iBiology.org Teaching Tools
Susan Wessler's Lecture Part 1:
The Dynamic Genome: Introduction to transposable elements

Teaching Tools were prepared by Jim Burnette with Susan Wessler.

Contents

---

1. Keywords and Terms

Transposable element, maize, Barbara McClintock, transposase, genome, genotype, phenotype, mutation

2. Lecture Notes

Wessler begins by introducing transposable elements (TEs), small pieces of DNA that can insert, and excise, throughout the genome. She describes their discovery by Barbara McClintock and their current role in the study or genetics.

Introduction

Barbara McClintock discovered transposable elements (TEs) while studying spotted maize (corn, *Zea mays*) kernels. The talk will cover three points:

1. The discovery of TEs by McClintock.
2. What TEs are and how they move.
3. The abundance of TEs in genomes.

1. Discovery of TEs.

Corn is a useful plant for studying genetics due to the many visible phenotypes. Many of these phenotypes can be seen in grocery stores around Thanksgiving (Indian corn).

In particular McClintock was interested in the genetics that caused spotted corn kernels.

McClintock's notebook entries show the detailed phenotypes found in the kernels and the typed notes show the genetics.

McClintock realized that the spotted kernel phenotype was caused by a newly identified type of mutation called a transposable element. Unlike DNA base pair changes and insertions and deletions, TEs are insertions of DNA sequence that can excise. Thus, TEs cause *reversible* mutations unlike other mutations that are not reversible.

Spotted corn kernels: The yellow parts of the kernel are due to a TE inserted inside a pigment gene preventing pigment from being produced. The purple areas have a wild type gene (no TE inserted) and pigment is produced. To obtain this phenotype the TE insertion must excise from the pigment gene in some cells during development so that both mutant and wild type phenotypes are present in the same kernel.

The sectors arise because plant cells, unlike animal cells, do not move during development.

Taking advantage of the identification of the corn kernel phenotypes McClintock was able to dissect the details of the genetics. The three genotypes described by her were:

1. Pigmented: fully colored kernel due to a wild type pigment gene.
2. Colorless: No pigment (yellow kernel) due to the insertion of a non-autonomous TE (nTE) or a TE that cannot move on it's own into the pigment gene.
3. Spotted Kernels: Some pigment because the nTE moved with the help of a second, autonomous TE that can excise the nTE.
4. Spotted Kernels: Some pigment because the TE in the pigment gene is autonomous, it can excise itself.

In summary McClintock described two new genetic elements: autonomous TEs and non-autonomous TEs (nTEs).

McClintock was way ahead of her time

She made her discoveries in 1940, but they were not recognized as significant by the scientific community until much later. TEs were thought to be an oddity of corn or perhaps limited to plants.

During the 40 years after her discovery TEs were found in almost all organisms:

- 1950's — *Drosophila melanogaster*, the fruit fly.
- 1960's — The bacteria *Escherichia coli*.
- 1970's — Humans (*Homo sapiens*)

TEs can cause striking phenotypes in flowers. Many garden plants have them. Note that TEs cause random sectoring.

McClintock discovered a new component in the genome

Pre-discovery, gene followed by gene by gene, etc. on chromosomes.
Post-discovery the genome contains TEs distributed in and among the genes on a chromosome.

This discovery was very important in understanding the organization of the genome of all organisms. For this reason McClintock was awarded the Nobel Prize in Physiology or Medicine in 1983. It is unusual for a botanist to win a prize in this category and demonstrates the importance of her discovery to genome science.

2. What are TEs and how do they move?

TEs are simple genetic units made of 100s –1000s of bps of DNA. Autonomous elements contain a gene for the enzyme transposase; the enzyme that moves TEs. The transposase gene is flanked by terminal inverted repeats (TIRs). This is a DNA sequence that reads forward on one end and is flipped over on the other end of the transposase gene. The whole element is flanked by a short direct repeat called the target site duplication (TSD). The non-autonomous element often contains a mutated transposase gene that is no longer functional. Otherwise nTEs contain the same features of TEs including TIRs and TSDs.

Do you want these subtitles italicized??
Transposase enzyme encoded by autonomous elements binds to the TIRs.
The transposase gene can be transcribed and translated to produce the multifunctional enzyme transposase. One function of transposase is to bind to the TIR DNA sequence. One transposase enzyme binds to one TIR at one end of the TE and a second enzyme molecule binds to the TIR at the other end of the TE. The two enzyme molecules then bind to each other, or dimerize, causing the TE DNA to form a loop. Dimerization is a second function of transposase. The transposase cuts, or cleaves (third function), the chromosome to excise the TE and inserts the TE into a different target site at a new location in the genome. Thus a single enzyme can perform three different functions: DNA binding, protein binding, and DNA cutting. It also performs a fourth function, ligation, described below.

TE families
A TE family has both autonomous and non-autonomous elements. A family is defined by having the same DNA sequence for the TIRs and TSD. The non-autonomous TEs are often different sizes and have different mutations in the transposase gene.

Generation of the Target Site Duplication (TSD)
The target site is a specific sequence of DNA that the transposase uses as the insertion site for an element. The target site is cleaved by the transposase creating a staggered cut. This results in overhanging ends where there the DNA is not hydrogen bonded to the opposite strand of DNA. This type of cleavage is similar to restriction enzymes. The TE is inserted between the staggered ends of the insertion site and ligated to the overhang by the TE. This leaves gaps of single stranded DNA, which are filled in by host DNA repair enzymes resulting in the TSD. The TSD length is determined by the length of the overhang formed when the transposase cuts the DNA.

Genomes contain many TE families
There are many different TE families in a genome. Shown on the slide are two families. The families have different TIR and TSD sequences and lengths.

TEs can increase their copy number during transposition
In addition to moving from location to another, TEs can increase their copy number. During DNA replication, the chromosome is duplicated so that both sister chromitids have a copy of the TE. The TE may then move from one chromatid to another place in the genome. The gap that remains after the TE moves, called an empty site, can be repaired by DNA repair enzymes using the sister chromatid as a template and thus copying a TE back into the original site. The result is three copies of the TE.

The second mechanism for increasing copy number is when the TE moves to a location in front of the replication fork. The result is that when the fork passes the TE a new copy on the sister chromatid is created, again resulting in three copies of the TE.

Two classes of TE.
The talk focuses on Class 2 or DNA TEs, so called because they transpose through a DNA intermediate. Class 1 TEs are retrotransposons, also known as RNA TEs, because they transpose through an RNA intermediate. Class 1 TEs have terminal direct repeats instead of TIRs; they do have TSDs.

How a retro increases its copy number
The RNA element is transcribed into RNA. The RNA is copied into a DNA copy by the enzyme reverse transcriptase, encoded by the retrotransposon. A second DNA strand is made and the element is inserted into the genome. RNA elements are like printing presses. One element can be transcribed many times and result in many new copies of the TE in the genome.

3. How abundant are TEs in genomes?
Genomics: the study of whole genomes at once.
Sequencing of genomes revealed that TEs are the largest component of the genome:

- 50% of animal genomes
- Plants even higher
    - 75% maize

- 85% barley
- 98% iris

Despite all these TEs, the organisms are fully functional.

An analogy: The human genome is 2.5 billion letters (A, G, C, T) or approximately 1000 textbooks (no pictures). 500 of the 1000 books are TEs while only to 20-40 contain the genes!

Where are TEs located?
A typical human gene can have hundreds of TEs located in the introns, but rarely in exons.

The size of cereal grass genomes:
Rice 300 Mb
Maize 2500 Mb
Sorghum700 Mb
Barley 5000 Mb

Yet, all have the same number of genes. The increase in the size of the genome is mostly due to TEs.

How do organisms live with so many TEs?
Most TEs are dead, that is, they can no longer move. Mutations have occurred in the transposase genes and thus, they no longer function. TEs also have evolved to insert into safe-havens or locations that do not affect host genes. These safe-havens are introns, between genes and, most often, within other TEs. In addition, most TEs are epigenetically silenced by the host which tightly condenses the region of the chromosome containing the TEs, thus preventing translation of transposase or other proteins required for TE activity.

McClintock's scenario for TEs as tools of evolution
McClintock hypothesized that TEs could be beneficial to the host if they usually do not move except under stress (i.e., climate change). This would increase the mutation frequency and the genetic diversity. This new diversity may be beneficial in changing environments.

3. Recommended Reading

1. Reece, JB, Urry, LA, Cain, ML, Wasserman, SA, Minorsky, PV, and Jackson, RB. (2011). Campbell Biology, Ninth Edition. Chapter 21.4 pgs. 434–437.

2. Griffiths, AJF, Wessler, SR, Carroll, SB, Doebley, J. (2012). Introduction to Genetic Analysis, Tenth Edition. Chapter 15. The Dynamic Genome: Transposable Elements. Pgs. 523-552.

4. Review Questions

1. Who discovered transposable elements?

    a. Rosalind Franklin
    b. Gregor Mendel
    c. James Watson
    d. Barbara McClintock

2. TEs cause a unique type of mutation because:

    a. The mutation kills the organism.
    b. The mutation is reversible.
    c. The mutation is permanent.
    d. The mutation has no effect.

3. Which statement describes the relationship of autonomous TEs vs. non-autonomous:

    a. They are completely different elements.
    b. Non-autonomous TEs never move while autonomous TEs move freely.
    c. Non-autonomous TEs rely on the transposase enzyme found in autonomous TEs.
    d. Non-autonomous and autonomous TEs describe two different families.

4. Select the functions of transposase (select all that apply):

    a. Hydrolysis
    b. DNA excision
    c. DNA cleavage
    d. Protein cleavage
    e. Dimerization
    f. Polymerization
    g. Ligation
    h. DNA binding
    i. RNA binding

5. The enzyme encoded by a TE that moves elements is called:

    a. Polymerase

b. Ligase
c. Transposase
d. Protease

6. The DNA sequences that define the ends of a DNA TE are called:

    a. TIRs
    b. SIRs
    c. Codons
    d. Splice sites

7. All of the TEs that have the same TIR and TSD sequences and share a transposase for movement belong to the same:

    a. Clade
    b. Unit
    c. Group
    d. Family

8. The most abundant genetic element in a eukaryotic genome is:

    a. Genes
    b. Exons
    c. TEs
    d. Introns
    e. Protein

9. What is the term used to describe how the host controls TEs?

    a. Silencing
    b. Euchromatin
    c. Modification
    d. Degradation

10. Choose the answer below that completes this sentence: Retrotransposons (Class 1 Elements) transpose using a(an) _____ intermediate while DNA transposons move using a(an) _____ intermediate.

    a. Protein, Protein
    b. RNA, RNA
    c. RNA, DNA
    d. DNA, RNA
    e. DNA, DNA

5. Answers to Review Questions

1. d. Barbara McClintock discovered TEs in the 1940s while studying spotted corn kernels.

2. c. TEs cause reversible mutations because they can excise from the insertion site.

3. c. A non-autonomous element does not encode a functional transposase while the autonomous element encodes functional transposase that can move both types of elements.

4. b, c, e, h, g. The act of transposition requires transposase to bind to DNA at the TIRs (h), dimerization of the two enzymes bound to the ends of the TE (e), cleavage of the DNA to excise the TE (c), DNA binding and cleavage at the new target site (h, c), and finally ligation of the TE to the chromosome at the new site (g).

5. c. Autonomous elements encode the enzyme transposase that is capable of excision and insertion of a TE.

6. a. TIRs are inverted DNA sequences that define the ends of the TE and serve as binding sites for the transposase.

7. d. Family. Members of a TE family use the same transposase for movement and usually have the same TIR and TSD sequences.

8. c. TEs comprise at least 50% of the DNA of most eukaryotic genomes and up to 98% in a few including iris.

9. a. Silencing is the mechanism of condensing chromatin into very dense heterochromatin in order to prevent access to the DNA.

10. c. Retrotransposons (Class 1 Elements) transpose using a(an) RNA intermediate while DNA transposons move using a(an) DNA intermediate.

6. Discussion Questions

1. How do the functions of transposase relate to TE structure and movement?

2. Discuss how TE activity results in spotted corn kernels during kernel development.

3. The title of the talk is the Dynamic Genome. Discuss why this title was chosen based on the content of the lecture.

7. Answers to Discussion Questions

1. Discussion Points:
    a. DNA Binding
        a. TIRS
            a. Inverted repeats
            b. Two enzymes bind to the two TIRS
            c. DNA cleavage at the ends of TIRs for excision
        b. TSD
            a. Binding site for transposase
            b. DNA cleavage site for insertion
            c. DNA ligation site
    b. Protein-Protein binding (dimerization)
            a. Transposase bound to TIRs bind to each other.
    c. DNA cleavage:
            a. Excision at existing locus between TIR and TSD.
            b. Insertion at new locus by staggered cut of the TSD.
    d. DNA ligation to paste the TE at the new location.
    e. Because the recognition of the TE is based on the TIRs, all elements with identical TIR sequences will be excised by the same transposase. This defines a TE family and results in two sub-categories: autonomous and non-autonomous. Autonomous elements contain a gene for transposase and are capable of moving themselves. Non-autonomous elements lack a functional transposase gene and rely on an autonomous one to move. An analogy is someone who owns a car versus someone who rides the city bus.
2. Discussion Points: At the start of development, the TE is inside a gene that encodes a protein necessary to develop a pigment. During seed development, the TE is excised and gene function is restored. In that cell and it's progeny pigment can be made. If the excision occurs early in development the sector or spot will be large relative to events that occur later in development with less time for cell division.
3. Discussion Points: The genome is not just static DNA as often described in textbooks. The most striking example of the dynamics is the activity of TEs which results in:
    a. New mutations that can result in phenotypes such as spotted corn kernels.
    b. New copies of TEs resulting in increased genome size.
    c. Differences in genome size despite the same number of genes.

8. Explain or Teach These Concepts to a Friend

1. Advanced Topic: Research DNA replication and explain the two mechanisms that result in the increase of TE copy number.

2. Explain how the function of Transposase and the DNA TE structure work together for TE movement.
3. Explain how the genome is a dynamic place.
4. Some people say that non-autonomous elements are parasites of autonomous elements. Why do they make this claim?

9. Research the Literature on Your Own

1. Are there any active TEs in the human genome? Research both RNA elements (Class I) and DNA elements (Class II). Do any insertions result in genetic diseases?
2. How do RNA elements insert into the genome?
3. What is ALU and the genetic locus PV92?
4. Most TEs have one gene for Transposase, but Ping is an exception. Read the paper by Yang G, Zhang F, Hancock C N and Wessler S R (2007). Transposition of the rice miniature inverted repeat transposable element mPing in Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA. 104: 10962 - 10967. Discuss how the two genes encoded by Ping are required for transposition.