LECTURE 26

DNA REPAIR

A. The capability for repair of damaged DNA is found in one form or another in all organisms. Prokaryotes (e.g., *E. coli*) have five repair systems, whereas higher organisms (e.g., mammals) have one less. The one present in *E. coli* but not mammals is *photoreactivation*, a system designed specifically to repair UV damage of cells exposed directly to UV light.

B. Mutations in genes involved with DNA repair can be very serious, and have been implicated in elevated mutation frequency (and associated problems) and higher incidence of carcinoma, i.e., a failure to repair DNA correctly generally produces a mutation. The paradigm example in humans is *xeroderma pigmentosum*, a phenotype of intense freckling and high melanoma incidence, related to UV sensitivity. The syndrome actually results from defects in any of nine genes involved in repair of UV damage.

C. A recent study of the human genome has already revealed at least 130 genes whose products participate in DNA repair. More undoubtedly will soon be identified.

REPAIRING DAMAGED OR INCORRECT BASES

1. Photoreactivation (light repair, prokaryotes only):

    a) A light-dependent repair mechanism carried out by an enzyme called DNA photolyase. The enzyme repairs UV damage (largely pyrimidine-pyrimidine dimers) by binding to the dimers and using light energy to cleave dimer cross-links.

2. Excision repair (dark repair):

    a) A light-independent repair mechanism that involves three steps:

        (i) recognition of, binding to, and removal of damaged DNA
        (ii) repair synthesis of excised region by DNA polymerase
        (iii) ligation by DNA ligase to seal the break

    b) There are two major types of excision repair:

Base excision repair: removal/repair of abnormal or chemically modified bases; serves as another proof-reading mechanism

    a) Base excision repair involves DNA glycosylases, enzymes that recognize abnormal bases. Different glycosylases recognize different types of abnormal bases. Removal of a damaged base is estimated to occur 20,000 a day in each cell by a DNA glycosylase and humans have at least eight genes encoding different DNA glycosylases that are responsible for identifying and removing a specific type of base damage.

b) Enzymes (DNA glycosylases) cleave the glycosidic bond between the base and the deoxyribose sugar, leaving an apurinic or apyrimidinic site (AP site) that, in turn, is recognized by an AP endonuclease that clips out the sugar-phosphate group. Humans have at least two genes encoding enzymes with this function.

c) DNA polymerase beta fills in the missing nucleotide and DNA ligase seals the nick. There are at least two ligating enzymes – both use ATP to provide the needed energy.

d) Base excision repair is involved in repairing bases altered by alkylation (addition of methyl and ethyl groups) and deamination (removal of amine groups)

Nucleotide excision repair: removal/repair of larger fragments (≥ 2 bases) of damaged DNA

a) Nucleotide excision repair involves removal of larger lesions (e.g., thymine-thymine dimers) and utilizes a special enzyme called an *excinuclease* that cuts on either side of the damage and excises an oligonucleotide containing the damage.

b) The damage is recognized by one or more protein factors that assemble at the damage location(s) and the damaged area removed

c) DNA polymerases delta and/or epsilon fills in the correct nucleotides using the intact (opposite) strand as a template, followed by ligation (ligase)

3. Mismatch repair:

a) Provides a "backup" to the replicative proofreading carried out by most (but not all) DNA polymerases during DNA replication.

b) Occurs subsequent to DNA synthesis, so *must* have some way to determine which of a mismatched base pair (e.g., an A-G base pair) is the correct one (i.e., from the template strand). Correct determination of the template strand in prokaryotes occurs on the basis the of methylation state. How such recognition occurs in eukaryotes is not yet known. Adenine (A) bases are methylated in *E. coli*, whereas cytosine (C) bases are methylated in eukaryotes.

c) The appropriate enzyme (a GATC-specific endonuclease) makes a "nick" in the unmethylated strand (i.e., the recently synthesized strand) at GATC sites either 5' or 3' to the mismatch.

d) The incision site can be ≥ 1,000 nucleotides from the mismatch. If the damage is 5' to the mismatch, a 5'→3' exonuclease is required. Alternatively, if the damage is 3' to the mismatch, a 3'→5' exonuclease is required.

e) DNA polymerase delta fills in the gap and DNA ligase seals the nick.

REPAIRING STRAND BREAKS

a) Ionizing radiation and certain chemicals can generate both single-strand and double-strand breaks in the DNA backbone

Single-strand breaks: Repaired using the same enzyme systems (polymerase and ligase) used in base-excision repair

Double-strand breaks:

a) Direct joining of the broken ends (also called 'non-homologous end-joining'). This requires proteins that recognize and bind to the exposed ends and bring them together for ligating.

b) Homologous recombination – this requires information on the intact sister chromatid (available after chromosome duplication). The process is not yet well understood.

c) Two of the proteins used in homologous recombination in humans are encoded by *BRCA*1 and *BRCA*2. Inherited mutations in these genes predispose women to breast and ovarisn cancers.

REPAIRING EXTENSIVE DAMAGE

Postreplication (recombination) repair:

a) Occurs after DNA synthesis and when damage (e.g., thymine dimers) were not removed prior to DNA replication. What happens is that DNA polymerase "jumps over" the damage (e.g., a thymine dimer) and restarts DNA synthesis somewhere past the damage.

b) A recombination protein (*Rec*A in *E. coli*) stimulates recombination and exchange of single strands between the strand with the UV-dimers and gap and the sister double helix.

c) The resulting gap in the sister double helix is filled in by DNA polymerase and sealed by DNA ligase.

d) The "original" strand with the UV-dimer now has a complete "other" strand and the UV-dimer can now be removed by "normal" mechanisms.

Error-Prone (SOS) repair system:

a) This is an "SOS" or "last-ditch" response when damage is extremely great.

b) In essence, what happens is that cells release a battery of at least 30 'inducible' genes that are activated DNA damage is extreme; several of these genes encode proteins (including several DNA polymerases) that allow DNA synthesis to proceed across damaged regions – referred to as *translesional DNA synthesis*

c) There are several such DNA polymerases – many are specific for a given lesion (e.g., thymine-thymine dimers); they have gone by several names

TLS polymerases	Mutases
SOS polymerases	Error-prone polymerases

d) Thought to have evolved to minimize cell death from replication blockage (which is when DNA synthesis ceases when ‘normal’ DNA polymerase comes upon damaged or degraded DNA)

d) Start replication on one side of the damaged DNA and terminate replication on the other side – note a few ‘undamaged’ bases on either side of the damaged DNA also are replicated by the TLS polymerases

e) Most TLS polymerases can copy cognate lesions with fairly high fidelity; however, the TLS polymerases have much *reduced* fidelity (sometimes an order-of-magnitude or more) when replicating ‘normal’ or non-damaged DNA, leading to a rather high number of mutations

ORGANELLE GENETICS

A. The specialized subcellular organelles in eukaryotes that possess their own DNA molecules are *chloroplasts* and *mitochondria*.

1. Chloroplasts photosynthesize sugars from $CO_2 + H_2O$ by using sunlight as energy. The sugars are fermented anaerobically for energy. The waste product of photosynthesis is O_2.

2. Mitochondria utilize O_2 produced via photosynthesis to extract (additional) energy from sugars via a process called oxidative or aerobic metabolism.

3. Inheritance of both chloroplasts and mitochondria is non-Mendelian, as there is unequal contribution by of two parents. Both molecules essentially are maternally inherited, and both molecules are generally haploid genetically.

B. Chloroplast DNA [cpDNA]

1. Very large DNA molecule, typically ranging in size from 85 to nearly 300 kilobases (kb). Some cpDNA molecules can be up to 2,000 kb in size. Most cpDNAs appear to be covalently closed circles.

2. cpDNAs appear to carry the same set of genes, including coding sequences for ribosomal RNAs, transfer RNAs, ribosomal and other proteins involved in capturing solar energy, and four subunits of a chloroplast-specific RNA polymerase.

3. The cpDNAs that have been sequenced in their entirety contain 130-150 genes.

B. Mitochondrial DNA [mtDNA]

1. MtDNA in far better known and understood that cpDNA, in large part because of its generally smaller size. MtDNAs contain genes that are involved in oxidative (aerobic) metabolism, as well as genes that function in transcribing and translating mtDNA genes. Note that some genes (gene products, actually) that function in mitochondrial processes (aerobic metabolism) are encoded in the nucleus.

2. MtDNAs vary tremendously in size, from 16-17 kb in most vertebrates to 2,500 kb in some flowering plants. Most mtDNA molecules are covalently closed circles, although linear mtDNA molecules are known.

3. MtDNA is best known in vertebrates, where there are 37 distinct genes encoding two ribosomal RNAs, 22 transfer RNAs, and thirteen proteins. One non-coding region, called the “control” or “D-loop” region, contains the origin of replication.

4. There are two transcription regions in the mtDNA of vertebrates. One is on the H strand (28 genes), and one on the L strand (9 genes). H and L strands differ in buoyant density (a function of base pair composition). Two long transcripts (one from each strand) are produced and then “processed” by cleavage to separate the rRNAs and tRNAs. Messenger RNAs are polyadenylated.

5. Translation proceeds as with nuclear-encoded mRNAs (using both mtDNA and nuclear-encoded tRNAs) except that some of the codons have a different meaning (i.e., call for a different amino acid).

6. There are a few described "syndromes" in humans that originate as mitochondrial defects. All described syndromes (phenotypes) are maternally inherited.

7. Available data indicate that mtDNA sequences evolve 5-10 times faster than analogous, nuclear-encoded sequences. Recent evidence indicates that mtDNA sequences are not "proofread" repaired (either by DNA polymerases or the mismatch repair system) as efficiently as nuclear DNA sequences. This may account for the perceived higher rates of DNA evolution in mtDNA as opposed to nuclear-encoded DNA.

8. MtDNAs have proven very useful in population genetics, as the effective population size (N_{ef}) is four times lower than that for nuclear-encoded genes: one factor of two is because of haploidy (versus diploidy), and one factor of two is because of maternal (uniparental) inheritance versus biparental inheritance.