Genetics, 1 of 6

01: The Science of Genetics

- **Genetics branches**

Classical Genetics: Also known as transmission genetics, initiated by Mendel's experiment with peas. Classical Genetics studies on the pattern of inheritance of traits.

Molecular Genetics: Study of genetic traits at molecular level, namely, how genes are regulated, genetic information is transcribed and translated, etc.

Population Genetics: Study of distribution and behavior of genes in a population, such as allele preference.

- **History of Genetics**

when	who	what
1865	Gregor Mendel	*Experiments on Plant Hybridization*
1910	Thomas Hunt Morgan	Chromosomes are carriers of genes
1941	E. L.Tatum and G. W. Beadle	Genes encode proteins
1953	J. D. Watson and F. Crick	DNA double helix structure
1961	Marshall W. Nirenberg	Genetic Code discovery
1968	M. Meselson and R. Yuan	Restriction Enzyme discovery
1996	Roslin Institute	Cloned Sheep

Forward Genetics: the process when a known phenotype is observed, the gene is attempted to be isolated for further studies.

Reverse Genetics: When a gene's DNA sequence is known, the gene can be knocked down in vivo and the phenotype is studied.

02: Mendelian Genetics and its extensions

Dominant: An allele that expresses its trait regardless of the other allele, usually designated as upper-case letters, for example, A.

Recessive: An allele that can not express its phenotype when a dominant allele presents, usually designated with lower-case letters, such as a.

Gamete: A mature reproductive haploid cell that is specialized for sexual fusion.

Genotype: The composition of two alleles of a gene or multiple genes, e.g. genotype RR, Rr or rr, or RrYy, RRyy, etc.

Phenotype: The outcome of the genotype, the expressed "observable" trait of a gene, e.g., green color and round shape of a pea seed.

Mendel's Laws:

Law of Segregation: There are two alleles for each gene; during the gamete formation, the two alleles of one gene segregate from each other independently

Law of Independent Assortment: During the gamete formation, genes from different chromosomes assort independently and combine randomly

Extensions to Mendel's Laws:

Maternal Inheritance; Co-dominance; Epigenetics; Multiple alleles inheritance

Problem solving skills:

a) Punnette Square:

write down both the male and female gamete type and ratio, list all the combination in Punnette Square.

b) Branch Method:

Start with one gene first, list all possible gametes and their ratio, these gametes randomly assort with a second gene, and write down each possible combinations. Multiply the ratio to obtain the final ratio.

03: Genes and Chromosomes

Nucleosome: A bead-string structure formed by DNA and histones, the basic DNA structure in a chromatin.

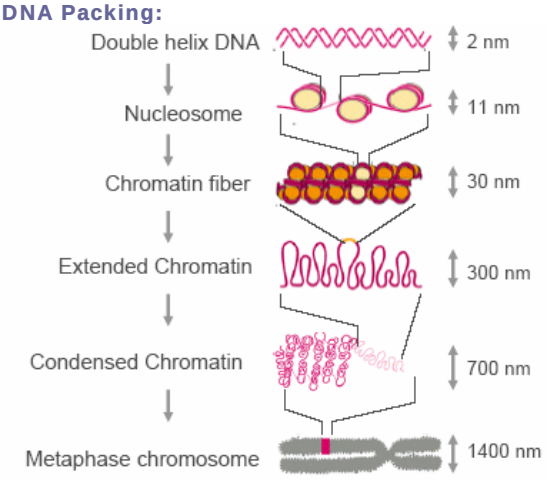
Chromatin: The general structure of any chromosome, the basic units are nucleosomes.

Chromatid One-half of a replicated chromosome. structure of arms, a centromere and two telomeres

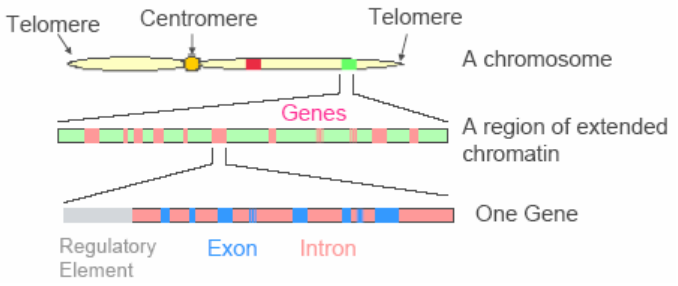
Euchromatin: Chromatin region stained lightly, usually are lightly packed and transcriptionally active.

Heterochromatin: Chromatin region stained dark, usually are heavily packed and transcriptionally inactive.

DNA Packing:



Genome Organization:

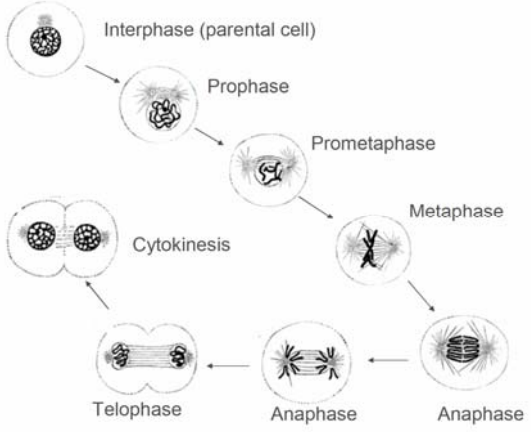


04: Mitosis and Meiosis

Centrosome: Made of two centrioles, it is the main microtubule organizing center (MTOC) and a regulator of cell-cycle progression.

Mitosis: The series events when a somatic cell is divided into two identical daughter cell. It normally contains prophase, prometaphase, metaphase, anaphase and telophase.

Meiosis: The series events when a cell is divided twice while its DNA only replicate once, resulting 4 haploid progeny cells –germ cells.



Genetics, 2 of 6

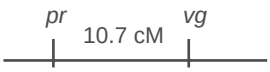
05: Genetic Mapping—Linkage and Recombination

Genetic markers: known DNA sequences or genes that encode phenotypes, with defined position on chromosome so they can serve as reference for other genes.
Linkage: Genes or markers on same chromosome that are within close distance are transmitted to next generation concomitantly.
Recombination: Linked genes or genetic markers from different chromosomes cross over during meiosis to produce new gamete types.
Crossover: During meiosis homologous chromosomes pair with each other and chromatids cross over to exchange genetic materials.
Testcross: When an individual is crossed to homozygous recessive pure breeding, the offspring phenotype reflects the gamete types of the tested individual.
Genetic Distance: The relative distance of two markers on same chromosome. Usually measured by recombination frequency, or map unit.
Linkage Detection:
Use Chi square Method to detect linkage

$$\chi^2 = \Sigma [(o-e)^2/e]$$

Df = n -1, calculate p value

Two-point testcross: Using one parent with homozyhous recessive alleles to test the other parent's genotyping, the offspring phenotype reflects the tested parent's genotype.



06: Pedigree and Sex Determination

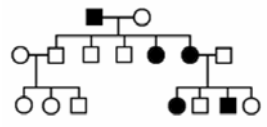
Sex Chromosomes: Chromosomes that are specialized for determining sex phenotype of an organism, normally males and females have different sex chromosomes.
Autosomal Chromosomes: Non-sex chromosomes. Usually males and Females have same autosomes.
XX-XO System:
- Female: XX – a pair of X chromosomes.
- Male: XO – one X chromosome
- Example: grasshopper

XX-XY system
- Female: XX – a pair of X chromosomes
- Male: XY – one X chromosome, one Y chromosome
- Example: humans, fruit flies (Drosophila)

ZW-ZZ system
- Female: ZW – one Z chromosome, one W chromosome
- Male: ZZ – a pair of Z chromosomes
- Example: Chicken

X-linked Recessive Traits:
Appear more often in males.
Not passed from father to son
Passed from carrier mother to son
Skip generations

X-linked Dominant Traits:
Appear in both males and females.
Affected males must have affected mother
Affected males pass the trait to all of their daughters
Do not skip generations



Y-linked Traits:
Appear only in males
Pass from father to son
Do not skip generations

07: Genetics of Mitochondrial and Chloroplast

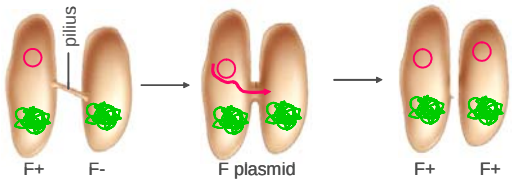
Heteroplasmy: A cell can have some mitochondria that have a mutation in the mtDNA and some that do not. –The mitochondria are not all the same in one cell.
Homoplasmy: A cell that has a uniform collection of mtDNA: either completely normal mtDNA or completely mutant mtDNA.
H strand: One strand of mtDNA is heavier than the other because it contains higher guanidine content.
L strand: cytosine-rich mtDNA strand, replicate later than H strand.
D-loop: a region of 1121 bp that contains the origin of replication of the H-strand (O_H) and the promoters for both strands P_L and P_H.
Endobiosymbiont Hypothesis: Origin of mitochondria and chloroplast is ancient free-living prokaryotes. These organisms invaded primitive eukaryotes and established symbiotic relationship.
Evidence:
- MtDNA and cpDNA are cicular, similar to bacteria
- Similar ribosomes
- Similar in size
- Similar double membranes
- Similar transcription mechanism

Extranuclear Inheritance: Traits encoded by mtDNA or cpDNA (both of which are located in cytoplasm, extranucleus) are determined by mother's genotype only. The genetic material is contributed by female parent only.
Rules of extranuclear inheritance:
- Results from reciprocal cross are different
- Extranuclear genes cannot be mapped to chromosomes in the nucleus
- Ratios typical of Mendelian inheritance is not found
- Replacing a nucleus with a different set of genes wouldn't affect the extranuclear inheritance traits.

Maternal Effect: Determination of characters of the progeny by the female parent. This can be caused by both extranuclear genes or nuclear genes.

08: Bacterial and Viral Genetics

Nucleoid: A region where the bacterial chromosome is located, no obvious boundary.
Plasmid: small circular DNA molecules found in bacteria cytoplasm, may carry certain genetic material.
Conjugation: Two bacteria exchange genetic material via contact.
Transformation: DNA from environment is picked up by a bacterial cell.
Transduction: Bacteria gain new genetic material upon virus infection.
F factor: also called F plasmid, contain information that enable a cell to make pilius and fuse with another recipient cell. Required for conjugation.



conjugating	results
F+ x F-	Two F+
Hfr x F-	Hfr and F-
F' x F-	Two F'

Genetics, 3 of 6

09: Genes and Environment

Phenotypic Variance: VP, Variability of a trait, calculated by statistical method for a group of individuals.

$$V_P = V_G + V_E + COV_{GxE} + V_{GxE}$$

Genetic Variance: VG, Difference of the Vp may be caused by different genotype within the group.

$$V_G = V_A + V_D + V_I$$

Environmental Variance: VE, Different environments experienced by individual in the group contribute to the phenotype Vp.

COV_{GxE}: Covariance of genetically-caused variance and environmentally-caused variance.

V_{GxE}: Genetic-environment interaction

Broad sense heritability: *H*2, The proportion of total phenotypic variation due to all genetic effects. It reflects the importance of genes in producing the difference.

H^2 = Total genetic variance / Phenotypic variance

$$= V_G / V_P = \frac{V_A + V_D + V_I}{V_p}$$

Narrow sense heritability: *h*2, the proportion of total phenotypic variation that is due to the additive effects of genes. It reflects natural selection.

h^2 = Additive genetic variance / Phenotypic variance
= VA/ VP

Penetrance: The probability of showing a phenotype given a genotype.

Expressivity: the variations of the expressed phenotype.

Co-dominant: Both allele each exhibits its own phenotype when they are heterozygous.

10: Quantitative Genetics

Quantitative traits: Continuous traits are often measured and given a quantitative value, so its other name is called quantitative traits.

QTL: Quantitative trait loci, genes controlling the quantitative traits. One trait may be controlled by multiple QTLs, each of which segregate according to Medelian's Law.

Mean: The average value of a distribution.

Variance: a measure of the variability of the distribution.

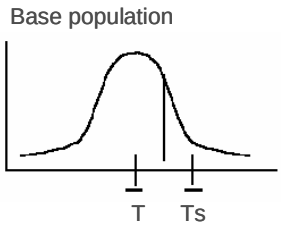
Standard Deviation: a measure of the variability of the distribution.

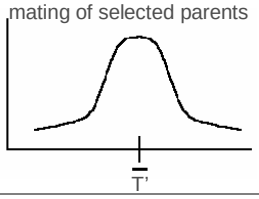
$$\text{Mean} = \bar{x} = \Sigma x_i / n$$

$$\text{Variance} = S^2 = \Sigma (x_i - \bar{x})^2 / (n-1)$$

$$\text{Standard Deviation} = s = \sqrt{s^2}$$

Response to Selection:
T: Population mean
Ts: mean for selected population
Selection differential
S = Ts - T
Selection response
R = T′ – T
T′: Mean of the pffspring
h^2 = R / S → R = h^2*S





Estimation of offspring phenotype:

Formula for predicted offspring value:

$T_o = T + h^2[(T_F+T_M)/2 - T]$

11: DNA: the Chemical Basis of Genetics

Nucleoside: Compounds with a ribose or deoxyribose and a base. Base includes four types G, C, A and T (or U).

Nucleotide: Compounds with a ribose or deoxyribose, a base and a phosphate group, it is the basic building unit for DNA (monomers of DNA).

Semiconservative Replication: when DNA replication finish, the newly-synthesized strand form a double strand with its template "mother" strand – therefore it is called "semi-conservative replication".

Replication Origin: Replication starts from particular sequence (origins) in E. coli and yeast.

Okazaki Fragment: short DNA fragments produced by the lagging strand during DNA replication.

RNA differ from DNA
RNA is single strand, DNA double strand
RNA contains G, C, A and T; DNA contains G, C, A and U.
RNA is normally short, DNA is long molecule.
RNA contains ribose, DNA contains dexoyribose.

DNA Double Helix Structure:

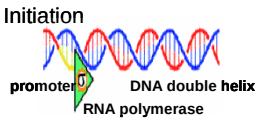
- The two strands wrap around each other and form double helix structure
- The DNA fiber measure 2.0 nm in diameter; 3.4 nm for each turn
- Each turn contains 10 base pairs
- The two strands run in opposite directions (anti-parallel)

12: Transcription, translation and genetic codon

Initiation: RNA polymemrase binds to promoter

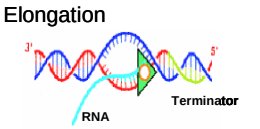
Elongation:

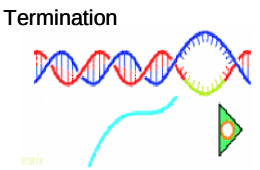
- Genes are read from 3′ to 5′ direction for the template strand
- RNA is synthesized from 5′ to 3′ direction
- Genetic information is copied to RNA according to base-pairing principle.



Termination:

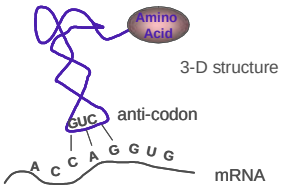
- When polymerase complex encounters a termination signal on DNA template, the polymerase dissemble and RNA molecules are released.





tRNA:

- 4 arms and three loops: D-loop, Tyc loop, anticodon loop and their arms, and amino acid arm.
- Anticodon reads information from mRNA by base-pairing
- Amino acid arm receives a matching amino acid



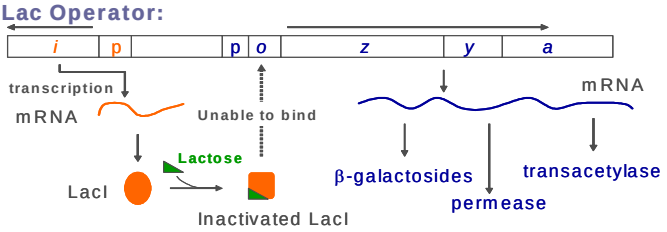
Genetic Codon:
Continuous, no comma
Non-overlapping, unambiguous
Almost universal with few exceptions
Degeneration
Special start codon (AUG) and stop codons (UAA, UAG and UGA); Wobble

Genetics, 4 of 6

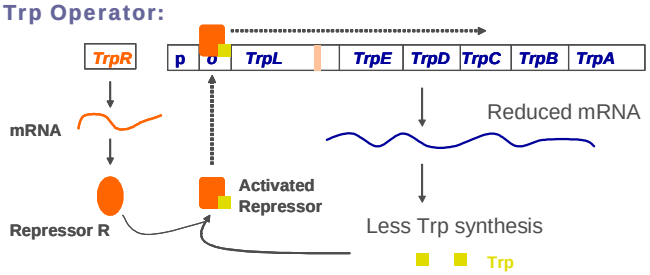
13: Control of Gene Expression in Prokaryotes

Repressor: Transcription factors that negatively regulate RNA polymerase activity.
Activator: Transcription factors that positively regulate RNA polymerase activity.
Structure genes: The genes in an operon that encodes the enzymes necessary for a metabolic pathway.
Regulatory genes: The genes in an operon that play regulatory roles, acting as either repressors or activators.

Lac Operator:

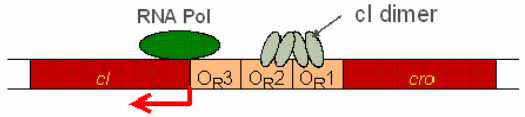


Trp Operator:

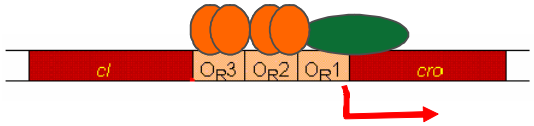


Phage Cycle control:

- High cI, low Cro → lysogeny

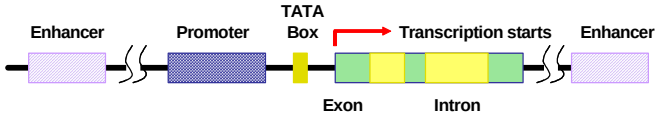


- High Cro, low cI → lysis

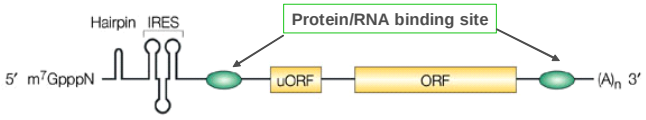


14: Regulation of Gene Expression in Eukaryotes

Activator: Transcription factors that binds to DNA and activate gene expression.
Co-activator: do not bind DNA but bind to activator to enhance activator function.
Repressor: Transcription factors that binds to DNA and repress gene expression
Co-repressor: Factors that coordinate with a repressor to inhibit gene expression.
Transcriptional Control:



Translational Control:

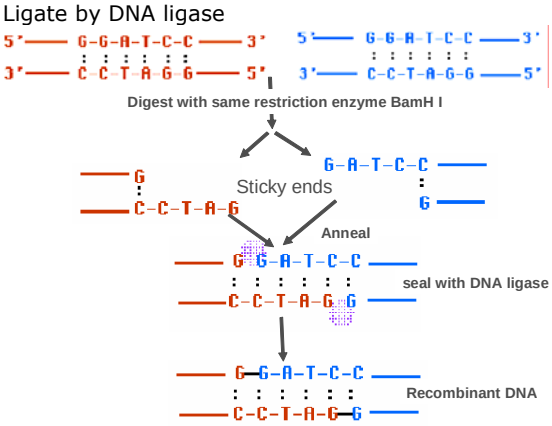


15: The Techniques of Molecular Genetics

cDNA: cDNA: complementary DNA, the DNA synthesized from a mature mRNA template. cDNA is often used to clone genes.
DNA vector: an agent that can carry a DNA fragment into a host cell and ensure its replication and sometimes, expression.

Recombinant DNA: Key Steps

Step 1: DNA isolation:
Genomic DNA or cDNA, or PCR amplified DNA fragment
Step 2: Cut and ligate:
Cut by restriction endonuclease
Ligate by DNA ligase



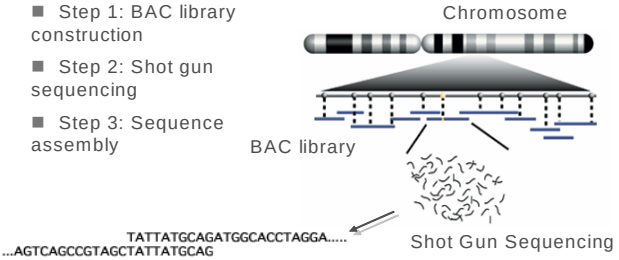
Step 3: transformation and identification
Selection through antibiotics resistance genes (markers)

16: Genomics and Proteomics

BAC: bacterial artificial chromosomes, which is based on F plasmid and can tolerate large inserts, widely used in genomic DNA library construction.
BAC library: large genomic DNA fragments ligated into BAC vectors.
Shot gun sequencing: Genomic DNA from a BAC clone is fragmented into smaller size, typically 1kb, and inserted into a sequencing plasmid. These sequences are then assembled based on the overlapping fragments.

Procedure Overview

- Step 1: BAC library construction
- Step 2: Shot gun sequencing
- Step 3: Sequence assembly



Functional Genomics:

Comparative genomics
- Sequence homology
- Phylogenetic Tree

Structural genomics
- Protein expression, purification and crystallization

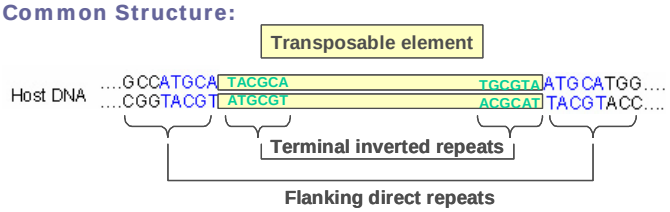
New experimental methodologies
- DNA microarray

Proteomics: two important tools:
- 2D gel electrophoresis
- MALDI-TOF Mass Spectrometry

Genetics, 5 of 6

17: Transposable Genetic Elements

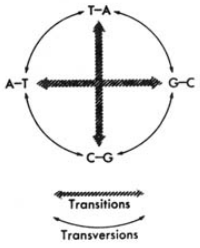
Transposable genetic elements: A heterogeneous class of genetic elements that can insert at new locations on chromosomes, also known as "jumping genes", "transposons".
Retrotransposon: Transposable elements that are translocate through RNA intermediates and "copy and paste" mechanism.
DNA transposon: Transposable elements that move wround the genome mainly by a "cut and paste" mechanism.
Transposon Tagging: a classical method for obtaining mutants and clone genes in plants and other organisms.
Common Structure:



Transposition Mechanism:
Cut and Paste – DNA transposons
Copy and Paste --retrotransposons

18: Mutation and DNA Repair

Transition mutation: One purine-pyrimidine base pair change to another purine-pyrimidine base pair, e.g., AT to GC.
Transversion mutation: One purine-pyrimidine base pair change to a pyrimidine-purine base pair, e.g., AT to TA.
Missense mutation: A mutation in which a base pair change in the DNA causes an altered amino acid in protein so that the phenotype changes.



Nonsense mutation: A mutation that changes an amino-acid encoding codon into a stop codon. This results in truncated protein.
Neutral mutation: A mutation that has no phenotypic effect. Often it is change from one amino acid to another with similar chemical property.
Silent mutation: DNA sequence changed but the encoded amino acid remains the same due to degenerative of codons.
Frame-shift mutation: Addition or deletion of one or a few base pairs that leads to a change in reading frame, Usually results in non-functional protein.

DNA Repair

Direct Correction of Lesions:
DNA polymerase:
- 5'→3' polymerase
- 3'→5' exonuclease

Light Repair: Thymidine dimmer repaired by light and photolyase.
Repair of Alkylating Damages: Alkylated DNA such as O6-methyl-guanidine can be repaired by an enzyme which removes the added methyl group from guanidine.
Repair involving base excision:
- Short-patched Repair:
- Repair by glycosylase and AP endonuclease
- Mismatch repair
- SOS response

19: Dissection of Gene Function: Mutational Analysis of Model Organisms

Homologous recombination: DNA cross over occurs between two homologous DNA molecules, leading to exchange of the paired DNA.
Chimeric mouse: a transgenic mouse that carry some sections from transgene and other sections from its orginal mother.
Forward Genetics:
Step 1: Mutant Identification.
If no mutant, make one.
- Transposon tagging
- Chemical mutagenesis

If two mutants have same phenotype:
- Cross the mutant to determine gene number

Step 2: Genetic Mapping: Locate a gene on chromosome
Step 3: Mutant Allele Tagging: Transposon tagging
Step 4: Gene Isolation
Reverse Genetics: Three approaches:
Gene Silence: RNAi or shRNA
Homologous recombination-based gene disruption:
- Disrupt the gene structure

Transposon-tagged stocks –genomics approach

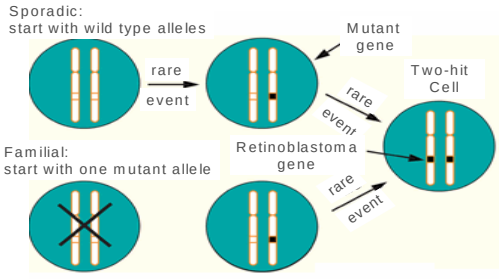
Gene Knockout in mice: homologous recombination.
- Introduce construct into ES cells
- Grow up the transfected ES cells and introduce it into isolated blastocysts.
- Place the blastocycsts into a pseudo-preganant female mouse
- Obtain the offspring chimeric mouse
- Cross with wild type but different coat colored mouse
- Select for knockout mouse.

20: The Genetics Basis of Cancer

Tumor: tissue composed of cells that deviate from normal program of cell division and differentiation.
Benign tumor: tumor cells remain together in a single mass and do not invade or disrupt surrounding tissues.
Malignant tumor: tumor cells invade and disrupt surrounding tissues (and are diagnosed as cancer).
Metastasis: spread of malignant tumor cells throughout the body (typically through the blood and lymphatic system).
RNA Tumor virus: retrovirus that cause tumors in vivo. Usually carry cell-originated oncogene.
DNA Tumor virus: DNA virus that can stimulate host DNA replication to replicate its own genome. Consequently the host cell proliferation increase and oncogenesis occur.
Tumor suppressors: genes that reduce the probability of tumorigenesis. A mutation or deletion of such a gene will increase the probability of the formation of tumor.

Genes mutated in cancer:
Proto-oncogene → oncogene → v-onc vs c-onc
Tumor suppressor genes

Two-hit model:



Genetics, 6 of 6

21: Cytogenetics and Chromosomal Mutation

Karyotyping: a standardized arrangement of all the chromosomes of a cell.

Chromosome Banding:

Q banding: stained with a fluorescent dye such as quinacrine

G banding: produced by staining with Giemsa after digesting the chromosomes with trypsin

C banding: treated with acid and base, then stained with Giesma stain

Variation in Chromosome number:

Aneuploidy:

- Nullisomy, 2N-2
  -- the loss of both pairs of homologous chromosomes
- Monosomy, 2N-1
  -- the loss of a single chromosome
- Trisomy, 2N+1
  -- the gain of an extra copy of a chromosome
- Tetrasomy, 2N+2
  -- the gain of an extra pair of homologous chromosomes.

Monoploidy: (haploidy): one set of chromosome. Most gametes are haploid.

Polyploidy: lethal in human, more tolerant in plants, caused by failure in chromosome segregation.

Variation in chromosome structure:

- Deletion
- Duplication
- Inversion
- Translocation

Fragile X syndrome: The most common form of inherited mental retardation, visible satellite regions at the ends of the long arms of metaphase **X** chromosomes , which is caused by a long series of CGG triplet repeats within a gene termed FMR-1.

Philadelphia chromosome: a specific chromosomal abnormality that is associated with chronic myelogenous leukemia (CML), resulting from reciprocal translocation between chromosome 9 and 22.

22: Developmental Genetics

Cell Determintation: The process by which a cell and its descendants become committed to a particular cell fate (e.g. myogenesis). This can only be established by perturbing the embryo experimentally.

Cell Differentiation: The expression of specific gene products, morphology, etc., characteristic of a particular fate (e.g. actin, myosin in a muscle cell).

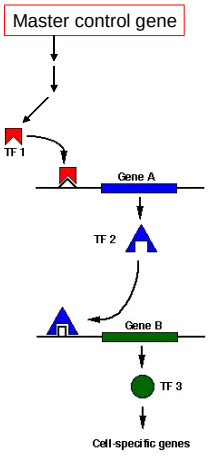
Theory of Differential Gene expression: master control gene

Totipotent: The ability of a single cell, usually a stem cell, to divide and produce all the differentiated cells in an organism, including extraembryonic tissues.

Pluripotent: The ability of giving rise to most tissues of an organism. These cells can differentiate into cells derived from the three germ layers.

Multipotent: The ability to produce only cells of a closely related family.

Unipotent: The ability to produce only one cell type, but have the property of self-renewal which distinguishes them from non-stem cells.



23: Population Genetics

The Hard-Weinberg Law: The genotype frequencies and gene frequencies of a large, randomly mating population remain constant provided immigration, mutation and selection do not take place.

Assumptions:

- Population is infinitely large
- Mating is random
- No natural selection
- No mutation.
- No migration.

Predictions:

Allele frequencies do not change over generations.

At equilibration: $p^2 + 2pq + q^2 = 1$

- p = allelic frequency of A
- q = allelic frequency of a
- p^2 frequency of AA
- 2pq frequency of Aa
- q^2 frequency of aa

Four Forces driving evolution:

Mutation:

- A → a rate: μ; a → A: ν
- At equilibrium: $\hat{p} = \frac{\mu}{\mu + \nu}$ $\hat{q} = \frac{\nu}{\mu + \nu}$

Genetic Drift:

Major cause is sampling error

Effective population size:

- equal sexed and equal reproductivity: Ne = N
- Otherwise:

Ne = (4 x Nf x Nm)/ (Nf + Nm)

Nf and Nm = numbers of breeding females and males.

Sampling variance:

- Sampling variance: $s_p^2 = pq/2Ne$
- Standard error: $s_p = \sqrt{(pq/2Ne)}$
- 95% confidence limit = $p \pm 2s_p$

Migration:

Migrants from population I: $f(A) = p_x$

Residents from population II: $f(A) = p_y$

After migration:

- $f(A) = m\,p_x + (1-m)p_y$
- $\Delta p = m(p_x - p_y)$; m = migrants/(migrants + residents)

Natural Selection: Increases or decreases genetic variation depending on the environment

24: Evolutionary Genetics

Fitness and Selection:

Selection coefficient (s) = 1 - *W*

After selection;

- Mean fitness = $1- sq^2$
- New frequency $p' = p/(1-sq^2)$
- Changes of gene frequency: $\Delta p = spq^2/(1-sq^2)$

When p is small: $\Delta p = sp/(1-s)$

When p is large: $\Delta p = sq^2$

Heterozytoe superiority: assume selection coefficient of AA is s and aa is t:

$P^* = s/(s+t)$

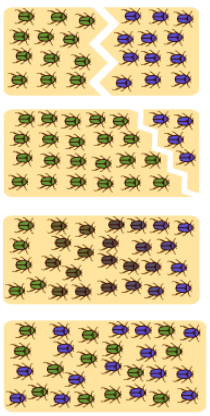
Four Modes of Speciation:

Allopatric: geographically isolated populations

Peripatric: mostly geographic,

Parapatric: a continuously distributed population

Sympatric: non-geographic, within the range of the ancestral population



Hypothesis:

- Gradualism
- Punctuated equilibrium