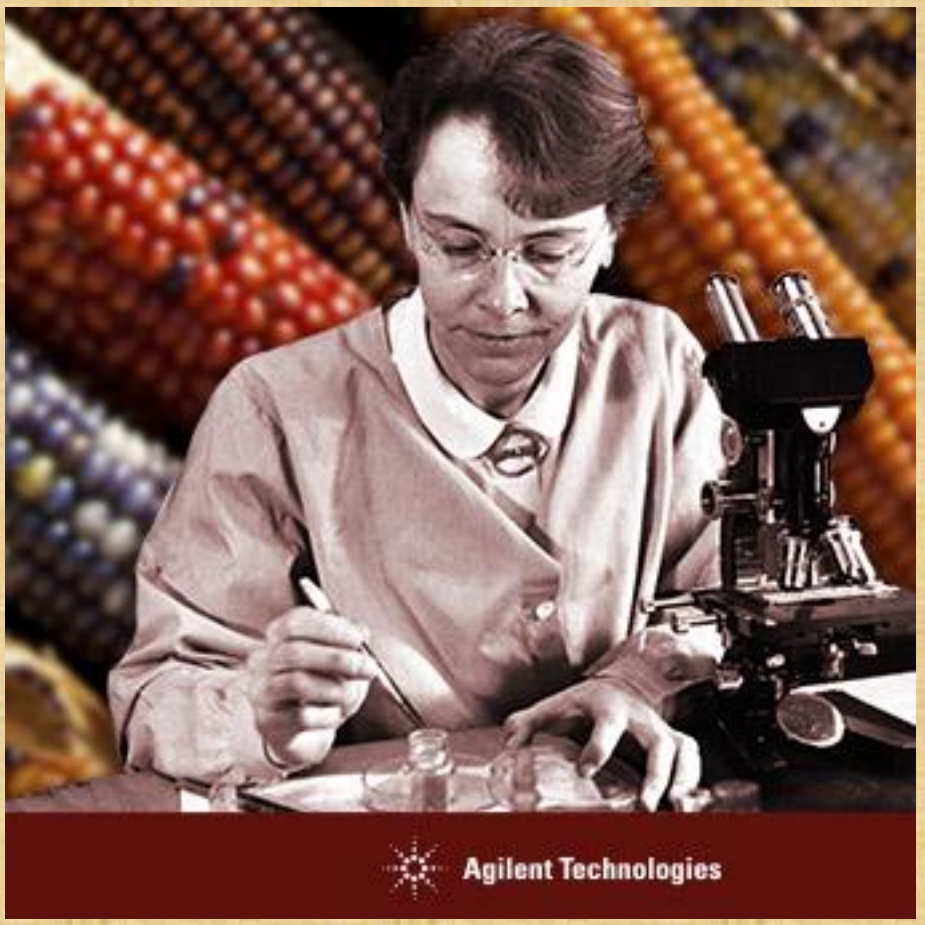
Transposons

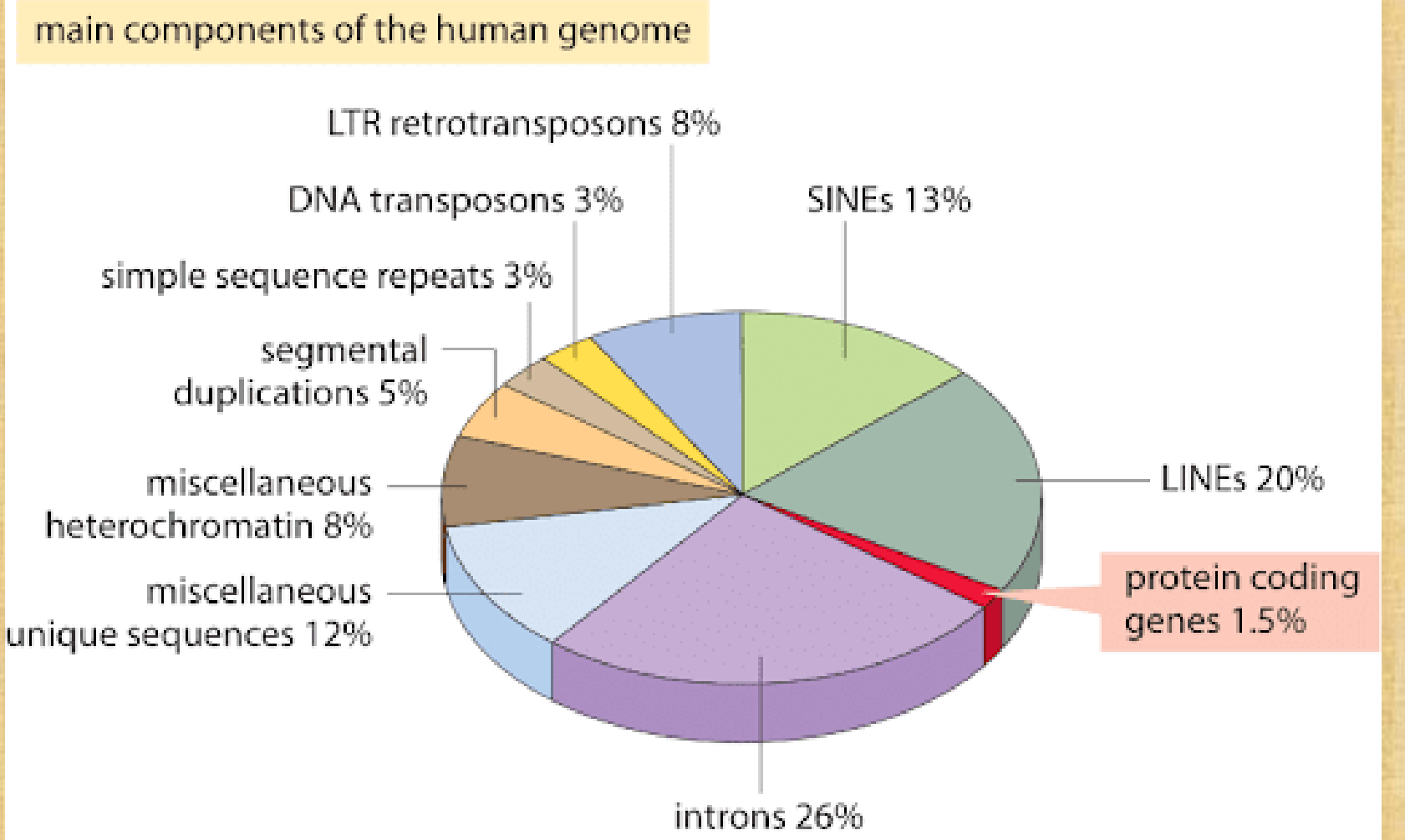
Introduction

- Transposable elements (TEs) are defined as DNA sequences that are able to move from one location to another in the genome. TEs have been identified in all organisms, prokaryotic and eukaryotic, and can occupy a high proportion of a species' genome. For example, transposable elements comprise approximately 10% of several fish species, 12 % of the *C. elegans* genome, 37% of the mouse genome, 45% of the human genome, and up to >80% of the genome of some plants like maize. From bacteria to humans, transposable elements have accumulated over time and continue to shape genomes through their mobilization.
- Barbara McClintock discovered the first mobile elements while doing classical genetic experiments in maize (corn) during the 1940s.
- Two categories: (1) those that transpose directly as DNA and (2) those that transpose via an RNA intermediate transcribed from the mobile element by an RNA polymerase and then converted back into double-stranded DNA by a reverse transcriptase (Figure).
- Mobile elements that transpose through a DNA intermediate are generally referred to as DNA transposons. Mobile elements that transpose to new sites in the genome via an RNA intermedi-ate are called retrotransposons because their movement is analogous to the infectious process of retroviruses.
- Mobile Elements That Move as DNA Are Present in Prokaryotes and Eukaryotes: Most mobile elements in bacteria transpose directly as DNA. In contrast, most mobile elements in eukaryotes are retrotransposons, but eukaryotic DNA transposons also occur.
- DNA transposons are useful tools to analyze the regulatory genome, study embryonic development, identify genes and pathways implicated in disease or pathogenesis of pathogens, and even contribute to gene therapy.

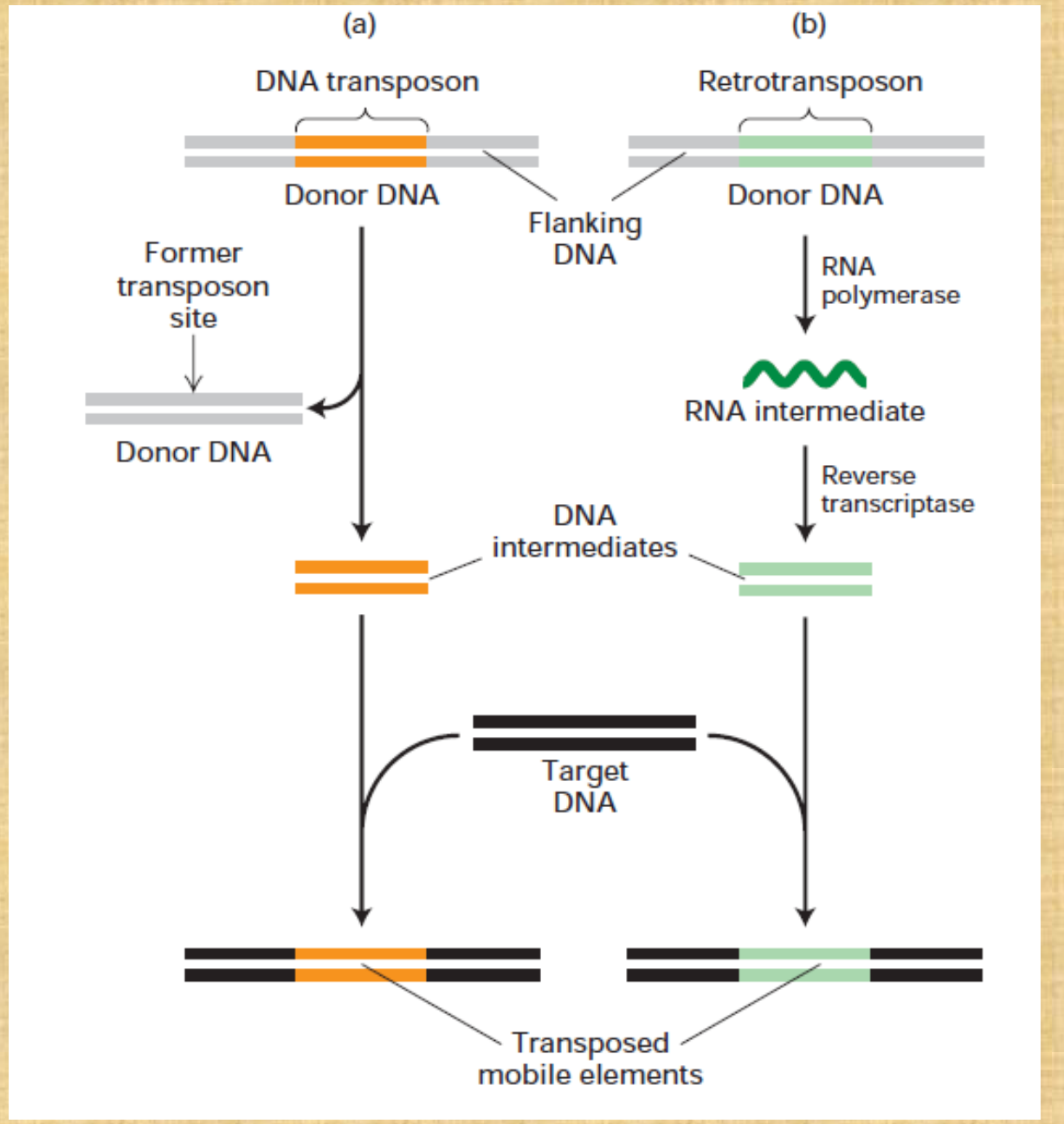
Barbara McClintock



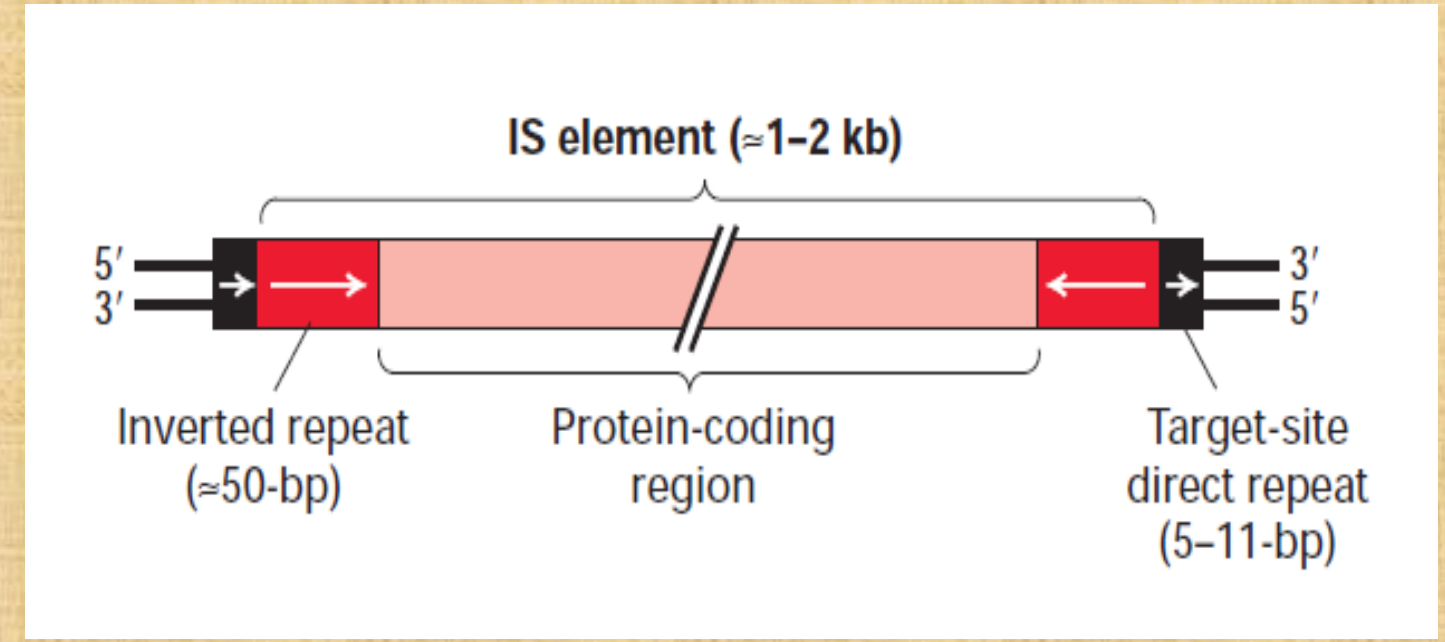
The transposition first discovered in 1948 by Barbara McClintock was initially thought to be unique to maize (corn) but later recognized to be common in eukaryotes, bacteria, viruses, phages and plasmids (a Nobel Prize in 1983).



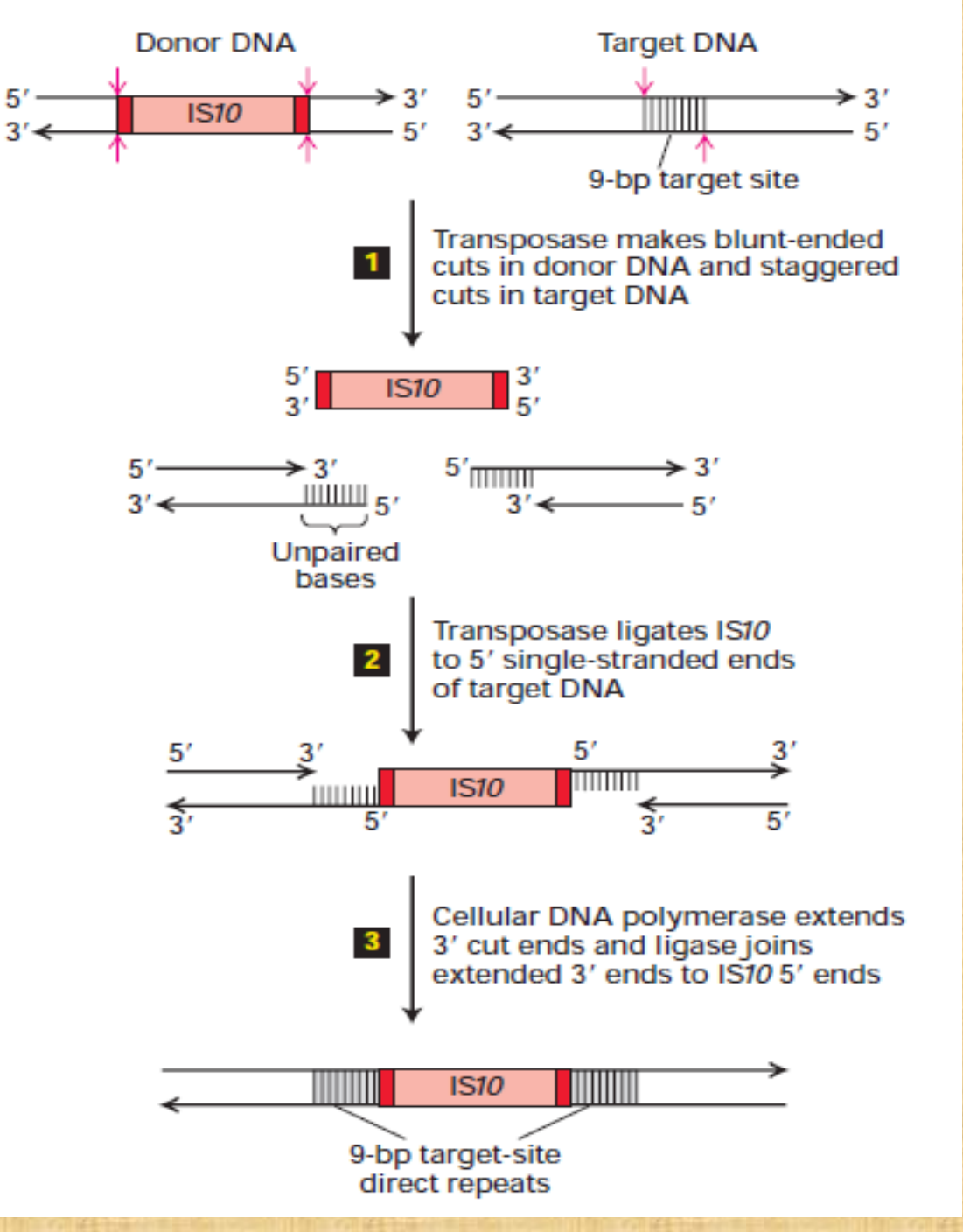
Short interspaced retrotransposable elements (SINEs), Long interspaced retrotransposable elements (LINEs)



Classification of mobile elements into two major classes. (a) Eukaryotic DNA transposons (orange) move via a DNA intermediate, which is excised from the donor site. (b) Retrotransposons (green) are first transcribed into an RNA molecule, which then is reverse-transcribed into double-stranded DNA. In both cases, the double-stranded DNA intermediate is integrated into the target-site DNA to complete movement. Thus DNA transposons move by a cut-and-paste mechanism, whereas retrotransposons move by a copy-and-paste mechanism.

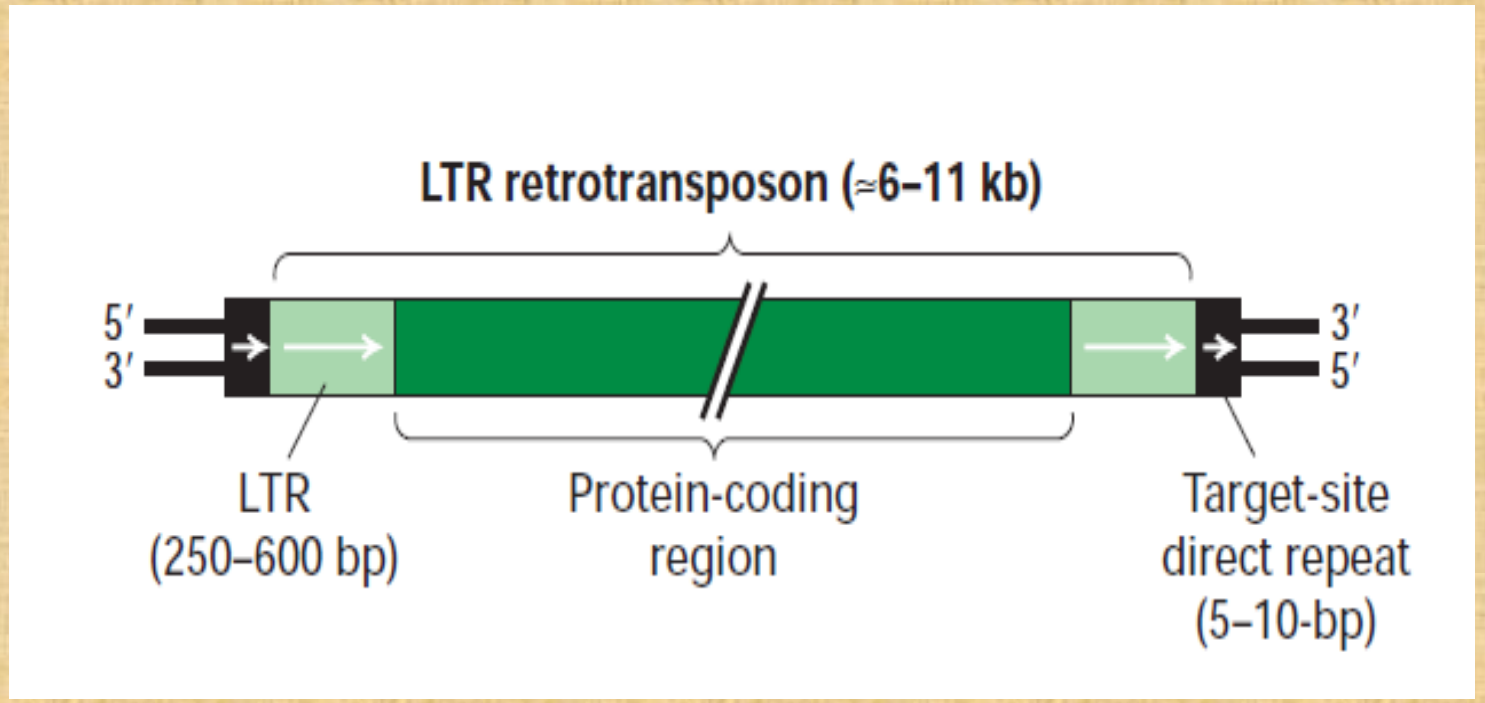


Bacterial Insertion Sequences: The first molecular understanding of mobile elements came from the study of certain *E.coli* mutations caused by the spontaneous insertion of a DNA sequence, ≈1–2 kb long, into the middle of a gene. These inserted stretches of DNA are called *insertion sequences,* or *IS elements.* So far, more than 20 different IS elements have been found in *E. coli* and other bacteria. General structure of bacterial IS elements. The relatively large central region of an IS element, which encodes one or two enzymes required for transposition, is flanked by an inverted repeat at each end. The sequences of the inverted repeats are nearly identical, but they are oriented in opposite directions. The sequence is characteristic of a particular IS element. The 5’ and 3’ short *direct* (as opposed to *inverted*) repeats are not transposed with the insertion element; rather, they are insertion-site sequences that become duplicated, with one copy at each end, during insertion of a mobile element. The length of the direct repeats is constant for a given IS element, but their sequence depends on the site of insertion and therefore varies with each transposition of the IS element. Arrows indicate sequence orientation.

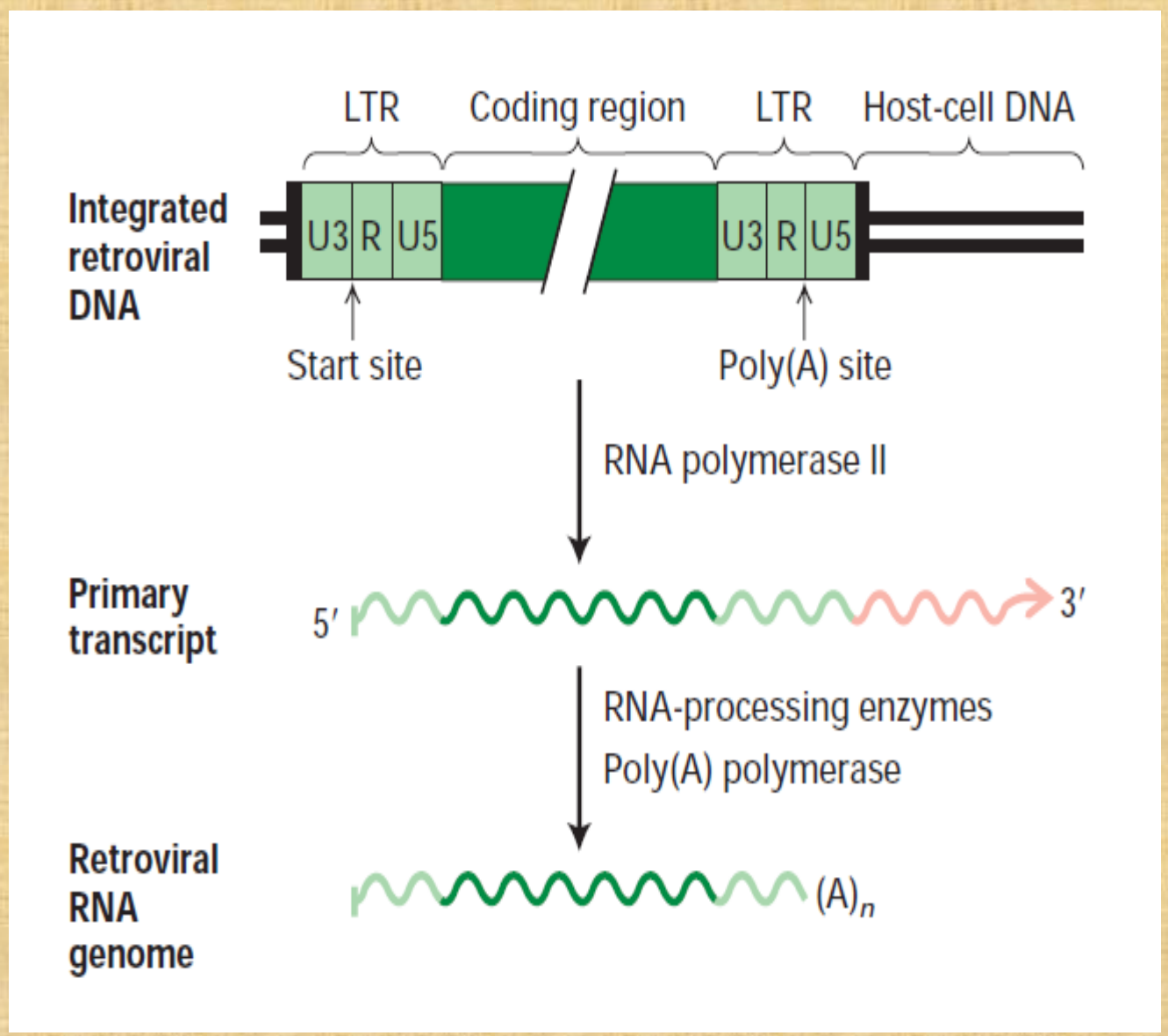


Model for transposition of bacterial insertion sequences.

Step 1 : Transposase, which is encoded by the IS element (IS*10* in this example), cleaves both strands of the donor DNA next to the inverted repeats (dark red), excising the IS*10* element. At a largely random target site, transposase makes staggered cuts in the target DNA. In the case of IS*10,* the two cuts are 9 bp apart. Step 2: Ligation of the 3’ ends of the excised IS element to the staggered sites in the target DNA also is catalyzed by transposase. Step 3 : The 9-bp gaps of single stranded DNA left in the resulting intermediate are filled in by a cellular DNA polymerase; finally cellular DNA ligase forms the 3’→5’ phosphodiester bonds between the 3’ ends of the extended target DNA strands and the 5’ ends of the IS*10* strands. This process results in duplication of the target-site sequence on each side of the inserted IS element. Note that the length of the target site and IS*10* are not to scale. [See H. W. Benjamin and N. Kleckner, 1989, *Cell* **59**:373, and 1992, *Proc. Nat’l. Acad.Sci. USA* **89**:4648.]

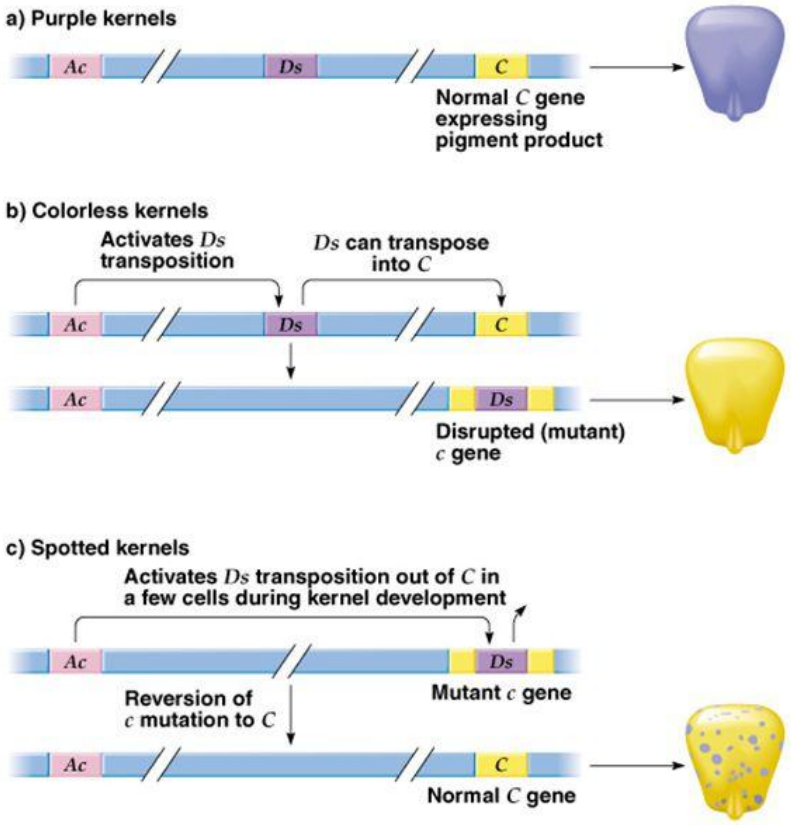


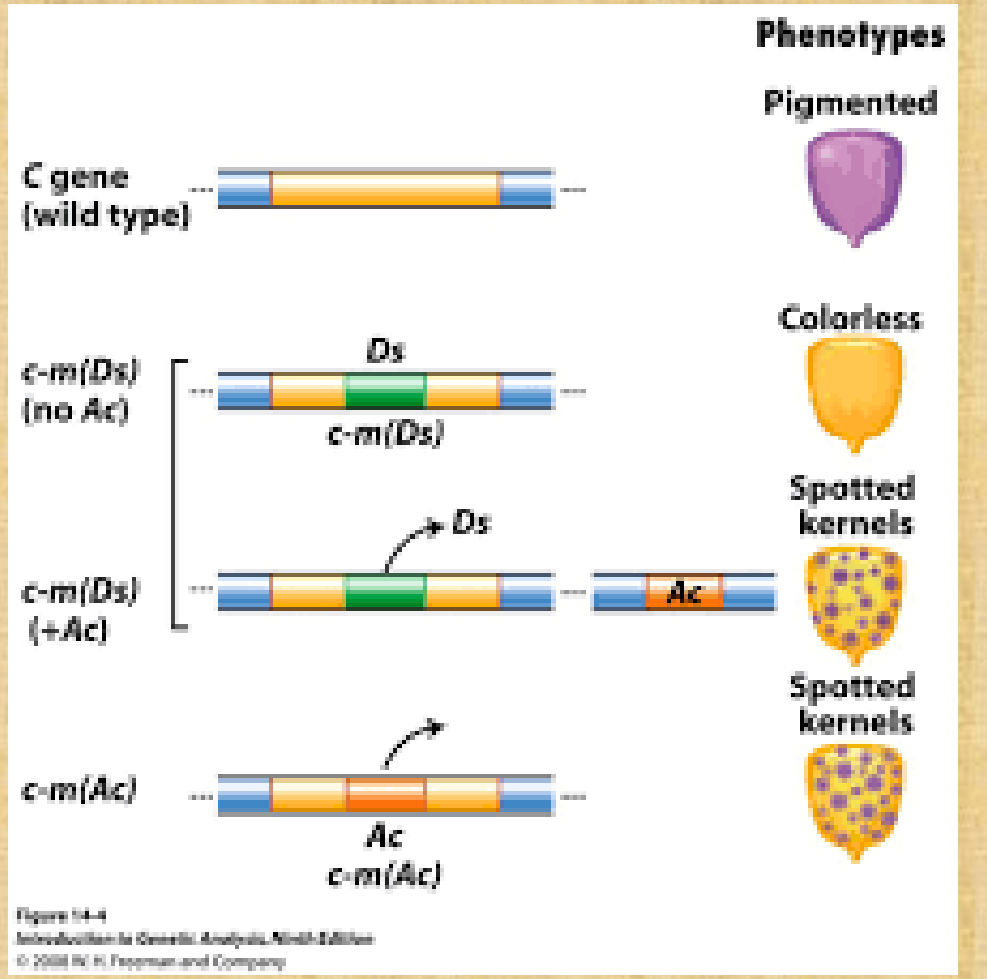
General structure of eukaryotic LTR retrotransposons. The central protein-coding region is flanked by two long terminal repeats (LTRs), which are element-specific direct repeats. Like other mobile elements, integrated retrotransposons have short target-site direct repeats at each end. Note that the different regions are not drawn to scale. The protein-coding region constitutes 80 percent or more of a retrotransposon and encodes reverse transcriptase, integrase, and other retroviral proteins.

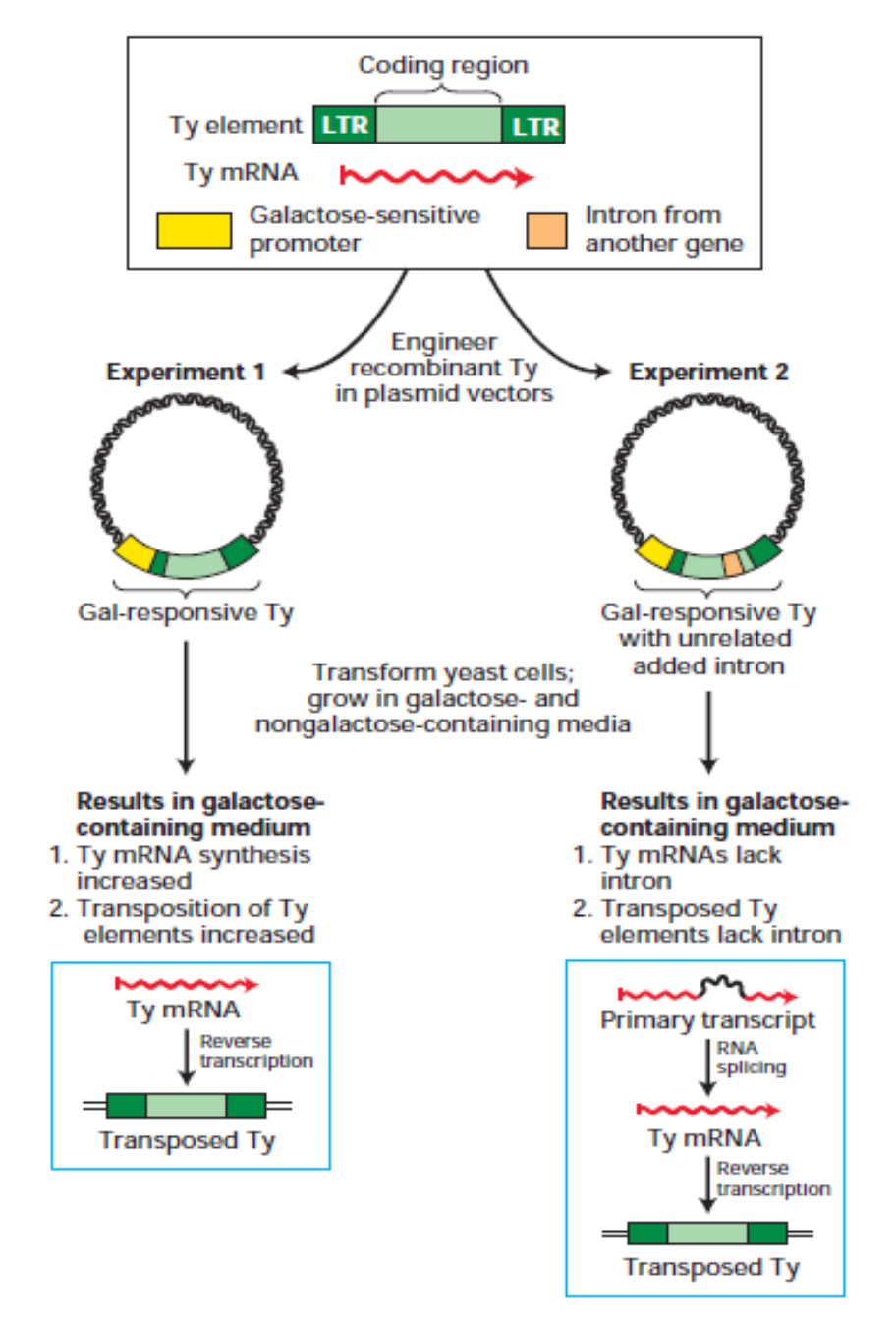


Generation of retroviral genomic RNA from integrated retroviral DNA.
The left LTR directs cellular RNA polymerase II to initiate transcription at the first nucleotide of the left R region. The resulting primary transcript extends beyond the right LTR. The right LTR, now present in the RNA primary transcript, directs cellular enzymes to cleave the primary transcript at the last nucleotide of the right R region and to add a poly(A) tail, yielding a retroviral RNA genome with the structure (Figure). A similar mechanism is thought to generate the RNA intermediate during transposition of retrotransposons. The short direct-repeat sequences (black) of target-site DNA are generated during integration of the retroviral DNA into the host-cell genome.

Fig. 7.24, Transposon effects on corn kernel color.







Recombinant plasmids demonstrate that the yeast Ty element transposes through an RNA intermediate. When yeast cells are transformed with a Ty-containing plasmid, the Ty element can transpose to new sites, although normally this occurs at a low rate. Using the elements diagrammed at the top, researchers engineered two different plasmid vectors containing recombinant Ty elements adjacent to a galactose-sensitive promoter. These plasmids were transformed into yeast cells, which were grown in a galactose-containing and a nongalactose medium. In experiment 1, growth of cells in galactose-containing medium resulted in many more transpositions than in nongalactose medium, indicating that transcription into an mRNA intermediate is required for Ty transposition. In experiment 2, an intron from an unrelated yeast gene was inserted into the putative proteincoding region of the recombinant galactose-responsive Ty element. The observed absence of the intron in transposed Ty elements is strong evidence that transposition involves an mRNA intermediate from which the intron was removed by RNA splicing, as depicted in the box on the right. In contrast, eukaryotic DNA transposons, like the Ac element of maize, contain introns within the transposase gene, indicating that they do not transpose via an RNA intermediate. [See J. Boeke et al., 1985, *Cell* **40**:491.]