The Central Dogma Revisited

Weimin Zhong, Ph.D.
weimin.zhong@yale.edu

Department of Molecular, Cellular & Developmental Biology
Yale University

Learning Goals for This Lecture

- To appreciate how each step of the central dogma of molecular biology itself is subject to regulation.

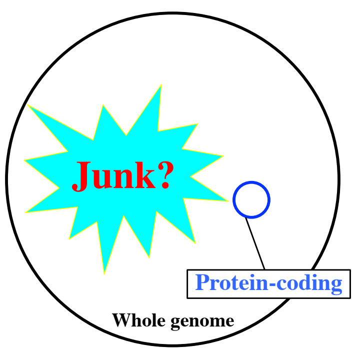
- To understand the differences between genetics and epigenetics.

- To appreciate how gene expression can be regulated through chromatin modifications.

- To appreciate that our genome is mostly used for producing regulatory RNAs (not proteins) and other functional elements.

- To appreciate how scientific dogmas are subject to revision as we gain new knowledge from new experiments.

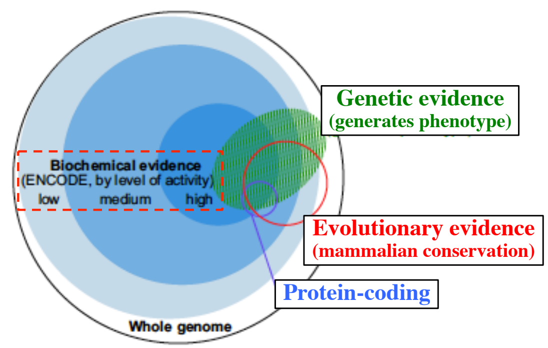
Genes Control Development, *but* 98% of Our Genome Does Not Encode Protein



- **Human genome has been sequenced, a triumph of the Human Genome Project (1986-2003; at a cost of $2.7 billion).**
- **Now everyone can have his/her genome sequenced for a ~$1,000 in one day! (personalized medicine).**

The Encyclopedia of DNA Elements (ENCODE) Project
(to identify all functional elements in the human genome sequence)

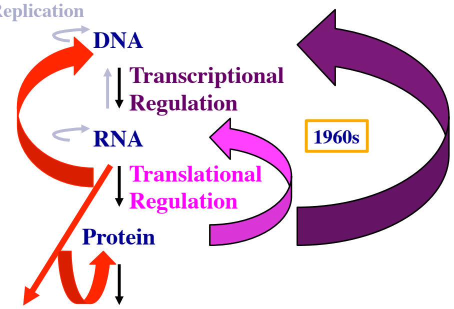
Most of Our Genome Does Not Encode Protein *but …*



Kellis *et al*. (2014) *PNAS* 111:6131-6138

DNA Is Life's Blueprint

(The Central Dogma of Molecular Biology)



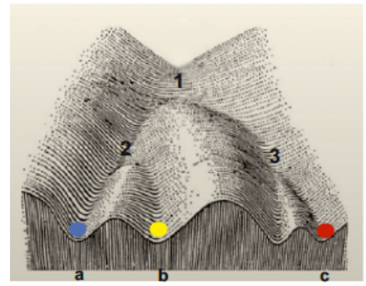
Today:
Regulatory RNAs?
- **siRNA**
- **miRNA**
- **lncRNA**

RNA as enzymes (Ribozymes; 1980s)

Biochemical Function

Every step in the central dogma is subjected to regulation!

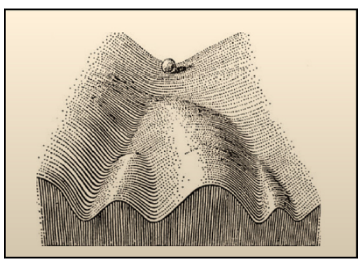
Cellular Differentiation Viewed as Decisions on Branching Points



Conrad H. Waddington (1957)

All cells of a multicellular organism contains the same genome, so why are genes differentially expressed?

Cellular Differentiation Viewed as A Result of Changing Epigenetic Landscape

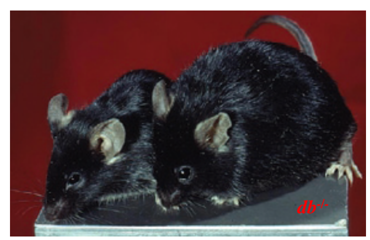


Epigenetics
("epi" = "on top of")

"… the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being" - Waddington (1942)

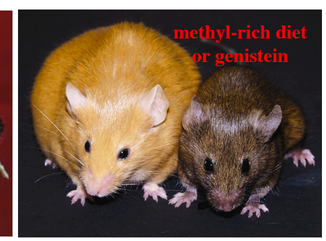
The study of any potentially stable and, ideally, heritable change in gene expression or cellular phenotype that occurs without changes in Watson-Crick base-pairing of DNA
(Goldberg *et al*. *Cell* 128: 635-638, 2007)

Genetics & Epigenetics



Genetically Different
(one homozygous for the *db* mutation)

Douglas L. Coleman
(Lasker Award, 2010)

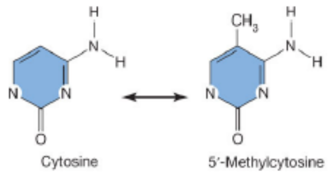


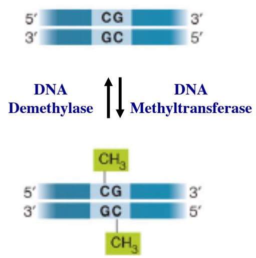
Genetically Identical
(pregnant mothers fed different diet)

Dolinoy *et al*. (2007)
*Reprod. Toxicol.*23:297-307

You can inherit something beyond DNA itself!

DNA Methylation: An Epigenetic Change
(occurs at CpG islands at promoters and silences transcription)



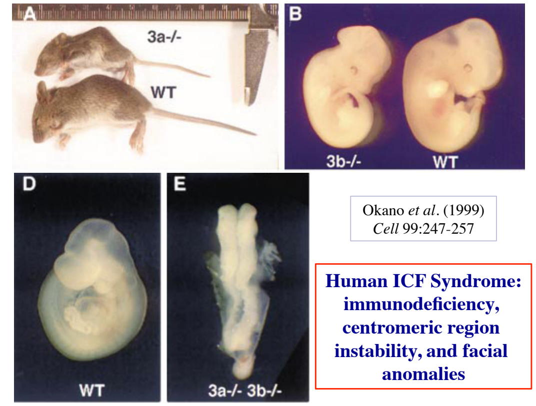


When DNA is replicated, a maintenance methylase catalyzes the formation in the new DNA strand

Bisulfite sequencing can detect the pattern of methylation
(teating DNA with bisulfite converts cytosine, but not 5-methylcytosine, residues to uracil)

Is heritable but also reversible, so the pattern can be altered
(in mammals usually occurs in C residues that are adjacent to G residues or CpG islands)

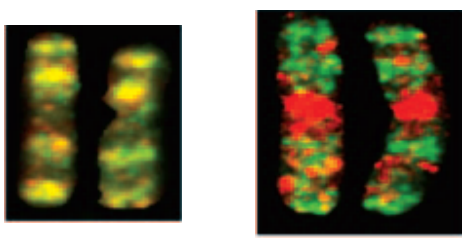
DNA Methylation Is Essential for Development
(DNA methyltransferase 3B mutation causes ICF Syndrome in human)



Okano *et al*. (1999) *Cell* 99:247-257

Human ICF Syndrome: immunodeficiency, centromeric region instability, and facial anomalies

DNA Methylation: An Epigenetic Change
(occurs at promoters and silences transcription)

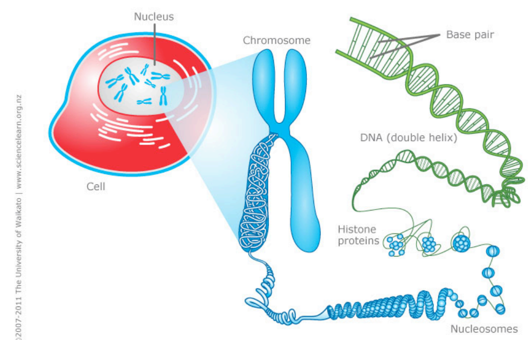


3-year-old twins **50-year-old twins**

Patterns of DNA methylation in monozygotic twins of different ages
(comparative genomic hybridization for methylated DNA; chromosome 3)

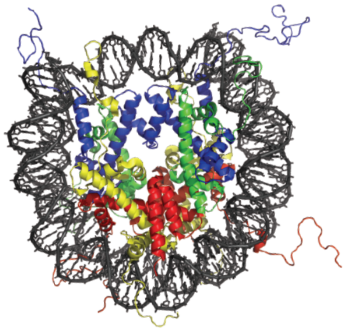
Fraga *et al*. (2005) *PNAS* 102:10604-10609

How Long Is the DNA in One of Our Cells?
(DNA needs to be compactly packaged to fit inside a nucleus)



~ 2 meters or 6 feet
(2 x 3 x 10^9 bp x 0.34 nm/bp = 2.04 x 10^9 nm)

The Nucleosome: Fundamental Unit of Chromatin

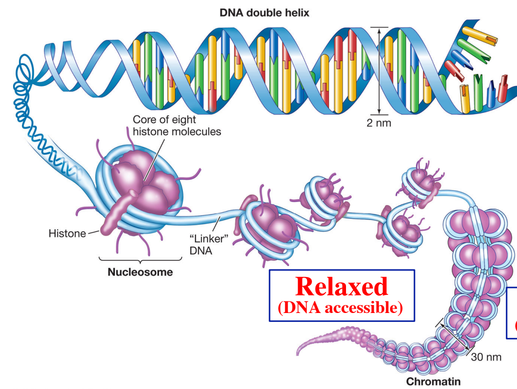


One nucleosome:
146 bp of DNA +
8 Histone proteins

- H2A
- H2B
- H3
- H4

x2 each

Two Flavors of Chromatin: Relaxed and Condensed



Two Flavors of Chromatin

(nuclear DNA exists in two forms that reflect the transcription activity in the cell)

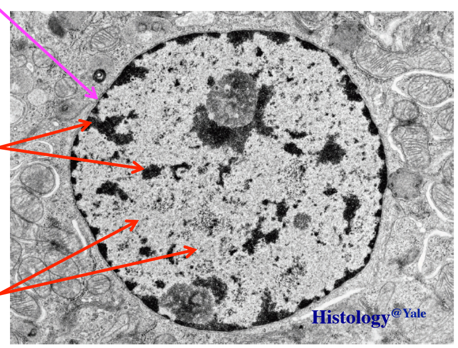
Nuclear Envelope

Heterochromatin

small, darkly staining and irregular particles, often accumulated adjacent to the nuclear envelope

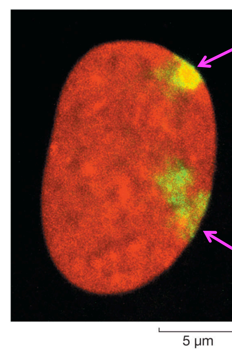
Euchromatin

dispersed and not easily stained, most abundant in cells with many genes in active transcription



Euchromatin & Heterochromatin

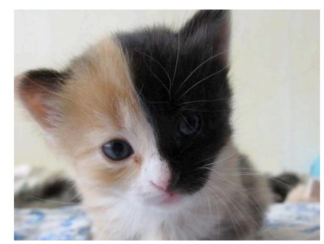
(nuclear DNA exists in two forms that reflect the activity in the cell)



Inactivated X (Barr body)

Active X

Tortoiseshell Cat

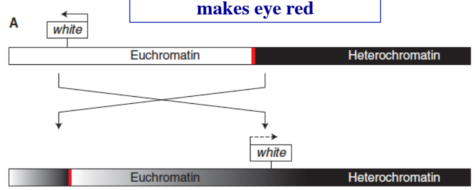


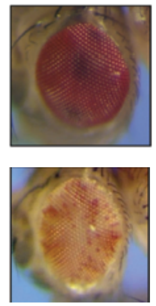
Only one of the two X chromosomes in somatic cells of mammalian (human) females is transcriptionally active

Position-Effect Variegation (PEV)

(when a gene normally in euchromatin is juxtaposed with heterochromatin)

Expression of the *white* gene makes eye red

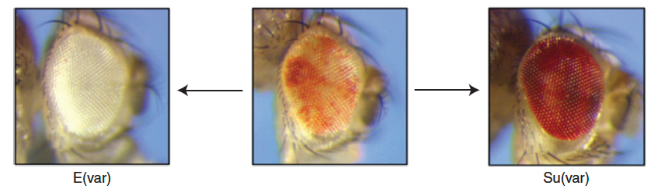




Cells of the same type can express the same gene differently!

Elgin & Reuter. (2013) *Cold Spring Harb. Perspect. Biol.* 5:a017780

Epigenetics: A Bridge Linking Genotype and Phenotype

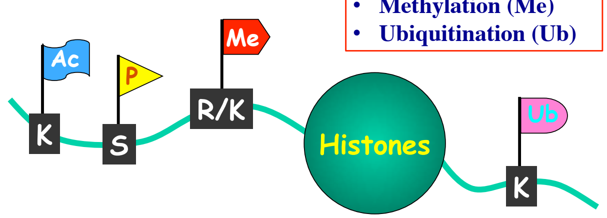


Enhancers and suppressors of PEV include chromatin modifiers

Genetic mutations as well as environmental impacts such as diet, toxins or stress can affect the readout of the DNA through affecting the levels and types of chromatin modifications

Histone Tails Are Heavily Modified After Translation

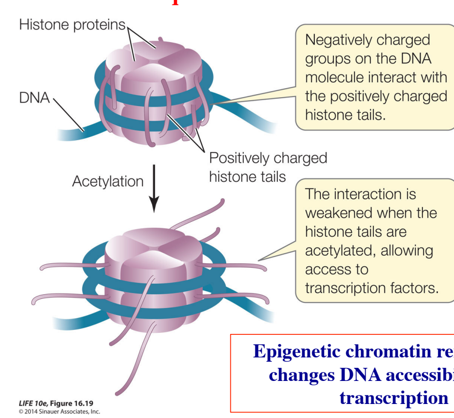
Covalent Modifications:
- **Acetylation (Ac)**
- **Phosphorylation (P)**
- **Methylation (Me)**
- **Ubiquitination (Ub)**



Courtesy of Yang Shi
(Harvard Medical School)

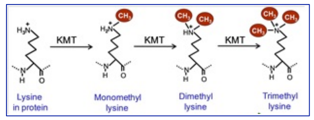
Remodeling of Chromatin for Transcription

(transcription initiation requires nucleosomes to be less compact)



Epigenetic chromatin remodeling changes DNA accessibility for transcription

Histone Lysine Methylation: Location & Degree Matter



Gene OFF

Heterochromatin

ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGIKK H3 K
4 9 27 36 79

Enhancers Promoters Gene Bodies

Gene ON

Histone Lysine Methylation Is Dynamically Regulated by Many Enzymes

Methylation

SET1, MLL1, MLL2, MLL3, MLL4, SMYD3 — SUV39H1, SUV39H2, G9a/GLP-1, SETDB1, RIZ1, ASH1 — EZH2 — HYPB (ySET2), NSD1, SMYD2 — Dot1

ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGIKK H3 K
4 9 27 36 79

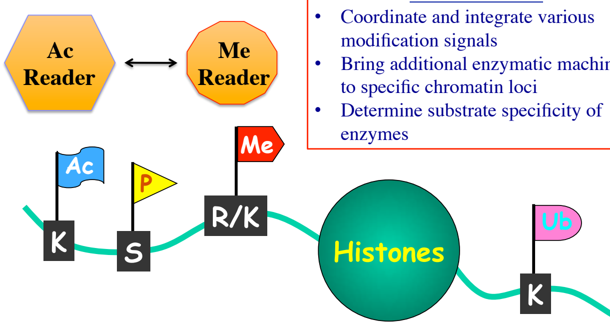
LSD1, SMCX, SMCY, RBP2, PLU-1 — LSD1/AR, JHDM2A, JMJD2A, JMJD2B, JMJD2C, JMJD2D — UTX, JmjD3 — JHDM1, FBXL10, JMJD2A, JMJD2B, JMJD2C

Demethylation

Heritable & Reversible

Courtesy of Yang Shi
(Harvard Medical School)

A Current Model for the Importance of Histone Tail Modification



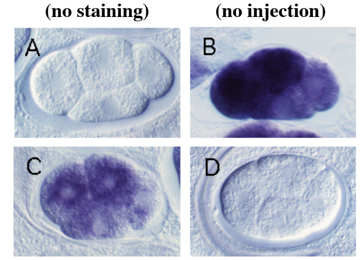
Reader Proteins:
- Coordinate and integrate various modification signals
- Bring additional enzymatic machinery to specific chromatin loci
- Determine substrate specificity of enzymes

RNA Interference (RNAi): Gene Silencing by Double-Stranded RNA (dsRNA)

Andrew Z. Fire — Craig C. Mello

Nobel Prize in 2006
(Physiology or Medicine)

Control (no staining) — *mex-3* Expression (no injection)



mex-3 Expression
(antisense injection) (dsRNA injection)

Fire *et al.* (1998) *Nature* 391: 806–811

Roots of RNAi: Very Strange Findings

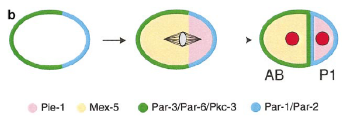
(sense and antisense RNA can both turn off the endogenous gene)

Experiment to Alter the Polarity of Worm Zygote

Molecule Injected	Embryonic Lethality (%)
ZC22 (*par-1*) antisense	52
ZC22 (*par-1*) sense	54
TS antisense	0
Z1 antisense	0
H20	0

***C. elegans* Embryogenesis**

(the first cell division is asymmetric)



"Surprisingly, injection of sense ZC22 RNA also induced *par-1* phenotypes. It is not clear what accounts for this effect … The basis for the sense effect is under investigation and will not be discussed further…"
(Guo & Kempheus. *Cell* 81:611-620, 1995)

Roots of RNAi: Very Strange Findings

(introducing a gene can turn off the endogenous gene)

Experiment to Alter Petunia Flower Color

wild type

introducing extra copies of a gene for chalcone synthase
(an enzyme required for flower pigmentation)

Instead of enhancing color, they got flowers with unusual patterns of colored and white tissues. This reflected the down regulation of both the transgene and the endogenous gene in the white tissues … (Napoli, *et al*. *Plant Cell* 2: 279–289, 1990)

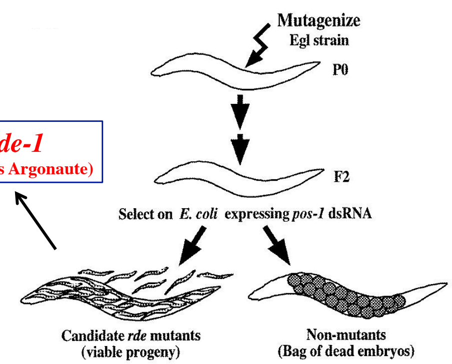
RNA Interference: Gene Silencing by Double-Stranded RNA (dsRNA)

- RNAi is a phenomenon in which the presence of double-stranded RNAs leads cells to specifically degrade any other RNAs with the same sequence, using small interfering RNAs (siRNAs).
- Viral replication often generates dsRNA, which can also be formed by aberrant transcription from genetic elements in the host genome.
- RNAi might have evolved as an RNA immune system and/or a mechanism to silence certain genomic regions and genes.
- RNAi is a promising new therapeutic for combatting diseases such as cancers or genetic diseases.

Searching for RNAi Mechanism

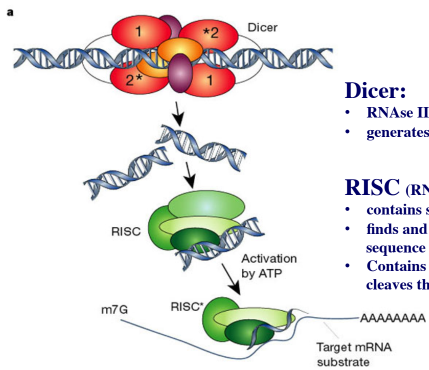
(feeding worms *E. coli* expressing a dsRNA or soaking them in dsRNA solution)

rde-1
(encodes Argonaute)



Tabara *et al*. (2003) *Cell* 99:123-132

RNA Interference Mechanism



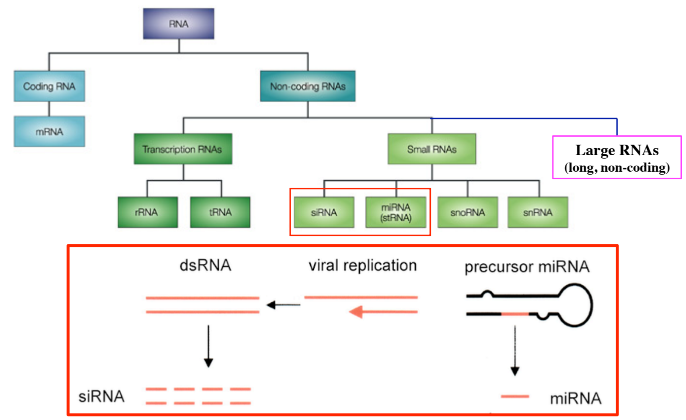
Dicer:

- RNAse III (dsRNA-specific endonuclease)
- generates siRNA (~22 nucleotides)

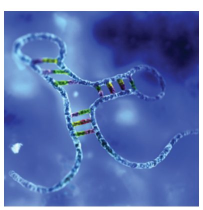
RISC (RNA-induced silencing complex):

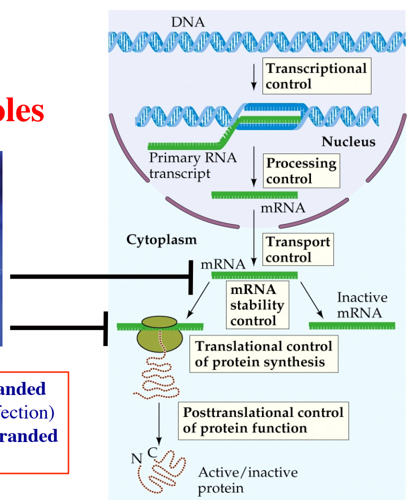
- contains siRNA
- finds and destroys mRNA of complementary sequence
- Contains Argonaute, an endonuclease, that cleaves the hybrid in the middle

An Expanding World of RNAs in Cells



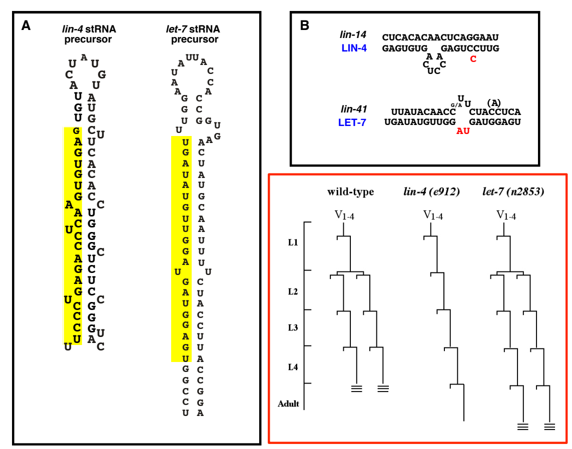
MicroRNA: Tiny but Big Roles





siRNA: exogenous, double-stranded
(taken up by cells, e.g. viral infection)
miRNA: endogenous, single-stranded
(i.e. made by cells themselves)

let-7 and lin-4 Are the Founding MicroRNAs



MicroRNAs Are Essential for Development

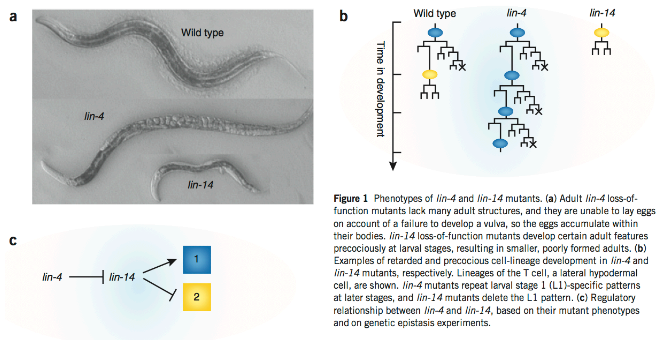
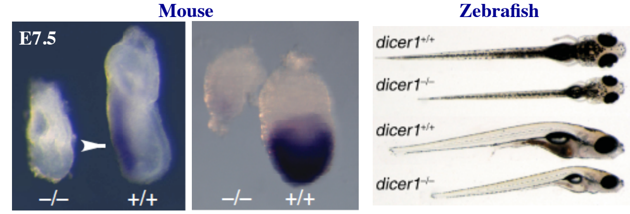


Figure 1 Phenotypes of *lin-4* and *lin-14* mutants. (**a**) Adult *lin-4* loss-of-function mutants lack many adult structures, and they are unable to lay eggs on account of a failure to develop a vulva, so the eggs accumulate within their bodies. *lin-14* loss-of-function mutants develop certain adult features precociously at larval stages, resulting in smaller, poorly formed adults. (**b**) Examples of retarded and precocious cell-lineage development in *lin-4* and *lin-14* mutants, respectively. Lineages of the T cell, a lateral hypodermal cell, are shown. *lin-4* mutants repeat larval stage 1 (L1)-specific patterns at later stages, and *lin-14* mutants delete the L1 pattern. (**c**) Regulatory relationship between *lin-4* and *lin-14*, based on their mutant phenotypes and on genetic epistasis experiments.

MicroRNAs Are Essential for Development
(mutating *Dicer1*, an enzyme essential for miRNA production)

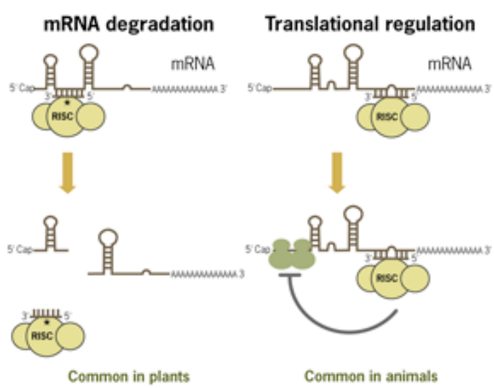
Mouse | **Zebrafish**



***T* (*bra*) Expression (primitive streak)** | ***Oct4* Expression (stem cells)** | **Growth Arrest**

Bernstein *et al*. (2003) *Nat. Genet*. 35:215–217
Wienholds *et al*. (2003) *Nat. Genet*. 35:217–218

How Do MicroRNAs Function?
(degrade mRNA or disrupt translation)



Genes that Can Jump: Transposable Elements (Transposons)

Courtesy of Cold Spring Harbor Laboratory Archives. Noncommercial, educational use only.

Discovered by Barbara McClintock in 1950s (Nobel Prize in 1983)

- **About 40% of our genomic DNA are transposons**
- **Are "specific" sequences of DNA (50–10,000 bp)**
- **Found in the genomes of many kinds of organisms**
- **Are structurally and functionally diverse**

Unusual Color Patterns

(each maize kernel is an individual embryo)

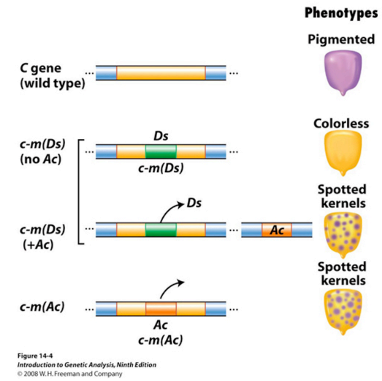
McClintock carried out many breeding experiments in maize. In this example, the expectation was for all the seed to be colorless. However, large and small patches of colored cells are apparent. She noted that these elements that induced this spotting were difficult to map to a discrete region in the genome, and that their presence could also affect the expression of other genes as well.

Two Types of Transposable Elements

(based on whether they can transpose on their own)

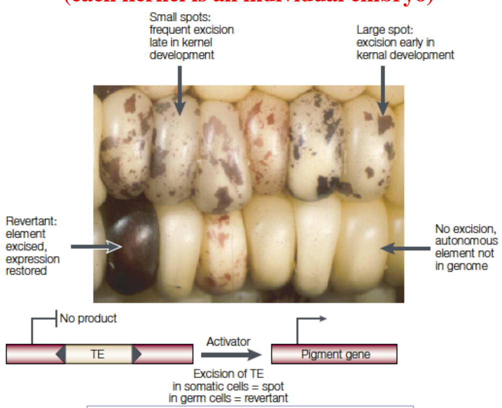
Non-autonomous Elements
(need the presence of another transposable element encoding a transposase)

Autonomous Elements
(transpose by themselves)



Unusual Color Patterns

(each kernel is an individual embryo)



Feschotte *et al*. (2002) *Nat. Rev. Genet*. 3:320-341

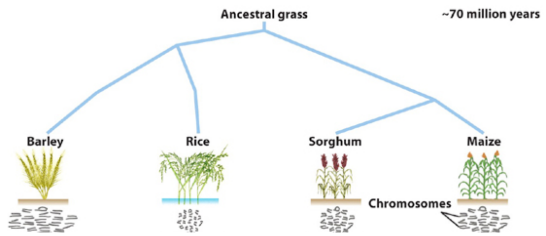
Types of Transposition

(transposons move from place to place in the genome)

- **In *cut-and-paste* transposition, an element is cut out of one site in a chromosome and pasted into a new site.**
- **In *replicative transposition*, an element is replicated, and one copy is inserted at a new site; one copy also remains at the original site.**
- **In *retrotransposition*, an element's RNA is used as a template to synthesize DNA molecules, which are inserted into new chromosomal sites.**

Transposon insertion can cause human diseases!
(Hemophilia B in *factor IX*, Neurofibromatosis in *NF1*, Duchenne Muscular Dystrophy in *dystrophin*)

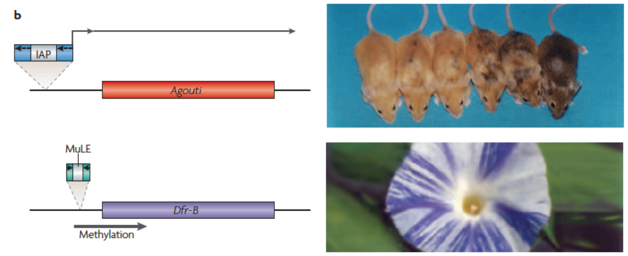
Differences in Genome Size Are Largely Attributable to Different Numbers of Transposons



5,100 Mb | 430 Mb | 700 Mb | 2,500 Mb

Except for the transposon regions, the different grasses show a great deal of synteny in their genomes.

Transposable Elements Can Disrupt the Epigenetic Landscape



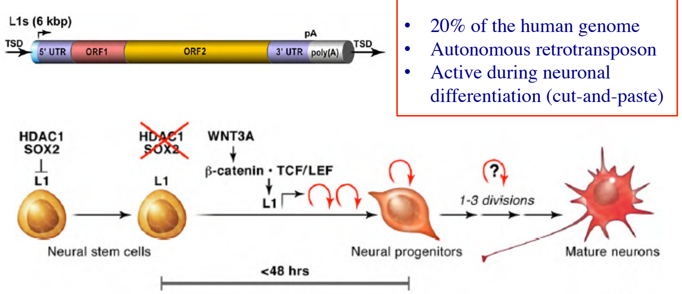
- In mice, the intracisternal A-particle (IAP) retrotransposon produces an outward-reading transcript that extends into the agouti coat-colour gene. The level of agouti transcript, and the colour of the coat, is subject to the epigenetic status of the retrotransposon and is heritable.
- In morning glory flowers, DNA methylation of a non-autonomous MuLE transposon can spread to the promoter of a flower-colour gene (Dfr-B), creating petal-colour streaks

Slotkin & Martienssen. (2007) *Nat. Rev. Genet.* 8:272-285

A Fun Idea: Explaining Neuronal Diversity?
(the human brain has 85-100 billion neurons)

L1 Elements

- 20% of the human genome
- Autonomous retrotransposon
- Active during neuronal differentiation (cut-and-paste)



The human neocortex has 15 quadrillion synapses

We are just beginning to understand how our genome works!

Key Concepts from This Lecture

- The central dogma, DNA > RNA > protein, can be modulated by feedback steps.
- Chromatin changes (DNA methylation and modifications of Histone proteins) influence whether or not a gene is transcribed.
- Non-coding small RNAs (like miRNAs) control development, disease state and metabolism by regulating gene expression post-transcriptionally (accumulation or translation of particular mRNAs).
- RNAi is the destruction of mRNAs corresponding to a dsRNA.
- Transposable elements can alter the expression of genes via insertion (causing mutations or changing epigenetic landscape).

The genome is dynamic!