a transposable element. The inverted repeats at the ends are characteristic of transposable elements that move via DNA intermediates. This is most likely a transposon, and not an insertion sequence, based on the size of the insert (3 kb) and the fact that the inserted DNA also confers resistance to an antibiotic.

You would expect to find a transposase and possibly a resolvase encoded in the transposon.

Answer 9.6. A cointegrate with the donor and recipient replicons fused, joined by two copies of the TE.

Answer 9.7. Excision of the TE from the donor replicon and movement of the TE to the recipient replicon, with duplication of the target site.

Answer 9.8. If each end of the transposon is ligated to the nick at the target on the same strand, then the process of replicative transposition will lead to a deletion. If each end of the transposon is ligated to the nick at the target on the opposite strand, then an inversion will be the result. This is diagrammed below.

Transposition to a 2nd site on the same DNA molecule
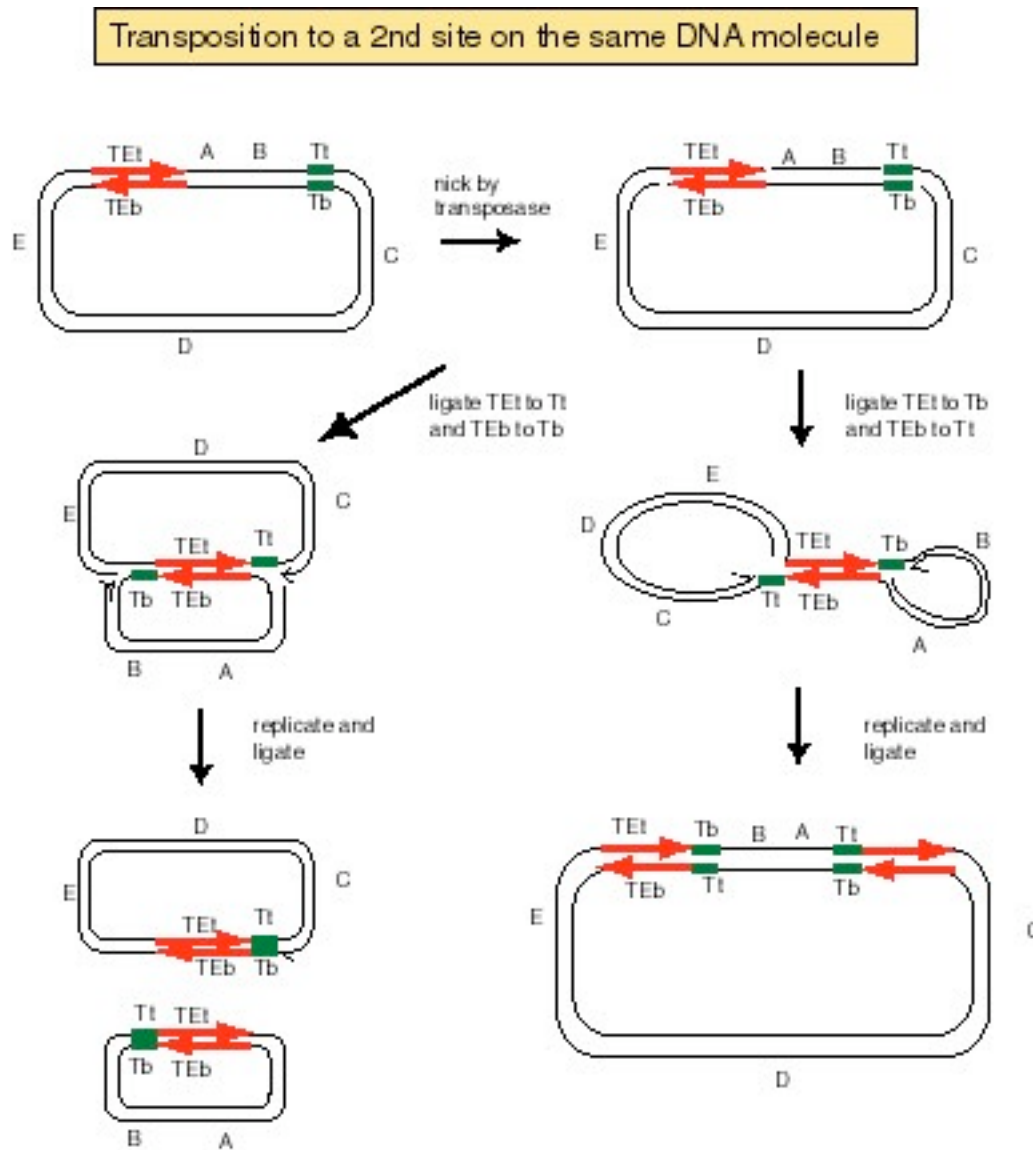
TEt is the top strand of the transposable element, TEb is the bottom strand.
Tt is the top strand of the target site, Tb is the bottom strand.

Answer 9.9. The L1 repeats of humans or other mammals. Actively retro-transposing elements have been isolated as integrating DNAs that mutate genes spontaneously, and the L1 repeats encode the enzymes needed for retrotransposition. *Alu* repeats could possibly be used, but a source of the enzymes needed for transposition needs to be provided.

H. Kazazian and his colleagues have developed the L1 repeat as a means for transposon tagging in mammalian cells (Moran et al., 1996, High frequency retrotransposition in cultured mammalian cells. Cell 87:917-927.