Chapter 2 – Lecture Notes

13-04-14 5:22 PM

Chromosome and Cellular Reproduction

- Cytologists and geneticists in the early 20th century determined that genes were carried on chromosomes
- In order to understand heredity, we need to understand the behavior of chromosomes during the cell cycle, mitosis and meiosis

Eukaryotic Chromosomes

- Genetic material of eukaryotes is often distributed among multiple linear chromosomes
- Chromosomes: a linear structure composed of a single DNA molecule complexed with protein. Each species has a characteristic number of chromosomes in the nucleus of a cell

Haploid and Diploid

- Haploid
    - N
    - One copy of genetic material subdivided into chromosomes
    - Non-homologous
- Diploid
    - 2N
    - Two copies of genetic material subdivided into chromosomes
    - Homologous
- **Diagram on Slide**
- Eukaryotes have 2 copies of each type of chromosome in their nuclei
    - Diploid or 2N
- Eukaryotes that have only one set of chromosomes…. Usually gametes
    - Haploid or N
- Example: human somatic cells have 46 chromosomes, while our gametes only have 23
    - $2N = 46, N=23$

- ***Homologous Chromosomes***
- In diploids, chromosomes that contain the same genes and that pair during meiosis are called homologous chromosomes
    - Each member is called a homolog
    - One homolog is inherited from each parent
        - In humans, one is the paternal homolog while the other is maternal homolog

- Chromosomes that contain different gene and that do not pair during meiosis are called non-homologous chromosomes

- ***Autosomes and Sex Chromosomes***
- Chromosomes that differ in the number or morphologically between two sexes are called sex chromosomes
  - One sex may have a matched pair of sex chromosomes, while the other sex may have an unmatched pair or just a single sex chromosome
- Chromosomes other then sex chromosomes are called autosomes
- **Diagram on Slide**

Variation in Chromosomes

- Chromosomes may also vary in length
  - In humans there is a wide range of chromosome lengths
  - In mice, all of the chromosomes are of similar size

Karyotype

- A complete set of all the metaphase chromosomes of a cell are calle da karyotype
- Notice for this human karyotype there are 22 autosomes and one of each X and Y
- Conventions
  - Largest chromosome is #1
  - Homologues are paired
  - Sex chromosomes often shown at the bottom together

Additional Nomenclature for Chromosomes

- When displayed in a karyotype the chromosomes are arranged from largest to smallest, and each chromosome is oriented with the largest chromosome down and the smallest on top
- The short arm is designated as p and the long arm is designated as q
  - Used for the location of markers along the chromosome
  - Eg. BRCA1 is at location 17q21 (on chromosome 17 on the long arm in region 21)

Interphase and the Cell Cycle

- In both uni and multicellular eukaryotes cellular reproduction is necessary for growth
- This is a cyclical process that includes growth and cell division (the cell cycle)
- The cell cycle is divided into two stages: interphase and cell division
- **Diagram on Slide**

Interphase

- The interval between cell division and is subdivided into
  - G1 –presynthesis stage where the cell prepares for DNA replication and grows
  - S – stage of DNA synthesis or replication
  - G2 – postsynthesis stage where the cell prepares for the onset of mitosis or cell division
  - G0 – a nondividing quiescent state some cells enter
    - These cells are non invading

Cell Cycle

- The length of G1 can vary among the different cell types
  - Minuets in rapidly dividing cells such as cancer and fetal cells and years for some differentiated cells such as nerve cells
- In contrast for S and G2 are the same for most cells
  - Once started a cascade of events occur to transition the cells through the stage

G0

- Series of controls determine if cells continue through G1 to S or withdraw from the cell cycle into G0
- G0 cells are metabolically active and may re-enter cell cycle if stimulated
- Cancer cells avoid G0 or spend only a short time in G0 compared to other cells

Cell Cycle Check Points

- The cell cycle has 3 major checkpoints in the mitotic cell cycle
- Controlled by the activity of kinases and cyclins
- **Diagram on Slide**

How do you Determine Length of the Stages

- The duration of the S phase can be determined with a pulse chase experiment

Mitosis

- Occurs in both haploid and diploid cells
  - Outcome – two cells, each with the same number and type of chromosome as the original parents cell
- The process is divided into 4 cytologically distinguishable phases

    - Prophase, metaphase, anaphase, and telophase
- **Diagram on Slide**

Prophase

- Chromosome condensation occurs
- Mitotic spindle apparatus assembles
- Nuclear envelope breaks down

Mitotic Spindle

- During prophase, the mitotic spindle assembles outside of the nucleus from microtubule organizing centres (MTOCS)
    - In animal cells the MTOCs are the centrioles which divide prior to S phase
        - The pair of centrioles separate during prophase aligning themselves at opposite poles of the cell
        - From each centriole a sunburst of microtubules are projected across the cytoplasm
        - The centrioles establish the axis for the upcoming mitotic division
- These microtubules bind to the centromere regions of the chromosomes forming a multiprotein complex called a kinetochore

Metaphase

- Begins when a nuclear envelope has completely disappeared
- Microtubules orient each centromere in one plane halfway between the 2 spindle poles
    - This place is called the metaphase plate
- **Diagram on Slide**

Anaphase

- Begins when joined centromeres or sister chromatids separate
    - Gives rise to 2 daughter cells
- Daughter chromosomes are pulled to each pole by a shortening of the microtubules attached to each kinetochore
- Shortest stage of mitosis
- Centromeres pulled first, arms trail behind

Telophase

- Migration of chromosomes to poles complete

- Chromosomes begin to uncoil
- Nuclear envelope reforms around chromosomes
- Spindle apparatus disappears

Cytokinesis

- Divisions of the cytoplasm
- Completed by the end of telophase
- In animal cells, the cell is constricted at the middle until 2 daughter cells are produced
- In plant cells with rigid cell wall, a new cell membrane and cell wall is produced between the 2 nuclei
- **Diagram on Slide**

Gene Segregation in Mitosis

- Following mitosis each daughter cell will have the same number of chromosomes as the parents cell
    - A diploid cell undergoing mitosis will produce 2 daughter cells that are both diploid
    - A haploid cell undergoing mitosis will produce 2 daughter cells that are both haploid
- In other words no segregation of genes occurs in mitosis

Meiosis

- Primary purpose
    - Reduce chromosome number from diploid to haploid in germ cells and spores
    - Create genetic variation in the process
        - Independent assortment and crossing over
- **Diagram on Slide**
- Occurs only in certain stages of development of an organisms life cycle
- Meiosis – two successive nuclear divisions of a diploid nucleus following one DNA replication that result in the formation of haploid gametes or spores having one half the genetic material of the original cell
- In animals results in the formation of haploid gametes which is part of the process of gametogenesis
- In plants meiosis results in the formation of haploid meispores. This divides mitotically to produce a multicellular gametophyste
- Before meiosis the chromosomes are replicate in S phase
- Meiosis consists of 2 division

  - In meiosis I the homologs separate, creating daughter cells that are haploid. This is the reductional division
  - In meiosis II the sister chromatids separate
- Each division usually produces 2 daughter cells
  - Thus a single diploid cell produces 4 haploid cells

Stages of Meiosis I

- Prophase I
- Metaphase I
- Anaphase I
- Telophase I
- These stages have the same names as the stages of meiosis, but there are subtle differences in what occurs during these stages in meiosis

Prophase I

- Prohphase I is the most complicated stage of meiosis and is the pivotal stage where the homologs find each other, pair together and exchange genetic information
- Due to this complexity, prophase I is subdivided into 5 stages
  - Leptonema
    - This is the initial stage of prophase
    - The extended chromosomes begin to coil and condense and now become visible as long threads in the nucleus
    - Once leptonema begins, the cell is committed to the meiotic process
    - Homology search begins
    - **Diagram on Slide**
  - Zygonema
    - The chromosomes continue to shorten
    - Homologs find each other and align along their lengths
    - Pairs undergo synapsis – the formation along the length of the chromatids of a zipper like structure called the synaptonemal complex
    - Chromosome pairs are maximally condensed before the synaptonemal complex forms
    - **Diagram on Slide**
  - Pahynema
    - Begins when synapsis is complete
    - Each synapsed set of homologs consists of 4 chromatids

            - These are now called a bivalent or a tetrad
        - Crossing over occurs during this stage
            - Reciprocal exchange of genetic material from one homolog to the other
            - If there are genetic differences between the homologs, then the new genetic combination will be produced
                - This is genetic recombination
        - **Diagram on Slide**
    - Diplonema
        - The synaptonemal complex begins to disassemble and the homologous chromosomes begin to separate
        - Chromosomes are still attached at the points where crossing over has occurred
            - These sites are called chiasma, cross shaped structures visible under the microscope
        - **Diagram on Slide**
    - Diakinesis
        - Here the nucleolus and nuclear envelope break down
        - The spindle apparatus forms
        - **Diagram on Slide**

How do the Homologs find Each Other?

- By the telomeres
- The telomeres are clustered on the nuclear envelope in one area producing a bouquet like structure
- The telomeres of the homologues come together in this region initiating synapsis

Why the Need for a Chiasma?

- Essential for the proper segregation of homologs during meiosis I
- Holds the homologs together until anaphase I
- **Diagram on Slide**

Meiosis I and Meiosis II

- **Diagram on Slide**

Consequences of Meiosis

- Generate 4 haploid cells

    - Two divisions preceded by only one cycle of DNA replication (S phase)
- In metaphase I an maternally and paternally derived homologs have an equal chance of aligning on one side or the other
    - This is called independent assortment
    - Many combos are possible creating immense variation of gametes

<u>Independent Assortment</u>

- General formula for the number of arrangments at the metaphase plate is 2^n-1 (n = number of chromosome pairs)
- General formula for the number of possible chromosome combinations in the resulting gametes is 2^n

<u>Further Variation in the Gametes</u>

- Crossing over between maternal and paternal homologs during prophase I generates even more combinations
    - Crossing over occurs between the homologs in every meiosis at random locations along the chromosomes
    - Large amount of variation is possible

<u>Meiosis in Animals</u>

- Most multicellular animals are diploid through most of their life cycle
- Meiosis produces haploid gametes
- Fusion of 2 gametes produces a diploid zygote
    - Gametes are only produced in specialized cells
- **Diagram on Slide**

<u>Spermatogenesis</u>

- In males, sperm cells (spermatozoa) are produced within the testes
    - Primary spermatogonia are the primordial germ cells
    - They divide by mitosis to give rise to secondary spermatogonia, which then transform into spermatocytes and enter meiosis I
- **Diagram on Slide**

<u>Oogenesis</u>

- In females, the ovary contains the primordial germ cells (primary oogonia)
- Primary oogonia divide mitotically to give rise to secondary oogonia

  - These then transform into primary oocytes
  - The diploid primary oocyte goes through meiosis I with unequal cytokinesis to produce 2 cells
    - A large cell called the secondary oocyte
    - A very small cell called the first polar body
- In meiosis II the secondary oocyte produces two cells by unequal cytokinesis
  - A large cell that rapidly divides into the ovum
  - a very small cell called a second polar body
- the polar bodies eventually degrade over time
- in the female only one mature gamete is produced by meiosis
- in humans, all oocytes are formed in the fetus
  - one oocyte completes meiosis I each month in an adult female
    - will only complete meiosis II when fertilized by a sperm
- **Diagram on Slide**

Chapter 3 - Lecture Notes

13-04-14 5:22 PM

Genetics

- Genetics is the study of the structure and function of genes
    - Early genetics focused on the passage of traits from parent to offspring
    - Advances in DNA technology and biochemistry allow the study of the structure and function of genes at the molecular level (molecular genetics)

Gregor Johann Mendel (1822-1884)

- In 1843 admitted to Augustinian Monastery in Brno
- Attended University of Vienna 1851-1854 majoring in physics and Botany
- In 1854 he began a series of breeding experiments with the garden pea, Pisum Sativum

Genotype and Phenotype

- Hereditary traits are characteristics that are passed down from one generation to the next
    - These traits are under the control of genes
- The genetic make-up of an organism is its genotype
- The phenotype is the observable characteristics of a set of characteristics of an organism produced by interaction between its genotype and phenotype
- **Diagram on Slide**

Mendel's Experiments

- From his results of crossbreeding pea plants with different characteristics. Mendel developed a simple theory to explain the transmission of heritable traits from one generation to the next
- He had no knowledge of meiosis, mitosis or that chromosomes even existed
- He published his conclusions in 1865, although their significance were not realized until the late 1800's and early 1900's

Why Mendel was Successful

- He kept extremely good records of all his observations
- He made simple interpretations of his results
- He made carefully controlled crosses between pea plants that had obvious differences
- He was creative
- He chose a good model organism

Pea Plants

- A good model organisms
    - Easy to grow

  - Bears flowers and fruit in the same year a seed is planted
  - Produces a large number of seeds
- Normally the plants reproduced by self-fertilization
- In order to cross the plants (cross-fertilization) Mendel had to transfer pollen mechanically from one flower to another flower

The Beginning of the Experiment

- Mendel obtained 34 strains of pea plants that differed in a number of traits
- He allowed each strain to self-fertilize for many generations
  - Ensure that traits he wanted to study were heritable
  - Produces true-breeding or pure-breeding strains
- From these plants Mendel decided to study 7 traits
  - Each trait had distinguishable alternate phenotypes

Mendel's 7 Traits

- Flower and see coat color (purple vs. white flowers and grey vs. white seed coats)(same gene controls flower and seed coat color)
- Seed color (yellow and green)
- Seed shape (smooth and wrinkled)
- Pod color (yellow and green)
- Pod shape (inflated and pinched)
- Stem height (tall and short)
- Flower position (axial and terminal)
- **Diagram on Slide**

Terminology

- P generation – parental generation
- F1 – or first filial generation, the progeny of P mating
- F2 – or second filial generation, the progeny of mating together of the F1 offspring
- F3, F4 – further interbreeding of the offspring from each generation

Monohybrid Cross

- Crosses btwn pea plants that differ in a single trait
- Mendel performed these crosses with his true-breeding plants
- **Diagram on Slide**

Reciprocal Cross

- Mendel also performed the reciprocal cross for each pair of traits he examined
    - Eg. Smooth (female) X Wrinkled (male) and Wrinkled (female) X Smooth (male)
    - In plants the female is given first in the crosses
- If the results of the reciprocal crosses are the same, then the trait does not depend on the sex of the organism

The 3:1 Ratio

- Mendel planted the seeds from his F1 plants and allowed them to self-fertilize to produce the F2 seeds
- Both smooth and wrinkled seeds appear in the F2 generation, even within the same pod
- Importantly, Mendel was meticulous in count and recorded of his results
- **Diagram on Slide**

Particulate Factors

- Mendel observed that the F1 generation was single phenotype, yet the F2 generation was a mixture of the two original parental types
- The problem was how the original parental type that disappeared in the F1, reappeared in the F2
    - This is where he was imaginative
    - He reasoned that a particulate factor could account for this pattern of heredity
        - We know that these factors today are genes

Alleles

- One of two or more alternative forms of a single gene that can exists at the same locus in the genome. All the alleles of a gene determine the hereditary trait (e.g. Seed shape), but each has a unique nucleotide sequence, which may result in different phenotypes (e.g. Smooth or wrinkled)

Dominant / Recessive

- The mating of two pea plants with different phenotypes (e.g. smooth x wrinkled seeds) brings together two alleles for the same locus in the plant
- In this case, even though both alleles were present, the plant displayed a single phenotype (e.g. smooth seeds)
    - The smooth allele masks the wrinkled allele
    - The smooth allele is dominant

    - The wrinkled allele is recessive

Conventions

- Capital letter represent the dominant allele
- Lowercase letter represents the recessive allele
- We can also denote the genotypes using this format as well
    - Parental smooth plants were SS (true breeding)
    - Parental wrinkled plants were ss (true breeding)
- The homozygous parental plants will produce only one type of gamete each, either S or s
    - When they fuse together they will produce a plant that is Ss… this plant has two different alleles for a particular trait and is heterozygous
- **Diagram on Next 4 Slides**

The Punnett Square

- In the F2 generation, three types of genotypes are produced
    - SS, Ss, and ss
    - They are present in a 1:2:1 ratio
    - The phenotype is 3:1
- The Punnett square is a representation of the gametes (on the top side) and the resulting zygotes (the center of the squares) that are formed
- **Diagram on Slide and Next 2 Slides**

Segregation

- Mendel's first lay is the principle of segregation
    - Two members of a gene pair (alleles) segregate (separate) from each other during the formation of gametes
    - Results in half of the gametes carrying one allele and the other half the other allele for each gene
- Segregation of the alleles occurs during meiosis when the pairs of homologous chromosomes separate in anaphase I
    - Segregation of alleles may occur in anaphase II If a crossover has occurred

Branch Diagrams

- **Diagram on Slide**

Determining the Genotypes of the F2

- Mendel wanted to know the genotypes of the smooth seed plants
    - They were either SS or Ss
    - He allowed these plants to self fertilize themselves and looked at the progeny
- **Diagram on Slide**

The Test Cross

- In the F2 of our pea example, the genotypes of the plants are SS, Ss, and ss
    - The SS and Ss have the smooth seeds
    - The ss plants have wrinkled seeds
- Of these plants, we know the genotype of the wrinkled seeds must be ss
- Yet we cannot tell if the smooth seed plants are SS or Ss
    - In order to determine their genotypes, a test cross is performed the plant is crossed to a plant that is homozygous recessive for the trait (e.g. ss)
- **Diagram on Slide**

What Causes Recessive Mutations

- Usually recessive mutations are loss-of-function mutations
    - Recessive since the wild type allele when present in a heterozygote is usually sufficient to produce enough functional product to produce the normal phenotype

Case Study: The Wrinkled Pea

- In SS peas, starch grains are large and simple
- In ss peas, the starch grains are small and deeply fissured
- SS seeds contain larger amounts of starch due to the activity of a starch-branching enzyme (SBEI) and contain less sucrose than ss seeds
    - More water is attracted to the ss seeds during development because of the sucrose so the seeds are larger (due to lack of SBEI activity)
    - When the ss seeds mature, they lose a larger proportion of their volume, leading to the wrinkled phenotype
- Seed shape gene encodes starch-branching enzyme (SBEI)
- SBEI is important in determining the starch content of embryos (without SBEI seeds accumulate excess sucrose in its place)
- Mendel's plants were found to have a 800bp insertion of the SBEI gene causing the s allele
    - Inserted DNA is a transposon

Dihybrid Crosses

- Mendel also analyzed a number of crosses in which two traits were involved
    - In each case, he observed the same results
- From these experiments, he proposed his second law, the principle of independent assortment
    - Factors for different traits assort independently of one another
    - Genes on different chromosomes behave independently in gamete production
- Using the Punnett square each parent produces only one type of gamete, either SY or sy
- All of the progeny are SsYy (smooth yellow seeded plants)
- **Diagram on Slide**

Dihybrid Cross Results SsYy X SsYy

- 9:3:3:1 phenotypic ratio is expected
    - 9 smooth, yellow (1 SSYY, 2 SsYY, 2SSYy, 4 SsYy)
    - 3 smooth, green (1 SSyy, 2Ssyy)
    - 3 wrinkled, yellow (1 ssYY, 2 ssYy)
    - 1 wrinkled, green (1 ssyy)
- This ratio is the expected ratio if the two genes are assorting independently
    - Mendle's actual results was 315 smooth yellow, 108 smooth green, 101 wrinkled yellow and 32 wrinkled green seeds… very close to predicted 9:3:3:1 ratio
- **Diagram on Slide (Branched Diagram)**

Testcross with SsYy

- SsYy plants will produce the following gametes
    - SY, Sy, sY and sy in a 1:1:1:1 ratio
- In the testcross SyYy x ssyy we would expect a SsYy, Ssyy, ssYy, and ssyy offspring all in a 1:1:1:1 ratio
    - this pattern is indictive of a testcross with an unknown parental strain that is a double herterozygote
- in the dihybrid cross, there are 9 different genotypic classes with only 4 phenotypic classes
    - To determine the genotype of the pants in the F2 testcrosses can be performed
- **Diagram on Next Slide**

Trihybrid Testcross

- SsYyCc X SsYyCc
- For 8 maternal and 8 paternal gametes, there are 64 combination of zygotes produced

  - 27 different genotypes
  - 8 different phenotypes
  - ratio in F2 is 27:9:9:9:3:3:3:1
- **Diagram on Slide and Next 2 Slides**

Statistical Analysis of Genetic Data

- Often, genetic crosses are quantitative
- To understand and interpret a set of data from crossing experiments it is important to understand the significance of any deviation of the observed results from the predicted results based on the hypothesis being tested
- Observed phenotypic values rarely match the expected ratios due to chance factors involved

Null Hypothesis

- In order to test the data, a hypothesis is developed called the null hypothesis
  - It states that there is no real difference between the observed and expected sets of data
  - If it is rejected, then a new hypothesis must be developed to explain the results
- How do we compare the two sets of data?
  - By the chi-square (x^2) test
    - A goodness – of – fit test

Chi-Square Analysis

- This test takes into account the observed deviation in each component of the expected ratio as well as the sample size
- The value of X^2 is then used to estimate how frequently the observed variation can be expected to occur strictly by chance
  - **Formula Description on Slide**
- **Example on Next Slides ** Be Able To DO!**

Interpretation of the P Value

- Our value of 0.3-0.5 tells us that 20-50% of the time, the variation we observed btwn our values is expected to be due to chance
  - So if we repeat the experiment, 20-50% of the trials would be expected to exhibit the same variation
- In this case, we would fail to reject the null hypothesis, in other words, the variation we observed in our values was likely due to chance

- If our P value was greater then the P value at 0.5, then we would reject the null hypothesis, suggesting that the observed variation was not due to chance alone
    - In other words, something else was going on
- **Tables on Next 2 slides**

<u>Mendelian Genetics in Humans</u>

- Geneticists discovered that Mendel's laws of inheritance of genes can be applied to all sexually reproducing eukaryotes, including humans
- W. Farabee, in 1905 discovered that brachydatyly which results in abnormally broad and short fingers is inherited
    - It is a simple autosomal dominant trait

<u>Pedigree Analysis</u>

- The study of inherited traits in humans is more difficult then those of plants
- In order to examine the inheritance of traits it is often necessary to study the family trees of individuals
    - This is called pedigree analysis
- **Diagram on Slide**

<u>Human Pedigrees</u>

- In many family trees of individuals will not know if they harbor any deleterious recessive mutations
- Often, interbreeding can lead to these recessive mutations becoming homozygous
    - Once identified it is often possible to work backwards and determine the genotypes of the parents, grandparents, and great grandparents
- **Diagram on Slide**
- Recessive traits usually skip generations
- Recessive autosomal traits appear equally in both sexes
- Can be autosomal or sex linked
    - Dominant / recessive
    - Gene linked
- **Diagram on Slide (Autosomal Recessive Trait)**
- Once lost (dominant trait) the next generations won't have it reappear, unless a new partner brings it in
- **Diagram on Slide (+ Notes Beside)**
- **Table on Next Slide**

Recessive Traits

- Albinism is a classic example of an autosomal recessive trait
- A typically loss of function mutation, these individuals do not produce melanin
    - Sensitive to sunlight
- Rare trait
    - 1 in 17 000 Caucasians
    - 1 in 28 000 African-Americans
    - 1 in 10 000 Irish
- **Diagram on Slide**

Some General Characteristics of Recessive Inheritance

- Most affected individuals have two normal parents
- Recessive traits often skip generations
- When both parents are affected, all of there offspring will be affected as well

Dominant Traits

- Dominant traits are often formed by gain-of-function mutations
    - Results in gene products with a new function
- Wooly hair is one example
    - Individuals hair is tightly kinked and brittle, breaking before it gets too long
    - Likely that person with traits is heterozygous… especially when it is a rare trait like wooly hair
    - In presence of functional genes, cause it to act differently phenotypically
    - Approximately 50% of children will have it
- Dominant traits have distinctive characteristics in pedigrees
    - Every affected person in the pedigree must have at least one affected parent
    - The trait usually does not skip a generation
    - On average, an affected heterozygous individual will transmit the mutant allele to half of his/her offspring

Probabilities

- Is a ratio of the number of times a particular event is expected to occur to the number of trials during which the event could have happened
- Product rule (AND)

    - The probability of two independent events occurring simultaneously is the product of each of their individual probabilities
- Sum rule (OR)
    - The probability of occurrence of any of several mutually exclusive events is the sum of the probabilities of the individual events
- Independent or dependent of one another
- **Examples on Slide**
- What are the expected ratios of the three possible genotypes?
- For smooth SS plants
    - S pollen must fuse with S egg
    - Frequency of S eggs is ½
    - Frequency of s pollen is ½
    - Thus, expected proportion of SS plants in the F2 is
        - ½ X ½ = ¼
        - expect proportion of wrinkled plants (ss) to be ¼ as well
- What would be the expected proportion of Ss plants?
    - Is it ½ X ½ ?
    - Almost there are two ways Ss plants can arise
        - From S (egg) X s (pollen)
        - From s (egg) X S (pollen)
    - So it is (1/2 X ½) + (1/2 X ½) = 1/12
    - Thus we predict that ¼ SS, ½ Ss, and ¼ ss which is the same result we get with the Punnett square (both 1:2:1)

Chapter 4 – Lecture Notes

13-04-14 5:22 PM

Some Terminology

- Wild type allele (+ or wt)
  - Allele that occurs most frequently in population
- Mutant allele (- or mt)
  - Loss of function mutations
    - Null mutation if complete loss of function
  - Gain of function mutations
    - Usually dominant
  - Neutral mutation
    - No effect
    - Recessive

Allele Symbols

- a+ / a+ (homozygous wild type, on different chromosomes/ homologs)
  - Or +/+
- a+ / a (heterozygous wild type)
  - Or +/-
- a/a (homozygous mutant)
  - Or -/-

Incomplete Dominance

- When one allele is not completely dominant to another then it is said to show incomplete or partial dominance
- The phenotype of the heterozygote is intermediate to those of the homozygotes
- Classic example
  - Flower color in snapdragons
    - Red X White
      - Progeny are all pink
      - 1:2:1 ratio (RR:Rw:ww)
- **Diagram on Slide**

Plumage Color in Chickens: Incomplete Dominance

- Crosses between true breeding black (C^BC^B) and white (C^WC^W) birds produces an F1 that is blueish grey (C^BC^W) called an Andulasian
- An Andulasian crossed to another one
  - 1 black: 2 Andulasian: 1 white fowl

  - ratio shows that you have incomplete dominance
- **Diagram on Slide**

Palomino Horses: Incomplete Dominance

- Do not breed true when bred together
  - Progency are ¼ cremellos (extremely light colored), ½ palominos and ¼ light chestnut (darker)
- C/C produces a horse that is light chestnul (or sorrel), C/C^cr palmino, and C^cr/C^cr the cremello
  - Other coat color genes are necessary but are at different loci
  - It is thought that there are up to 7 different loci that contribute to the coat color of horses
- In other words the palamino is a slight dilution of the sorrel, and the cremello is a further dilution to give an even paler color
- In order to get a palomino you would cross a chestnut and a cremello

Codominance

- The heterozygote displays both phenotypes of the two homozygotes
  - Incomplete dominance displayed intermediate phenotype between the 2 homozygotes
- The ABO blood group i^A and i^B alleles are codominant alleles
  - Individuals can be AB blood type
- Human M-N blood group
  - L^M/L^M and L^M/L^N and L^N/L^N
  - Antigens present on the surface of RBCs

Multiple Alleles

- So far we have only considered genes with two alleles
  - One mutant form and the other the wild type
- Some genes may have a wild type form and more than one mutant form
  - These genes are said to have multiple alleles
- Importantly, for all diploid individuals, it is only possible to have two of these alleles present
- **Diagram on Slide**

Drosophila Eye Color: Multiple Alleles

- All on the same gene, just pigment levels are different (except for the white)
- All variations of red pigments in eyes

- **Diagram on Slide**

The Nature of the ABO Blood Groups

- Red blood cells contain complex polysaccharides linked to the lipids of their membranes
    - These polysaccharides are on the outside of the RBCs
    - These are antigens, since in organisms where this molecule is foreign, antibodies will be produced against it
- People with type A have A antigens on their RBCs while type B individuals have type B antigens. Type AB individuals have both antigens present on their RBCs
- ABO locus encodes glycosultransferases, enzymes that add sugar groups to existing polysaccharides
- RBCs have a precursor glycolipid called the H antigen present
    - The A allele produces a glycosyltransferase enzyme that adds a alpha-N-acetylgalactosamine (sugar group) to the H antigen
        - Produces the A antigen
    - The B allele produces a different glycosyltransferase that adds galactose to the H antigen
        - Produce the B antigen
    - Type O individuals lack these glycosyltranferase enzymes, so the H antigen is unmodified
- **Diagram on Slide**

Antibodies to the ABO Antigens

- People with blood type O – have anti-A and anti-B antibodies
- People with blood type A – have anti-B antibodies
- People with blood type B – have anti-A antibodies
- People with blood type AB – have no antibodies to A or B
- What about the antigen?
    - It is shared amongst all these individuals so antibodies are not raised against this antigen
- **Diagram on Slide**

The H Antigen

- Is produced by the H allele (FUT1 gene) at a distinct locus from the ABO locus
- People who are homozygous h/h do not make this H antigen
    - Regardless of ABO locus, these individuals are similar to blood type O individuals

    - They lack A and B antigens
  - Produce anti-O antibodies (antibodies against the H antigen)
  - These people have the Bombay phenotype
  - Will attack all blood types
  - Can only get blood from another person that is Bombay
- **Diagram on Next Slide about Eye Color in Drosophila Melanogaster**

Lethal Alleles

- Genes products essential to an organisms survival need to be present
  - One copy may be sufficient, although in some cases it may lead to developmental abnormalities
    - E.g. stunted growth
- Two types of lethal alleles
  - Recessive lethals – essentially as described above
    - E.g. yellow coat color allele in mice
  - Dominant lethals – only need one copy for lethality
    - E.g. Huntington's disease
    - How? – would get one from mom and one from dad (heterozygous)
- **Diagram on Slide**

Huntington Disease

- Due to autosomal dominant allele H
- Results in gradual nervous and motor degeneration until they die
- Late onset disorder, typically beginning or manifesting itself by age 40
- All children have a 50% chance of inheriting the lethal allele

Epistasis

- Gene interaction where several genes may influence a particular trait
- Epistasis is where one gene or gene pair modifies or masks the expression of another gene
  - Three types
    - Recessive epistasis – E.g. Bombay phenotype over ABO locus
    - Dominant epistasis – E.g. fruit color in squash
    - Duplicate recessive – E.g. white flowered sweat peas
      - Also called complementary gene action

Coat Color in Mice

- Agouti X albino
    - F1 all agouti
    - F2 9/16 agouti, 3/16 black, 4/16 albino
- Color in mice is determined by C gene
    - C_ are black
    - cc are albino
- Agouti pattern (yellow banding) is controlled by the A gene
    - A_ are agouti
    - aa are nonagouti
- A_C_, aaC_ black and A_cc + aacc albino
- **Diagram on slide and next slide**

Dominant Epistasis

- A_B_ and A_bb individuals have the same phenotype
    - Observe 12:3:1 ratio rather then the 9:3:3:1 ratio
    - One gene when dominant is epispastatic to the other gene
- Summer squash example
    - Three common fruit colors; white, yellow, green
        - White X yellow gives white fruit
        - Yellow X green gives yellow fruit
            - Yellow is recessive to white, but dominant to green

Summer Squash: Dominant Epistasis

- Two gene pairs W and Y
- W_ have white fruit regardless of other locus
- ww plants will be yellow if Y_ or green if yy
- **Diagram on Next 2 Slides**

Dominant Epistasis: Greying in Horses

- Horses with GG or Gg genotype will progressively grey as they age
    - Skin and eye pigmentation remains that same, but the coat becomes whiter
- Gg horses remain the same color throughout their life
    - Well, they might get some greying due to old age

Epistasis and Duplicate Genes

- Consider the pathway below (**Diagram on Slide)** for flower color in peas
    - The C gene converts a white compound to a white intermediate and then a second gene, gene P converts this intermediate to a purple compound
- C_P_ plants are purple (9/16)
- ccP_ and C_pp plant are white (7/16)
- This is the result of complementary gene action
- So in our pea plans
    - Cross two white pea plants and the F1 is all purple. The F2 is 9 purple: 7 white????
        - Giving an F1 that was CcPp (purple)
        - F2 was 9 C_P-: 7 (ccP_ +C_pp +ccpp)
            - Cross 2 that don't show the final product
                - Cross them and get it
- **Diagram on Slide**

Epistasis and Duplicate Genes: Shepherd's Purse Plant

- Fruit shape in shepherd's purse plant can either be heart shaped or narrow
    - True breeding heart shaped fruit plant X narrow fruit plant
        - F1 plants produce heart shaped fruit
        - F2 plants show a ratio of 15 heart to 1 narrow
        - Modification of the 9:3:3:1 ratio
            - Heart shaped A_bb or aaB_ or A_B_
            - Narrow shaped aabb

Fruit Shape in Summer Squash

- Long fruit are always true breeding
- In some crosses between two different varieties of sphere shaped plants the F1 are disk shaped
    - F2 of disk shaped plants is
        - 9/16 disk shaped
        - 6/16 sphere shaped
        - 1/16 long shaped
- this is a modified mendelian ration, so two genes are likely involved
    - A_bb or aaB_ results in sphere shaped fruit (3+3)
    - A_B_ interact together to give disk shaped fruit (9)
    - aabb gives the long shaped fruit (1)

- **Diagram on Slide**

Coat Color in Labrador Retrievers

- One gene B- specifies black pigment, bb specifies brown pigment
- At another independent gene, E, allows the expression of the B gene, while ee does not allow expression of the B gene
    - B_E_ black lab
    - bbE_ chocolate
    - -/- ee yellow lab
        - if B_ee yellow lab with dark lips and nose
        - if bbee yellow lab with pale lips and nose
- E gene – transport protein for pigment from the cell into the hair follicles
- If recessive ee the pigment is still produced but it is not making it to the hair follicles

***Know Figure 4-7**

Complementation Analysis

- Case one – mutations are in separate genes
    - F1 has one normal copy of each gene present – complementation occurs
    - Flies are wild type and develop wings
    - Where are the mutant genes located?
    - Ho many genes are involved?
    - Must be homozygous
    - Turns heterozygous (compliment each other)
- Case Two – mutations are in different locations within the same gene
    - Gene 1 is mutant in all cases, while gene 2 is normal – no complementation occurs
    - Flies are mutant and do not develop wings
    - Both are in the same gene
    - Lead to homozygote for gene 1
    - If you cross it with itself you will get some with wings
- **Diagram on Slide**

Pleiotrophy

- A single gene with multiple phenotypic effects
- Marfan syndrome
    - Autosomal dominant

  - Gene encoding connective tissue protein fibrillin which is found in many tissues
    - Lens of eye, lining of vessels and bones
    - Risk of aortic aneurism, eye defects and lengthened bones in limbs

Porphyria Varigata

- Autosomal dominant disorder
- Defect in break down of porphorin component in hemoglobin
  - Accumulation leads to red colored urine, abdominal pain, muscular weakness, insomnia, headaches, delirium etc.

X Linkage

- When genes are located on the X chromosome
- In humans, females have two X chromosomes while males only have one X and one Y
  - Females can be homozygous or heterozygous for alleles on the X chromosome
  - Males only have one X, so they are hemizygous for X-linked traits
- **Diagrams on Next 4 Slides**

Sex

- Expression of particular genes may be influenced by the sex of an individual
  - In the case of sex-linked genes the differences in the phenotypes between the two sexes are related to the differences in their complement of the sex chromosomes
- Some genes that are on the autosomes affect a particular character in one sex and not the other
  - These are called sex-limited traits

Some Sex Limited Traits

- Milk production in dairy cattle
- Horns in certain species of sheep
  - Males with genes for horns have horns, while females do not
- Distribution of facial hair in humans
- Cock feathering in fowl
- Not on X chromosome

Sex Influenced Genes

- Often controlled by autosomal genes, such traits appear in both sexes, but either the frequency is different between the two sexes, or the relationship between the genotype and phenotype is different
- Pattern baldness is a classic example
    - Sex-influenced autosomal gene
    - Dominant in males, but recessive in females
        - BB individuals are bald
        - Bb males are bald, females are nonbald
        - bb individuals are not bald
- Other sex influenced traits
    - Cleft lip/palate (2:1 male to female ratio)
    - Clubfoot (2:1)
    - Gout (8:1)
    - Rheumatoid arthritis (1:3)
    - Osteoporosis (1:3)
    - Systemic lupus erythematosus (1:9)

Environmental Influences: Penetrance

- Not all individuals of a certain genotype will show the expected phenotype
    - Frequency with which a dominant or homozygous recessive gene manifests itself in individuals of a population is called penetrance
    - Depends on both the genetic background of the individual and the environment
    - If all individuals carrying a dominant allele show the mutant phenotype then the gene shows complete penetrance

Incomplete Penetrance

- Braxhydactyly an autosomal dominant trait causing shortened and malformed fingers
    - Shows 50-80% penetrance
        - Only 50-80% of people with the dominant mutant allele will have the phenotype
- **Diagram on Slide**

Expressivity

- Degree to which a penetrant gene is phenotypically expressed in an individual
- Depends on both genetic background and environment

  - E.g. Osteogenesis imperfect
    - Causes blueness of sclera (whites of eyes), very fragile bones and deafness
    - Autosomal dominant disorder with 100% penetrance
    - Individuals show variable expressivity
      - Some individuals have one or more of the traits
      - Fragility of bones may be more severe in some individuals
- **Diagram on Slide**

Incomplete Penetrance and Variable Expressivity

- Some genes exhibit both incomplete penetrance and variable expressivity
  - E.g. Neurofibromatosis
    - Autosomal dominant
    - Shows 50-80% penetrance
    - Develop tumor like growths over the entire body
    - In milder forms causes café-au-lait spots on the skin
    - Severe manifestations can lead to curvature of the spine, tumors of the eyes and brain, high blood pressure and more

Temperature

- Biochemical reactions carried out by enzymes
  - Normally these enzymes operate over a wide range of temperatures
  - Some may be temperature sensitive
    - May work over a certain temperature range, but not another
- Classic example is the Himalayan rabbit and Siamese cat

Himalayan Rabitts

- Dark fur develops in extremities
  - Ears, paws, tail and nose
- Hypothesized that differences due to lower temperature in these regions
  - High body core temperature inhibits the action of a color pigment gene to create white body color
- Experiments
  - Rabbits in 30 degree incubators are white
  - Rabbits in 25 degree incubators are Himalayan
  - Rabbits in 25 degree incubators with cold patches developed black spots where cold patch was applied

The Effects of the Environment: Age of Onset

- Age of organism creates internal environmental changes that can affect gene function
    - E.g. pattern baldness appears between 20-30 years of age
    - Duchenne muscular dystrophy appears in children between 3-5 years old
    - Huntington disease appears at age 40
- Sometimes the accumulation of toxic products of metabolism give a delayed effect
    - E.g. Tay-Sachs disease, an autosomal recessive
        - Lipid metabolism disorder, hexaminidase A enzyme deficiency
        - Build up of toxic lipids over time leads to rapid developmental retardation, paralysis and death by age 3
        - Prevalent in people of European Jewish decent
- Lesch-Nyan syndrome
    - X-linked recessive
    - Abnormal nucleic acid metabolism
        - Accumulation of nitrogenous purine bases
        - Mutation in gene encoding hypoxanthine-guanine phosphoribosyl transferase (HGPRT) which is toxic and not incorporated into DNA
    - Uric acid in blood, mental retardation, palsy and self-mutilation
    - More common in males
    - Indications that genes can play a role in behavior

Genetic Anticipation

- Some disorders show a progressively earlier age of onset and increased sensitivity with each successive generation
- Myotonic Dystrophy (DM)
    - Adult dystrophy caused by autosomal dominant mutation
    - Mild symptoms include, cataracts as adults with little muscular weakness
    - Severely affected have extensive myopathy and mental retardation
    - Most extreme form may result in death soon after birth
    - Wide range of symptoms
    - Cause of genetic anticipation due to a trinucleotide (CGG) repeat
- The cause
    - A particular region of DM gene had the presence of a short trinucleotide DNA repeat
        - Normal have 5-10 copies
        - Mildly affected has 50 copies

        - Severely affected have over 1000 copies
    - Similar trinucleotide repeats play a role in fragile-X syndrome and Huntington disease
        - Repeats increase with each generation

<u>Genomic Imprinting</u>

- During early development some genes may be silenced
    - The gene is essentially turned "off"
    - Not permanent, as these genes may be expressed again in subsequent generations or later in development
- Usually one or the other homologs in "imprinted", depending on their parental origin
    - Thus only the allele not silenced will be expressed
- Epigenetics – not altering code just if it is on/off
    - Due to methylation – causes no transcription or translation
- **Example of Imprinting of Igf2 Gene**

<u>Imprinting: PSW and AS</u>

- A deletion on chromosome 15 appears to be responsible for two syndromes
    - Prader-Willi syndrome (PWS)
        - Only when deletion is inherited from mother
            - Good copy from father shut down by imprinting
        - Small hands and feet, short stature, mental retardation and often obese and compulsive eaters
        - Are poor eaters until about age 3
    - Angelman Syndrome (AS)
        - Only when deletion is inherited from father
            - Good copy from mother shut down by imprinting
        - Red cheeks, large jaw and mouth, happy disposition, severe mental and motor retardation

<u>Extranuclear Inheritance</u>

- Inheritance of organelles that contain DNA
    - E.g. mitochondria and chloroplasts
        - Organelle inheritance (maternal inheritance)
- Organisms phenotype determined by genetic information (usually mRNA or protein) in the gamete of one parent

  - Usually from the mothers egg cell
  - Zygotes early development determined solely by mothers genotype – maternal effect

<u>Examples of Organelle Heredity</u>

- Shoot variegation in the 4 o'clock
  - Variegated shoot phenotype in 4 o'clock involves non-mendelian inheritance of chloroplasts in the shoots (stems, leaves, and flowers)
    - Green shoots have normal chloroplasts
    - White shoots have only leucoplast, which lack chlorophyll and are incapable of photosynthesis
    - Variegated shoots received both chloroplasts and leucoplasts, which segregated during cell division. Progeny cells are therefore green or white, in a mixed pattern
- **Diagrams on Next 2 Slides**

<u>Mitochondrial Mutations: The poky Mutant of Neurospora</u>

- Is an aerobe, so it requires mitochondrial functions to grow, the poky mutation in mtDNA has an altered cytochrome complement, leading to slow growth of the fungus
- Experimental crosses in Neurospora involve fusion of nuclei from mating type A and a parents. Crosses can occur in 2 ways
  - Place both parents on the medium at the same time
  - Inoculate one parent onto the medium, and add the second parent several days later. The first parent produces all the protoperithecia (fruiting bodies containing ascospores)

<u>Poky</u>

- Protoperithecia have much more cytoplasm then conidia (asexual spores)
  - The strain producing this is similar to the female parent
  - The second strain which contribute to conidia is analogous to the male parents
- Assigning these roles allows reciprocal crosses to be made
  - Protoperithecia from poky parent and wt conidia results in all poky progeny
  - Protoperithecia from wt parents and poly conidia results in all wt progeny
  - Results indicate maternal inheritance
- **Diagram on Slide**

Human Genetic Diseases and Mitochondrial DNA Defects

- Leber's hereditary optic neuropathy (LHON). Optic nerve degeneration results in complete or partial blindness in midlife adults (average age 27)
    - LHON is caused by mutations in mtDNA genes for electron transport chain proteins
        - 50% of cases affect NADH dehydrogenase
    - LHON results from defects in the enzymes of oxidative phosphorylation. Without ATP production, the optic nerve dies
- Kearns-Sayre Syndrome produces 3 types of neuromuscular defects
    - Progressive paralysis of certain eye muscles and hearing loss
    - Abnormal pigment accumulation on the retina, causing chromic inflammation and degeneration of the retina
    - Heart disease
    - Results from deletions in mtDNA. A model for the disorder is that tRNA genes are removed, disrupting mitochondrial translation
    - Many KSS patients are symptom free as children, but display progressive symptoms as adults
        - The proportion of mtDNAs with deletion mutations increases as the severity of symptoms increase
- Myoclonic epilepsy and ragged-red fibre disease
    - Symptoms include
    - Microscopic tissue abnormality, raged-red fibers
    - Myoclonic seizures (jerking spasms)
    - Ataxia (uncoordinated movement)
    - Accumulation of lactic acid in blood
    - Many additional symptoms
    - Mitochondrial have abnormal appearance
    - Caused by a single nucleotide substitution in the lysine tRNA gene. Mitochondrial protein synthesis is affected, and in some way this phenotype is produced
        - Most patients have a mixture of normal and abnormal mitochondria – heteroplasmy

Maternal Affect

- Some maternally derived phenotypes are produced by the maternal nuclear genome (maternal effect), rather then inherited as extranuclear genes (organelle inheritance)

    - Proteins and or mRNA deposited in the oocyte before fertilization direct early development in embryo
    - The genes encoding these products are on nuclear chromosomes. NO mtDNA IS INVOLVED

__Ephestia Pigmentation__

- Maternal effect seen in Mediterranean meal moth
- Pigment derived from precursor molecule kynurenine (a derivative of tryptophan)
    - Mutation "a" unabe to synthesize kynurenine
    - Homozygous larva have little pigmentation and red eyes
- **Diagram on Slide**

__Shell Coiling in Limnaea Peregra__

- Shell coiling is determined by a pair of nuclear alleles, with the dominant D allele producing a dextral (right) coil, and the recessive d allele producing the sinistral (left) coil
- The shell phenotype is always determined by the mother's GENOTYPE
- In all crosses of true breeding dextral and sinistral snails, the F1's have the same genotype (D/d) but the reciprocal crosses produce different phenotypes
    - A dextral female (D/D) crossed with a sinistal male (d/d) produces a dextral F1 (D/d)
        - The F2 genotype have a 1:2:1 ration (DD:Dd:dd). All F2 snails, including those with genotype dd have dextral shells
        - Selfing the F2produces an F3 that is ¾ dextral and ¼ sinistral. The sinistral snails are the progeny of the F2 dd mothers (who had dextral shells)
- **Diagram on Slide**
- A sinistral female (d/d) crossed with a dextral male (D/D) will produce a sinistral F1 (D/d)
    - The F2 genotypes also have a 1:2:1 ration (DD:Dd:dd). All F2 snails have dextral shells
    - Selfing the F2 produces an F3 that is ¾ dextral and ¼ sinistral
- **Diagram on Slide**
-

Chapter 5 – Lecture Notes

Some Terminology

- Heteromorphic chromosomes
    - E.g. X and Y
- Unisexual, dioecious, and gonochoric
    - Contains only male OR female reproductive organs
- Bisexual, monoecious and hermaphroditic
    - Contain both male AND female reproductive organs
- Intersex
    - Individuals of intermediate sexual differentiation
    - Usually sterile
- **Diagrams on Next 3 Slides**

Modes of Sex Determination

- Protenor mode (XX and XO)
- Lygaeus mode (XX and XY) or (ZZ – male and ZW – female)
- In both cases one sex produces two types of gametes (heterogametic sex) while the other produces only one type of gamete (homogametic sex)
- For XY males determine the sex of the offspring while ZW females determine the sex of the offspring
- **Diagram on Next 2 Slides**

Y Chromosome in Humans

- Correlation of the presence of Y chromosomes in males and not females
    - Males 2n = 46 (XY)
    - Females 2n = 46 (XX)
- **Diagram of Female and Male Karyotypes**
- Chromosomes 1-22 – autosomes
- Chromosome 23 – sex chromosomes

Klinefelter Syndrome 47, XXY – 2n = 47

- Genitalia and internal ducts are generally male
- 1 in 1000 male births
- Testes underdeveloped and fail to produce sperm
- Tall, long arms and legs, larger hands and feet
- Masculine development occurs, but some feminine sexual development may occur
    - Slight breast development and rounding of hips

- Slight decrease in intelligence (below normal range)
- Result of nondisjunction event in parent
    - Can occur in either parent
    - For males it occurs in meiosis I
    - For females it occurs in meiosis I and II
- **Diagram on Next 2 Slides**

<u>Turner Syndrome 45, XO – 2n = 45</u>

- Female external genitalia and internal ducts
- 1 in 10 000 female births
- Ovaries are no fully developed – semi sterile
- Short stature, skin flaps on back of neck and underdeveloped breasts. Shield-like chest
- Intelligence is normal
- Result of nondisjunction event in parent
- **Diagram on Slide**

<u>47, XXY Syndrome – 2n = 47</u>

- Occurs in about 1 of every 1200 female births
- Observe variable expression
    - Some XXX individuals are phenotypically normal, while others show underdeveloped secondary sex characteristics and mental retardation can occur
- Also see 48, XXXX and 49 XXXXX individuals, more pronounced symptoms

<u>47, XYY Condition – Super Males</u>

- Phenotypically male, usually over 6 feet tall, sometimes associated with lower intelligence and behavioral problems
- Considerable attention was given to these individuals in the 1960's and 70's
    - Patricia Jacobs in 1965 found an unusually high proportion of XYY individuals in a Scottish maximum security prion (9 of 315 male inmates)
    - Suggested that they were more likely to be criminally minded
    - Research in this area was abandoned in 1970's to avoid harm on the subjects in this study

<u>The Y Chromosome</u>

- Originally thought not the have genes or homology to other chromosomes
    - Pseudoautosomal regions (PARs) shave homology with X chromosome

        - Allow pairing during meiosis
    - The rest of the chromosome is called the male specific region of the Y (MSY)
        - Some genes in common with X, and some novel genes
- There are about 164 genes on Y, out of them 60 are functional
    - 104 are pseudo genes – they were functional but now they are not
- why is PAR at the telomeres?
    - Because that it where the homologs search is
- **Diagram on Slide**
- Presence of Y chromosome determines sex in humans
    - When present the organism is a male
    - When not present the organism is a female
- There is at least one very important gene that determines maleness on the Y chromosome
    - The sex determining region Y (or SRY) gene
    - SRY product is called the testis determining factor (TDF) and triggers the undifferential gonadal tissue of the embryo to form testes (start to produce testosterone)

Evidence

- There are females that are XY
    - Upon closer examination they found that the Y chromosome lacked the SRY gene
- There are males that are XY but they develop as females
    - Called testicular feminization syndrome with a frequency of 1 in 65 000 male births
    - Have external female genitalia, a blind vagina, and no uterus. Testes present in labia or abdomen
    - Lack the androgen receptor so the androgens that determine maleness have no effect
    - Essentially like Turners (XO) default is to form a female
- **Diagram on Slide**

Ration of Males to Females

- Expect a ratio of 1
- In the US population the ratio is 1.06 males to females born (secondary ratio)
    - Primary ratio (conception ratio) is more difficult to determine
- May be the result of Y bearing sperm being able to swim faster then X-bearing sperm
    - This is because they weigh different

Dosage Compensation

- Organisms with sex chromosomes have an inequality in gene dosage
  - The number of gene copies between the sexes varies
  - 2 copies in human females, while only one copy in males from the X chromosome
  - To equalize this expression, systems for dosage compensation have evolved
  - Males can survive due to only having one X chromosome

Barr Bodies

- In mammals, somatic cell nuclei of XX females contain a highly condensed mass of chromatin
  - Called a Barr body
  - Not found in XY males
- In 1961, Mary Lyon and Lillian Russell expanded upon this observation to propose a concept known as the Lyon hypothesis
- **Diagram on Slide**

The Lyon Hypothesis

- In XX females, the Barr body is a highly condensed and genetically inactive X chromosome. It has become lionized in a process called lionization
  - This leaves just one genetically active X chromosome, just like in males
- The X chromosome inactivated is randomly chosen from the maternally and paternally derived X chromosomes
  - Once it has chosen, all of the descendants of that cell will have the same X chromosome inactivated

X – Inactivation

- Is a heritable change in gene expression which occurs without a change in DNA sequence
  - An example of an epigenetic phenomenon
    - Heritable modification in gene function not due to changes in the DNA sequence
  - X inactivation is an epigenetic silencing of one X chromosome
  - Occurs at day 16 after fertilization in humans (at 500 to 1000 cell stage)
  - If heterozygous for traits on the X chromosome, then females will be genetic mosaics
    - Some cells chow the phenotype of one X chromosome, while other show the phenotype of the other X chromosome
- **Diagram on Next Slide**

Calico and Siamese Cats

- A female cat
    - heterozygous for the X linked black and orange alleles
        - same mechanism as tortoiseshell cat
    - white patches are due to the autosomal spotting color gene S
        - S determines the timing of melanocyte migration
            - Early migration, then black or orange pigmentation
            - Late migration, then white or less intense black or orange pigmentation
        - Interestingly, tortoiseshell cats lack this gene

More Examples of Mosaics

- Human females heterozygous for X-linked mutation that causes the absence of sweat glands (anhidroitic ectodermal dysplasia)
    - Mosaic of skin patches lacking sweat glands
- Red-green color blindness
    - Females may have retinas with patches of defective color perception

X Chromosome Tolerance

- X chromosome inactivation explains how mammals can tolerate changes in the numbers of X chromosomes
    - E.g. XXX females and XXY males
- Lionization in cells with extra X chromosomes operates by inactivating all but one of the X chromosomes
    - Thus an XXX female will have 2 Barr bodies
- **Diagram on Slide**

Sex Determination in Drosophila

- Determined by the ratio of the number of X chromosomes to the number of sets of autosomes
    - This in an X chromosome-autosome balance system of sex determination
    - The Y chromosome has no effect on sex determination
        - It is only required for male fertility
- Females are homogametic sex (XX) and males are heterogametic sex (XY)
    - However XXY flies are female and XO flies are male (sterile due to the lack of the Y chromosome)

- A normal female has 2 X's and 2 sets of autosomes for a ratio of X:A or 1
- A normal male has one X and 2 sets of autosomes for a ratio of X:A of 0.5
    - If the X:A ratio is greater then 1 its female
    - If the X:A ratio is less then or equal to 0.5 it is male
    - If the ratio is between 0.5 and 1, then it is intersex
        - Variable in appearance with characteristics of both sexes
        - Sterile
- **Diagram on Slide**

Chapter 6 – Lecture Notes

Variation in Chromosome Number

- An organism or cell is euploid when it has one complete set of chromosomes, or exact multiples of each set
    - Eukaryotes that are haploid or diploid are euploid, as are other organism with variable number of chromosome sets
- Aneuploidy results from variations in the number of individual chromosomes (not sets), so that the chromosome number is not an exact multiple of the haploid set of chromosomes

Generation of Aneuploidy

- Can occur due to nondisjunction during meiosis
    - In meiosis I will produce 4 gametes, two with a chromosome duplicate and two that are missing a chromosome
        - Fusion of a normal gamete with one containing a chromosomal duplication will produce a zygote with 3 copies of that chromosome, and two of all the others
        - Fusion of a normal gamete with one missing a chromosome will result in a zygote with only one copy of that chromosome and 2 of all the others
- **Diagram on Slide**
    - In meiosis II produces two normal gametes and two that are abnormal (one with two sibling chromosomes and one with that chromosome missing)
        - Fusion of abnormal gametes with normal ones will produce the genotypes discussed previously
        - Normal gametes are also produced when fertilized will produce normal zygotes
- **Diagram on Slide**

Changes in One or a Few Chromosomes

- More complex gametic chromosome composition can result when
    - More than one chromosome is involved
    - Nondisjunction occurs in both meiotic division
    - Nondisjunction occurs in mitosis (results in somatic cells with unusual chromosome complements)

Aneuploidy

- Autosomal aneuploidy is not well tolerated in animals, and in mammals is detected mainly after spontaneous abortion. Aneuploidy is much better tolerated in plants

- There are 4 main types of aneuploidy
    - Monosomy – loss of a single chromosome (2N-1)
    - Nullisomy – involves loss of one homologous chromosome pair (2N-2)
    - Trisomy – involves one extra chromosome, so the cell has three copies of one and two of all the others (2N+1)
    - Tetrasomy – involves an extra chromosome pair so that a cell has 4 copies of one and 2 of the others
- More then one chromosome or chromosome pair may be lost or added
    - A double monosomic aneuploidy has 2 separate chromosomes present in only one copy each (2N – 1 – 1)
    - A double tetrasomic aneuploidy has two chromosomes present in 4 copes each (2N + 2 +2)
- **Diagram on Slide**

<u>Meiotic Consequences</u>

- Some types of aneuploidy have serious mitotic consequences
    - A monosomic cell (2N-1)
        - May produce gametes that are N (normal) and N-1 (monosomic)
        - Or the unpaired chromosome may be lost completely, producing gametes that are all N-1
- A trisomic cell (2N+1) with the genotype +/+/a is an example (assuming organism can tolerate trisomy and no crossing over occurs)
- Gametes produced belong to four genotypic classes in these proportions
    - 2 gametes +/a
    - 2 gametes +
    - 1 gamete +/+
    - 1 gamete a
- the cross of a +/+/a trisomic to an a/a individual will produce a phenotypic ratio of 5 wild type: 1 mutant (a)
- **Diagram written below and Next 2 Slides**

<u>Human Examples of Aneuploidy</u>

- Sex chromosome aneuploidy is found more often then autosome aneuploidy because lionization compensates from chromosome dosage
- Autosomal monosomies are rarely found in humans, presumably because they are lost early in pregnancy (Turner's is an example)

- Autosomal trisomies account for about half of fetal deaths, and only a few are seen in live births. Most (trisomy 8,13,and 18) result in early death with only trisomy-21 (down syndrome) surviving to adulthood

<u>Trisomy-21 (Down Syndrome)</u>

- Occurs in an estimated 3510 per 1 million conceptions and 1430 per 1 million live births
- Down syndrome individuals are characterized by
    - Low IQ
    - Epicanthal folds over eyes
    - Short and broad hands
    - Below average height
    - Simian crease
- **Diagram on Slide**

<u>Correlation of Maternal Age and Probability of Trisomy-21</u>

- a female fetus before birth produces primary oocytes in her ovaries that stop their development at prophase I of meiosis
- after puberty secondary oocytes develop from the primary ones entering the second meiotic division but again arresting this time at metaphase II
- if fertilized, the second meiotic division is completed
- the probability of nondisjunction increases with the length of time the primary oocyte is in the ovary
- **Diagrams on Next 2 Slides**

<u>Additional Risks for Down Syndrome</u>

- Increased paternal age
- Smoking in mothers who have an error in meiosis II, especially if they use oral contraceptives (oral contraceptives alone do not increase the risk)
- Amniocentesis or chorionic villus sampling can determine whether the fetus has a normal complement of chromosomes

<u>Trisomy-13 (Patau Syndrome)</u>

- Occurs in 2/1000 live births, most die within the first 3 months
    - Cleft palate and lip
    - Small eyes
    - Polydactyly (extra fingers and toes)

  - Mental and developmental retardation
  - Cardiac and other abnormalities
- **Diagram on Slide**

Trisomy-18 (Edwards Syndrome)

- Occurs in 2.5/10 000 live births
- 90% will die within 6 months
- 80% are female
- Characteristics
  - Small size with multiple congenital malformations throughout the body
  - Clenched fists
  - Elongated skull
  - Low set ears
  - Mental and developmental retardation
- **Diagram on Slide**

Changes in Complete Sets of Chromosomes

- Monoploidy and polyploidy involve complete sets of chromosomes, so both cases are euploidy
  - Euploidy is lethal in most animal species but often tolerated in plants where it has played a role in speciation and diversification

Monoploidy and Polyploidy

- Can result when either round of meiotic division lacks cytokinesis, or when meiotic nondisjunction occurs for all chromosomes
  - Complete nondisjunction in meiosis I will produce ½ gametes with normal chromosomes, ¼ with two sets and ¼ with no sets of chromosomes
  - Gametes with 2 sets of chromosomes fused with a normal gamete produces a triploid (3N) zygote
  - Fusion of 2 gametes that each have two sets of chromosomes produces a tetraploid (4N) zygote
- Polyploidy of somatic cells can result from meiotic nondisjunction of complete chromosome sets
- **Diagram on Next 2 Slides**

Monoploidy

- Rare in adults of diploid species due to recessive lethal mutations
    - Males of some species (e.g. wasps, ants, and bees) develop from unfertilized eggs and are monoploid
    - Mutants are easily produced from monoploid individuals since mutants can be isolated directly

Polyploidy

- Involves 3 or more sets of chromosomes and may occur naturally (e.g. by breakdown of mitotic spindles) or induction (e.g. with chemical such as colochicine)
- Nearly all plants and animals have some polyploidy tissues
    - Plant endosperm is triploid
    - Liver of mammals (and perhaps other vertebrates) is polyploidy
    - Giant abdominal neurons of Aplysia has about 75 000 copies of the genome
    - Wheat is hexaploid (5N) and the strawberry is octaploid (8N)
    - North American sucker fish, salmon, and some salamanders are polyploidy
- Two classes based on number of chromosomes
    - Even number polyploids are more likely to be at least partially fertile, because the potential exists for equal segregation of homologs during meiosis
    - Odd number polyploids will always have unpaired chromosomes, Balanced gametes are rare and these organisms are usually sterile or have increased zygote death

Triploids

- Unstable in meiosis because random segregation means that balanced gametes (either N or 2N) are rare
    - The probability of a triploid organism producing a haploid gamete is (1/2)^n where n is the number of chromosomes
    - Triploidy is always lethal in humans, accounting for 15-20% of spontaneous abortions and 1/10 000 live births, most dying withing the first month
    - Tetraploidy in humans is also lethal, usually before birth, accounting for 5% of spontaneous abortions
- **Diagram on Slide**

Polyploidy in Plants

- Is more common in plants, probably due to self fertilizaiotn
    - Allows even number polyploids to produce fertile gametes and reproduce

- Two types
    - Autoploidy (self) and alloploidy (from outside)
- **Diagram on Slide**

Autopolyploidy

- Results when all sets of the chromosomes from the same species
    - Usually due to meiotic error
    - Fusion of a diploid gamete with a haploid one produces a triploid organism
        - E.g. grasses, garden flowers, crop plants and forest trees
        - E.g. seedless fruits such as bananas, grapes and watermelons
            - Odd number of set of chromosomes
                - No seeds because it cannot complete meiosis

Allopolyploidy

- Results when the chromosomes are from 2 different organism
    - Fusion of haploid gametes followed by chromosomes doubling
        - E.g. fusion of a haploid gamete from plant 1 and plant 2 produces N1 + N2 hybrid plants. No chromosomal pairing occurs in meiosis, viable gametes are not produced and the plants are sterile
        - Rarely, division error doubles the chromosome sets (2N1 + 2N2). The diploid sets function normally in meiosis and fertile allotetraploid plant results
- **Diagram on Slide**
- **Example!!!**
    - Crosses of cabbages and radishes, which both have a chromosome number of 2n=18
        - The F1 hybrids have 9 chromosomes from each parents, and have a morphology intermediate between cabbages and radishes. They are mostly sterile
        - A few seeds, some sterile, can be produced by the F1 thought meiotic errors
            - Somatic cells in the resulting plants have 36 chromosomes, a full diploid set from both cabbages and radishes
            - These fully fertile plants look much like the F1 hybrids and are named Raphanobrassica
                - Root like cabbage and shoot like radish… they were hoping for the reverse

Types of Chromosomal Mutations

- Variations in chromosomes structure or number can arise spontaneously or be induced by chemicals or radiation. Chromosomal mutations can be detected by
    - Genetic analysis (observing changes in linkage)
    - Microscopic examination of eukaryotic chromosomes in mitosis and meiosis (karyotype analysis)

Chromosomal Mutations

- Contribute significantly to human miscarriages, stillbirths and genetic disorders
    - About ½ of spontaneous abortions result from major chromosomal mutations
    - Visible chromosomal mutations occur in about 6/1000 live births
    - About 11% of men with fertility problems and about 6% of people institutionalized with mental deficiencies have chromosomal mutations

Variation in Chromosomal Structure

- Mutations involving changes in chromosomes structure occur in four common types
    - Deletions
    - Duplications
    - Inversions (changing orientation of a DNA segment)
    - Translocations (moving a DNA segment)
- All chromosomes structural mutations begin with a break in the DNA, leaving ends that are not protected by telomeres, but are 'sticky' and may adhere to other broken ends

Deletions

- In a deletion, part of a chromosome is mission
- Start with a chromosomal break induced by
    - Heat or radiation (ionizing)
    - Viruses
    - Chemicals
    - Transposable elements
    - Errors in recombination
- Importantly, deletions do not revert, because the DNA is missing
    - Don't go back to being the wild type
- **Diagram on Slide**

Effects of a Deletion

- A deletion in one allele of a homozygous wild type organism may give a normal phenotype, while the same deletion in the wild type allele of a heterozygote would produce a mutant phenotype
- Deletion of the centromere results in an acentric chromosome that is lost (centromere is gone), usually with serious or lethal consequences.
    - No known living human has an entire autosome deleted from the genome
- Large deletions can be detected by unpaired loops seen in karyotype analysis

Deletions and Human Disorders

- Human disorders caused by a large chromosomal deletions are generally seen in heterozygotes since homozygotes die
    - The number of gene copies is important
    - Syndromes result from the loss of several to many genes

Examples of Human Disorders

- Cri-du-chat (cry of the cat) syndrome resulting from the deletion of part of the short arm of chromosome 5 (46,5p-)
    - Results in severe mental retardation and physical abnormalities

Duplications

- Result from the doubling of chromosomal segments and occur in a range of sizes and locations
    - Tandem duplications are adjacent to one another
    - Reverse tandem duplications result in genes arranged in the opposite order of the original
    - Tandem duplications at the end of a chromosome is a terminal tandem duplication
    - Heterozygous duplications result in unpaired loops and may be detected cytologically
- **Diagram on Next 2 Slides**

Duplication Example: Bar Eye

- Drosophila eye shape allele, Bar, reduces the number of eye facets, giving the eye a slit like rather than an oval appearance
    - The bar allele resembles an incompletely dominant mutations
        - females heterozygous for Bar have a kidney shaped eye that is larger and more faceted than a female homozygous for Bar

        - males hemizygous for Bar have slit like eyes like those of a Bar/Bar female
    - cytological examination of polytene chromosomes showed that the Bar allele results `from duplication of a small segment (16A) of the X chromosome
- isolated from salivary glands – go through repeated S phase without division
- **Diagram on Slide**

Multigene Families

- Result from duplication. Hemoglobin (Hb) is an example
    - Each Hb contains two copies of two subunits (e.g. 2 alpha globin's and 2 beta globin's) and the identity of the subunits changes with the organisms developmental stages
    - Genes for the alpha type polypeptides are clustered together on one chromosome, and those for the beta type are clustered on another
    - Alpha type genes have similar sequences as do beta type. They probably arose from duplication and subsequent sequence divergence

Inversions

- Results when a chromosome segment excises and reintegrates oriented at 180 degree from the original orientation. There are two types
    - Paracentric inversions do not include the centromere
    - Pericentric inversions include the centromere
- Generally do not result in lost DNA, but phenotypes can arise if the breakpoints are in genes or regulatory regions
- **Diagram on Slide**
- Linked genes are often inverted together. The meiotic consequence depends on whether the inversion occurs in a homozygote or a heterozygote
    - A homozygote will have normal meiosis
    - For a heterozygote it depends on crossing over
        - If there is no crossing over, no meiotic problems arise
        - If crossing over occurs within the inversion, unequal crossover may produce serious genetic consequences

Inversions and Crossing Over

- Paracentric inversions
    - Result in visible inversion loops between homologous chromosomes

    - Crossover in the inversion region results in unbalanced set of genes, and gametes or zygotes derived from recombined chromatids may not be viable due to abnormal gene dosage
    - Without crossover in the looped region, gametes receive complete sets of genes (2 gametes normal and 2 gametes inverted) and are viable
    - Effects of a single cross over within an inverted segment in a heterozygote include
        - Joining of homologous regions of two chromatids to produce a dicentric bridge and corresponding loss of an acentric fragment
            - During anaphase the 2 centromeres of the dicentric chromosome migrate towards opposite poles causing the bridge to break and producing two chromatids with deletions
            - In meiosis II one chromati is distributed to each gamete
                - Two gametes carry normal sets of genes (one is normal order and the other is the inverted order)
                - Two gametes are missing many genes and are inviable
            - In female mammals the dicentric chromosomes are often shunted to the polar bodies so fertility is impacted less then in males
- **Diagram on Slide**
- Pericentric inversions
    - If a single cross over occurs within the inversions then we would expect
        - Two gametes that are viable, one with the genes in normal order and the other with the inversion
        - Two gametes that are not viable, each with some genes deleted and some duplicated
- **Diagram on Slide**
- Some crossover events within the inversion loop do not affect gamete viability
    - A double crossover close together involving the same 2 chromatids (two strand dco)
    - Changes where duplicated and deleted segments do not affect gene expression (very small segments)
    - In mammals, inverted segments may remain unpaired and so avoid crossing over

<u>Translocations</u>

- A change in the location of a DNA segment within the genome is a translocation. NO DNA is lost or gained (not the different with duplication). Simple translocations are of two types:
    - Intrachromosomal – with a change of position within the same chromosome
    - Interchromosomal – with transfer of the segment to a non-homologous chromosomes

    - If a segment is transferred from one chromosome to another, it is non-reciprocal
    - if segments are exchanged it is recirprocal
    -
- **Diagram on Slide**

Translocations and Gamete Formation

- In homozygotes with the same translocation on both chromosomes altered gene linkage is seen
  - Meiosis is fine like in inversions
- Gametes produced with chromosomal translocations often have unbalanced duplications and/or deletions and are not viable, or produce disorders such as familial Down Syndrome
- Strains that are homozygous for a reciprocal translocation form normal gametes
- Strains that are heterozygous for a reciprocal translocation must pair a set of normal chromosomes (N) with a set of translocated ones (T)
  - Result is a cross like configuration in meiotic prophase I of 4 associated chromosomes, each partially homologous to two others in the group
    - At the metaphase plate
      - 50% of the time inversions on opposite poles = viable
      - 50% of the time inversions go to the same poles = not viable
      - 50% infertility
    - Independent assortment can occur
- **Diagram on Next 2 Slides**

Translocations

- From these results it is clear that heterozygotes for a reciprocal translocation are considered semi-sterile
- Animal gametes with large duplications or deletions may function, but the zygote generally dies. Only small duplications or deletions may be viable

Robertisonian Translocation

- Centric function – produces 3 copies of the long arm of chromosome 21, resulting in familial Down Syndrome
- **Diagram on Slide with Example**
- Two non-homologous acrocentric (centromeres near the end) chromosomes break at the centromeres

  - Both long arms fuse at the same centromere, creating a chromosome with the long arm of chromosome 21 and the long arm of chromosome 14 (or 15)
  - The short arms fuse forming a reciprocal product that is usually lost within a few cell division
  - The heterozygous carrier of this chromosome is phenotypically normal since the two copies of each major chromosome arm supply two copies of all essential genes
- Mating of a heterozygous carrier and a normal individual has a high risk of Down Syndrome offspring
  - The normal parent produces normal gametes, with one copy of each of chromosomes 14 and 21
  - The heterozygous carrier parents produces three reciprocal pairs of gametes, each produced by different segregation of the 3 chromosomes involved
- Resulting zygotes
  - 1/6 with normal 14 and 21
  - 1/6 heterozygous carriers with normal phenotype
  - 1/6 not viable due to monosomy 14
  - 1/6 not viable due to monosomy 21
  - 1/6 in viable due to trisomy 14
  - 1/6 trisomy 21. Have normal chromosome # (46) but 3 copies of the long arm of 21, sufficient to cause Down Syndrome
- **Diagram on Slide**

Chromosomal Mutations and Human Tumors

- Most human malignant tumors have chromosomal mutations
  - The most common are translocations
  - Much variation in chromosome abnormalities, however, and they include simple rearrangements to complex changes in chromosome structure and number
  - Many tumor types show a variety of mutations
  - Some however are associated with specific chromosomal abnormalities

Examples of Specific Abnormalities

- Chronic myelogneous leukemia (CML) involves a reciprocal translocation of chromosomes 9 and 22
  - Myeloblasts (stem cells and white blood cells) replicate uncontrollably
  - 90% of CML patients have the Philadelphia chromosome (Ph^1) reciprocal translocation

    - reciprocal translocation causes transition from a differentiated cell to a tumor cell by moving a proto-oncogene from chromosome 9 to chromosome 22 and probably converting it the ABL oncogene
    - the hybrid gene arrangement causes expression of a leukemia producing gene product
- **Diagram on Slide**
- Burkitt Lymphoma (BL) involves a reciprocal translocation of chromosomes 8 and 14
    - Induced by a virus, this disease is common in Africa
    - B cells are affected and secrete antibodies as they proliferate
    - The reciprocal translocation positions the MYC proto-oncogene next to an active immunoglobin gene, resulting in over-expression of MYC and development of the lymphoma

Position Effect

- Sometimes inversions of tranlocations change phenotypic expression of genes by the postion effect
    - Moving a gene from a euchromatin (Active) to heterochromatin (not active)(transcription generally occurs in euchromatin but not in heterochromatin)
    - This is an example of an epigenetic effect since the DNA sequence of the gene is not effected
    - ABL and MYC are examples of this

Position Effect Example

- White eye (w) locus in Drosophila
    - An inversion moves the w+ gene from a euchromatin region on the X chromosome to a position in a heterochromatin
    - In a w+/Y males, or a w+/w+ female, where the w+ is involved in inversion, the eyes will have red eyes with white spots resulting from the cells where the w+ allele was moved and inactivated
- Associated with some human diseases. Aniridia (without iris) is one example
    - Is a severe hypoplasia of the iris, usually associated with cataracts and clouding of the cornea
    - The cause is early termination of eye development resulting from lost function of the PAX6 gene by deletion, mutation, or translocation

Fragile Site and Fragile X Syndrome

- Chromosomes in cultured human cells develop narrowing or unstrained areas (gaps) called fragile sites
    - Over 40 human fragile sites have been identified
- A well known example is fragile X syndrome, in which a region at position Xq27.3 is prone to breakage and mental retardation may result
- **Diagram on Slide**

Fragile X Syndrome

- Has an incidence in the US of about 1/1250 males and 1/2500 females (heterozygous)
- Inheritance follows Mendelian patterns, but only 80% of males with fragile X chromosome are affected. The 20% with fragile X chromosome but a normal phenotype are called normal but transmitting males
    - Can pass the chromosomes to his daughters
    - Sons of those daughters frequently show mental retardation
    - Higher degree as you go down through the generations

Molecular Bases for Fragile X

- Presence of a repeated 3-bp sequence, CGG, in the FMR-1 (fragile X mental retardation-1) gene at the fragile site
    - Normal individuals have 6-54 CGG repeats, with an average of 29
    - Normal transmitting carrier males, their daughter and some other carrier females have 55 to 200 copies, but do not show any symptoms
    - Individuals with fragile X syndrome have 200-1300 copies
        - Indicates that tandem amplification of this sequence is tolerated until a threshold number of copies is reached
    - Amplification even only occurs in females
        - Likely a mispairing process during DNA replication
    - FMR-1 gene product an RNA-binding protein. The triplet repeat affects gene expression rather than protein coding, resulting in loss of gene activity

Chapter 7 – Linkage and Chromosome Mapping in Eukaryotes

13-04-14 5:22 PM

What's Ahead

- Genetic recombination results from physical exchanges between homologous chromosomes in meiosis. A chiasma is the site of crossing-over where this reciprocal exchange takes place
- Map distance between genes is measured in map units (mu). 1 mu is the interval in which 1% of crossing over takes place
- As the distance between genes increases, the incidence of multiple crossovers causes recombination frequency to be an underestimate of the true map distance
- **Diagram on Slide**

Linked Genes

- Up until now, we have only looked at the inheritance of genes present on different chromosomes
- However many genes are located together on the same chromosome and can be inherited together
- Genes that do not assort independently are usually located on the same chromosome and exhibit what is called linkage and are called linked genes

Recombinants

- Crosses between two parents with different genetic characteristics result in progeny that have the same make-up as the parents (parental type) and some will have new combinations of the parental characteristics (recombinants)
- The process that produces the recombinants is called recombination

Early Studies of Genetic Linkage

- Morgan's experiment with Drosophila
    - By 1911 Morgan had identified several X-linked genes including white eye (w) and miniature wing (m)
    - Crossed a white miniature female (wm/wm) with a wild type male (w+m+/Y)
- **Diagram on Slide**

Morgan's Crosses

- All F1 males were white eyed and miniature winged (wm/Y) and all of the females were heterozygous and wild type for both eye color and wing size (w+m+/wm)
- The F1 flies were interbred, which in this case was essentially a test cross since the male was homozygous recessive for both genes

- 2441 F2 flies were analyzed
- **Diagram on Slide**

Parentals

- In the F2 the most common class of phenotypes was the parental type (or grandparental type in this case)
    - White eyed, miniature winged and wild type, wild type wings for a total of 1541 flies
    - These are the parentals or parental classes since they have the same genotype as the original parents

Recombinants

- Of the F2 flies 900 or 36.9% had nonparental phenotypic combinations
    - White eyed, wild type wings and wild type eyes, miniature wings
    - Nonparental combination of linked genes are called recombinants
- 50% recombinant frequency would be expected if independent assortment was occurring
    - Lower percentage observed above is an indication that the genes are linked
    - These recombinants are formed in meiosis with the exchange of genetic material in the F1 female
- **Diagram on Slide**

Recombination Observations

1. The parental classes are the most frequent phenotypes observed
2. Approximately equal numbers of the 2 parental classes are observed. Similarly, equal numbers of the 2 recombinant classes are also observed
3. The closer two genes are on the chromosome, the less likely that there will be a recombination event between them

Crossing Over

- A chiasma is the place on the homologous pair of chromosomes where the physical exchange takes place…. The site of crossing over
- Crossing over is a reciprocal exchange (usually) at positions along the homologous. The chromosomes are broken and rejoined in the process
- Crossing over occurs at the 4-chromatid stage in prophase I of meiosis. Each crossover involves only 2 chromatids, although multiple crossovers can occur along the whole length of the chromosome

DNA Recombination

- Crossing over involves the breaking and rejoining of DNA
- No loss or addition of bases
- Robin Holliday presented a model for recombination in the 1960's
    - Since the other researchers have modified this original model to fit their experimental observations

Holiday Model for Reciprocal Genetic Recombination

1. The 2 homologous DNA become aligned properly
2. A single strand in the 2 double helices is broken
3. Each broken strand invades the other double helix, and base pairs with the complementary nucleotides of the invaded helix
4. The cruciform structure separates to produce either parental or recombinant combinations

- **Diagram on Next 2 Slides**

Constructing Genetic Maps

- Genetic recombination experiments can be used in genetic or linkage mapping
    - We can examine the relative positions of genes along a chromosome
- When 2 genes are located on the same chromosome such they tend to be inherited together they are said to be linked genes

Detecting Linkage through Testcrosses

- A way to test for the linkage of a pair of genes is to analyze the results of a testcross
    - If the 2 genes assort independently (e.g. when they are on different chromosomes) we will observe rations associated with independent assortment
    - Data that deviate significantly from these ratios suggests that the genes may be linked

Detecting Linkage

- P1 (AABB) X P2 (aabb) – F1 all AaBb
- Testcross AaBb X aabb (homozygous recessive)
- Expect: AaBb:Aabb:aaBb:aabb in a 1:1:1:1 ratio when A and B are unlinked
- If significant deviation in this ratio is observed in the direction of too many parental types and too few recombinant types this suggests that these genes are linked
- How do we determine the significance… chi-square analysis
- Chi-square analysis is used to analyze testcross data and determine whether a deviation is significant

    - A null hypothesis (the genes are not linked) is used because it is not possible to predict phenotype frequencies produced by linked genes
- If 2 genes are not linked, a testcross should yield a 1:1 ratio of parentals:recombinants
- X^2 = sum of (d^2/e) (d=deviation value = (observed-expected)

<u>Chi-Square</u>

- The X^2 value and the degrees of freedom (df) for the data set are used with a table of chi-square probabilities to determine the probability (P) that the deviation of observed from expected values is due to chance
    - If P > 0.05 then the deviation not considered statistically significant
    - If P < 0.05 then the deviation is considered statistically significant and not due to chance. The null hypothesis is likely invalid (rejected)
    - If the hypothesis "the genes are not linked" is rejected, the only remaining option is that the genes are linked

<u>An Example</u>

- In drosophila
    - b is an autosomal recessive gene resulting in black body color, wild type flies have grey bodies
    - vg is an autosomal recessive gene resulting in vestigial (short and crumpled) wings, wild type wings are long and uncrumpled
- True breeding strains were crossed
    - Black, normal wings (b/b vg+/vg+) and
    - Grey, vestigial wings (b+/b+ vg/vg)
    - The F1 were all grey, normal winged (b+/b vg+/vg)
- An F1 female was testcrossed to a black, vestigial b/b vg/vg male
    - NOTE: the testcross was set up with the female as the heterozygous because crossing over does not occur in males
- Cross is b+/b vg+/vg x b/b vg/vg
    - Progeny
        - o 283 grey, normal
        - o 1294 grey, vestigial (parental)
        - o 1418 black, normal (parental)
        - o 241 black, vestigial
    - Total = 3236
        - o Parental = 2712 – should be 1618

        - Total /2 = 1618 but it is not
        - X^2 = 1479.4
        - Reject the null hypothesis that they are assorting independently
    - Recombination = 525
- Expect a 1:1 ratio of parental to recombinants if the genes are unlinked
- Our hypothesis is that these genes are assorting independently and we should observe a 1:1 ratio of parental to recombinant phenotypes
- Since chi-square value corresponds to a $P < 0.05$ on the table the deviation we observed is statistically significant
    - Reject the null hypothesis, so the genes are linked

The Concept of a Genetic Map

- In an individual heterozygous at two loci, there are 2 arrangements of alleles
    - Cis (coupling) arrangement has both wild type alleles on one homologous chromosome and both mutants on another
    - Trans (repulsion) arrangement has one mutant and one wild type on each homolog
- A crossover between homologs in the cis arrangement results in homologous pair with the trans arrangement. A crossover between homologs in the trans arrangement results in cis homologs
- **Diagram on Slide**
- Drosophila crosses showed that crossover frequency for linked genes (as measured by recombinants) is characteristic for each gene pair
    - This frequency stays the same, whether the genes are in coupling or in repulsion
- In 1913 Morgan and Sturtevant used recombination frequencies to make a genetic map
    - 1% crossover rate is a genetic distance of 1 map unit (mu). May also be called a centimogran (cM)
- Geneticists use recombination frequency as a way to estimate crossover frequency. It is not an exact measure
- However the further apart the two genes are on the chromosome, the more likely it is that that crossover will occur between them, and therefore the greater their crossover frequency
- With this data the genes on a chromosome can be represented with a genetic map showing the linear order of the genes, and their relative positioning along the chromosome

First Genetic Maps

- Based on crosses in Drosophila involving the three sex linked genes
    - W gives white eye

    - M gives miniature wings
    - Y gives yellow body
- Crosses gave the following recombination frequencies
    - W x m was 32.6
    - W x y was 1.3
    - M x y was 33.9
- **Diagram on Slide**

Genetic Mapping with a 2-Point Testcross

- With autosomal recessive alleles, when a double heterozygote is testcrossed, 4 phenotypic classes are expected. If the genes are linked the 2 parental phenotypes will be about equally frequent and more abundant than the two recombinant phenotypes
- **Diagram on Slide**
- Mapping of genes with other mechanisms of inheritance is also done with 2-point testcrosses
    - For x-linked recessives, a female double heterozygote is crossed with a male that is hemizygous for the recessive alleles
- The following formula is used to calculate the recombination frequency:
    - RF = (# of recombinants)/(# of test cross progeny) X 100
- **Diagram on Slide**

Recombination Frequencies

- The recombination frequency is used directly as an estimate of map units
- The measure is more accurate when the alleles are closer together
- Scoring large numbers of progeny increases the accuracy
- Mapping in all organisms show genes arranged with a 1:1 correspondence between linkage groups and chromosomes

Generating a Genetic Map

- A genetic map is generated from estimating the crossover rate in a particular segment of the chromosome
    - It may not exactly match the physical map because crossing over is not equally probable at all sites along the chromosome
- Recombination frequency is also used to predict progeny in genetic crosses
    - A 20% crossover rate between 2 pairs of alleles in a heterozygote (a+b+/ab) will give 10% gametes of each recombinant type (a+b and ab+)

Recombination Frequencies

- The recombination frequency in the progeny can never exceed 50%
    - If the genes are assorting independently (eg on different chromosomes) then we would expect an equal number of parental and recombinant classes (so the recombination frequency would be 50%)
    - If you see a recombination of 50% it shows that the genes are unlinked. There are 2 ways that genes can be unlinked
        - They may be on separate chromosomes
        - They may be far apart on the same chromosome
- **Diagram on Slide**
- Even when two genes are located very far apart on the same chromosome, it is generally impossible to have a recombination frequency greater then 50%
    - All combinations of a single and double crossover events leads to a 50% recombination frequency
- To determine if the genes are on the same chromosome, or on a different chromosome, other genes in the linkage group may be mapped in relation to a and b and used to deduce their location
- Single cross over – ½ and ½
- Double cross over – all parental
- 3 strand cross over – ½ and ½
- 4 strand cross over – all recombinants
- **Diagram on Slide**

Gene Mapping with 3-Point Testcrosses

- Typically, geneticists design experiments to gather data on several traits in one testcross
    - An example of a 3 point testcross would be p+r+j+/prj x prj/prj
- In the progeny, each gene has 2 possible phenotypes. For 3 genes there are 2^3 = 8 expected phenotypic classes in the progeny
    - Note that not all of the phenotypic classes may be seen in the progeny, which may give some insight into the rearrangement of the genes
- **Instructions Below on Slide (handwritten)**
- **Diagram on Next Slide**

Establishing the Order of the Genes

- The first step in mapping the three genes is to establish the order of the genes on the chromosome

- In a testcross, a parent that is heterozygous for all three genes will determine the phenotype of the progeny
    - If the heterozygous parents had all three genes in coupling, then the progeny that are triply heterozygous or triply homozygous recessive are the parental types (e.g. no crossover)
- The order of the genes can be deduced from results of the cross. Of the 8 expected progeny phenotypes:
    - 2 classes are the parental types and will be the most abundant
    - of the 6 remaining classes, 2 will be present at the lowest frequency, resulting from an apparent double crossover
        - notice that one will cross over to the other chromatid
        - the one that moves is the one that is in the middle
- **Diagram on Slide**

Calculating the Recombination Frequencies for Genes

- Cross data is organized to reflect the gene order now and in this example, the region between p and j is called region I and between j and r is region II
- **Diagram on Slide**
- Recombination frequencies are now calculated for two genes at a time. It includes single crossovers in the region under study, and double cross overs, since they occur in both regions
- Recombination frequencies are used to position genes on the genetic map (each 1% recombination frequency =1 map unit) for the chromosomal region
- Recombination frequencies are not identical to crossover frequencies, and typically underestimate the true map distance
- The recombination frequency is determined by (sco in region + dco)/(total progeny) X100
    - Where sco is the number of recombinant progeny resulting from a single crossover and dco is the number of recombinant progeny resulting from a double crossover
- **Diagram on Slide**

Interference and Coincidence

- Characteristically double crossovers do not occur as often as expected from the observed rate of single crossovers. Crossover appears to reduce formation of other chiasmata nearby, producing interference
    - Interference =1 is total interference with no other crossovers occurring in this region
- In our previous example we determined a recombination frequency between p and j to be 20.8 and between j and r to be 10.0

- If crossing over in region I is independent to crossing over in region II, then the probability of a double crossover is the product of the probabilities of the 2 events occurring separately.
    - (recombination frequency region I)/ 100 X (recombination frequency region II)/100
    - =0.208 X 0.100 = 0.0208 or 2.08%
    - However only 6/500 = 1.2% double crossovers occurred in this cross
- Usually the double crossover progeny do not appear as often as one would expect based on the map distance between the genes
- In some ways, the presence of one crossover interferes with the formation of another crossover event nearby…. Interference
- The extent of interference is expressed as a coefficient of coincidence
    - Coefficient of evidence = (observed dco frequency)/(expected dco frequency)
    - Interference = 1 – coefficient of evidence
- In the example we gave earlier the coefficient of coincidence is 0.012/0.0208 = 0.577
- Coefficient of coincidence
    - When it is 1 it means that there is no interference since the observed and the expected number of dco events are equal
    - When it is 0 that means that there is total interference such that none of the expected dco events occurring. Here, is a single crossover interference with other crossover events in the region, interference =1
- Note that the coefficient of coincidence and interference values are inversely related

Calculating Accurate Map Distances

- Recombination frequency generally underestimates the true map distance
    - Double crossovers between two loci will restore the parental genotype as will any even number of crossovers. These will not be counted as recombinants even though crossovers have taken place
    - A single crossover as well as any odd number of crossovers will produce recombinant chromosomes. Progeny analysis assumes that all recombinants are produced by only sco events
    - Map distances for genes that are less then 7 mu apart are very accurate. As distances increase the accuracy declines because crossovers go uncounted
- **Diagram on Slide**

Somatic Cell Hybridization

- Two cells in culture can be induced to fuse into a single hybrid cell
- Produces a heterokaryon

  - Two nuclei in one cell
  - Nuclei eventually fuse
  - Chromosomes from one parent species is gradually lost
- Product is a cell with a full complement of chromosomes from one species and just a few chromosomes from another species
- Way to find insulin gene on chimeric line
- Fine the chromosome that is common
- In human-mouse hybrid cell lines, humans chromosomes are lost
- Many of these cell lines are produced to create a library of cell lines with different combinations of human chromosomes present
  - Can assign genes to chromosomes
- May look where # is not expressed and look backwards (e.g. not a one-three)
- Process of elimination
- **Diagram on Slide**

Gene to Centromere Mapping

- In organisms that form an ordered tetrad we can map the genes to its centromere
- In neurospora, a single gene can be mapped to its centromere
  - If a crossover occurs between the gene and the centromere a second division segregant is produced (meiosis II)
  - If no crossover occurs then a first division segregant is produced (meiosis I)
  - Spore color
- **Example on Next 2 Slides**

Chapter 9 – DNA Structure and Analysis

Four Characteristics of Genetic Material

- Some substance must be responsible for passage of traits from parents to offspring. For a substance to do this it must be able to
    1. Replicate
    2. Store information
    3. Express the information
    4. Change by mutation (the driving force of evolution)
- **Diagram on Slide**

The Search for Genetic Material

- In the early 1900's chromosomes were shown to be the carriers of hereditary information. In eukaryotes they are composed of both DNA and protein, and most scientists initially believed that protein must be the genetic material
    - Proteins are far more complex then DNA
        - 20 different amino acids
    - DNA is less complex
        - 4 different nucleotides

Griffith's Transformation Experiment

- 1928 experiment with streptococcus pneumonia bacteria in mice showed that something passed from dead bacteria into nearby living ones, allowing them to change their cell surface
- he called this agent the transforming principle, but did not know what it was or how it worked
- **Diagram on Slide**

Avery's Transformation Experiments

- In 1944, Avery, MacLeod, and McCarty published results of a study that identified the transforming principle from S. pneumnoiae. Their approach was to break open dead cells, chemically separate the components (e.g. protein, nucleic acids) and determine which was capable of transforming like S. pneumonia cells
- Only nucleic acid fraction was capable of transforming bacteria
- Critics noted that the nucleic acid fraction was contaminated with proteins. The researchers treated this fraction with either RNase or protease and still found transforming activity, but when it was treated with DNase, no transformation occurred, indicating that the transforming principle was DNA
- **Diagram on Slide**

The Hershey-Chase Bacteriophage Experiments

- In 1953, more evidence for DNA as the genetic material resulted from Alfred Hershey and Martha Chase's work on E. coli infected bacteriophage T2

T2 Phage Biology

- Never enters the cell only introduced the DNA
    - Was it DNA or protein?
- **Diagram on Slide**

The Hershey-Chase Bacteriophage Experiments

- In one part of the experiment, T2 proteins were labeled with S35 and in the other part, T2 DNA was labeled with 32P
    - Then each group of labeled viruses was mixed separately with the E. coli hos. After a short time, phage attachment was disrupted with a kitchen blender, and the location of the label determined
- Results
    - The 35S-labeled was found outside the infected cells, while the 32P-labeled DNA was inside the E. coli, indicating that DNA carried the information needed for virual infection
        - This provided additional support for the idea that genetic inheritance occurs via DNA
- Cysteine has a sulfur group on it
- Grew in S (medium) and P (medium)
- Looked for the phages to inject DNA
- Wanted to get rid of phages and separate
- Wanted to see radioactivity in cells
- **Diagram on Slide**

Indirect Evidence: Distribution of DNA

- DNA doubled in diploid cells
- If you look at protein the same pattern occurs
- **Diagram on Slide**

Indirect Evidence: Mutagenesis

- Action spectrum

  - Shows that UV light is optimal mutagenic wavelength
- Absorption spectrum
  - DNA absorbs light optimally in UV range… unlike protein
- Therefore DNA is the genetic material
  - Genetic material must also absorb at that wavelength (DNA)
  - Protein is at 280 nm
- **Diagram on Slide**

The Discovery of RNA as Viral Genetic Material

- All known cellular organisms have DNA as their genetic material. Some viruses, however, use RNA instead
- Tobacco mosaic virus (TMV) is composed of RNA and protein; it contains no DNA. In 1956 Giere and Schramm showed that when purified RNA from TMV is applied directly to tobacco leaves, they develop mosaic disease. Pretreating to purified RNA with RNase destroys its ability to cause TMV lesions
- In 1957 Fraenkel-Conrat and Singer showed that in TMV infections with viruses containing RNA from one strain and protein from another, the progeny viruses were always of the type specified by the RNA, not by the protein
- **Diagram on Slide**

The Composition and Structure of DNA and RNA

- DNA and RNA are polymers composed of monomers called nucleotides
- Each nucleotide has 3 parts
  - A pentose (5-carbon) sugar
  - A nitrogenous base
  - A phosphate group (link at 5’ and 3’)
- The pentose sugar in RNA is ribose, and in DNA it is deoxyribose. The only difference is at the 2’ position, where RNA has a hydroxyl (OH) group, while DNA only has a hydrogen
- **Diagram on Slide**

There are Two Classes of Nitrogenous Bases

- Pyrimidine’s (one ring, six membered structure) include cytosine (C), thymine (T) in DNA, and uracil (U) in RNA
- Purines (double ring, nine membered structure) include adenine (A) and guanine (G)
- **Diagram on Slide**

The Structure of Nucleotides

- The base is always attached by a covalent bond between the 1' carbon of the pentose sugar and a nitrogen in the base (specifically, the 9 nitrogen in purines and the 1 nitrogen in pyrimidines)
- The sugar base combination is a nucleoside. When a phosphate is added (always to the 5' carbon of the pentose sugar), it becomes a nucleoside phosphate, or simply a nucleotide
- **Diagram on Slide**
- Oligonucleotides (chains of up to 20 nucleotides) and long chains, polynucleotides of both DNA and RNA are formed by stable covalent bonds (phosphodiester bonds) between the phosphate group on the 5' carbon of one nucleotide, and the 3' hydroxyl on another nucleotide. This creates the "backbone" of a nucleic acid molecule
- The asymmetry of phosphodiester bonds creates a 3'-5' polarity within the nucleic acid chain
- Antiparallel: 5'-3' and 3'-5'
- **Diagram on Slide**

Chargaff's Rule

- After hydrolyzing DNA for a number of organisms, he showed that in all doubled stranded DNAs, 50% of the bases were purines and 50% were pyrimidines
- The amount of A equals the amount of T and the amount of G equals the amount of C
    - A=T and G=C
    - (A+G)/(T+C)=1
    - (A+T)/(G+C)=species dependent
- **Diagram on Slide**

The DNA Double Helix

- James Watson and Francis Crick published the famous double helix structure in 1953. When they began their work, it was known that DNA was composed of nucleotides, but how the nucleotides were assembled into nucleic acid was unknown. Two important additional sources of data assisted Watson and Crick with their model
    - Chargaff's base composition studies
    - Rosalind Franklin's X-ray diffraction pattern of DNA
- Didn't do any experiments – used work of others

Rosalind Franklin's X-Ray Diffraction Studies

- While working in the same lab as Maurice Wilkins, was able to create high resolution X-ray diffraction pattern for DNA

- From these diffraction patterns, she was able to determine that DNA was helical with two regularities of 0.34nm and 3.4nm along the axis of the molecule
- Dehydrate at a controlled rate
- This is how it crystalized
- Crystal – all molecules line up at the same orientation
- Shoot an x-ray through and see how they're diffracted
- Space = pass through, atoms = diffract
- **Diagram on Slide**

Features of the Watson and Crick Model

- It is two polynucleotide chains wound around each other in a right handed helix
- The two chains are antiparallel
- The sugar phosphate backbones are on the outside of the helix, and the bases are on the inside, staked perpendicularly to the long axis of the step of a spiral staircase
- Between bases is 3.4 A
- For a complete turn it is 34 A
- There are minor and major grooves
- Why is DNA refereed to as a double helix?
  - Both strands have a helix, 10 bases make up one turn
- **Diagram on Slide**

The DNA Zipper

- The bases of the two strands are held together by hydrogen bonds between complementary bases (2 for A-T pairs, and 3 for G-C pairs)
- Individual H bonds are relatively weak and so the strands can be separated (by heating, for example)
- Complementary base pairing means that the sequence of one strand dictates the sequence of the other strand
- **Diagram on Side**

Features of Watson and Crick Model: Recap

- The base pairs are 0.34nm apart, and one full turn of the DNA helix takes 3.4nm, so there are 10bp in a complete turn. The diameter of a dsDNA helix is 2nm
- Because of the way the bases H-bond with each other, the opposite sugar phosphate backbones are not equally spaced, resulting in a major and minor groove. This feature of DNA structure is important for protein binding

The Nobel Prize for Medicine 1962

- Watson, Crick, and Wilkins won
- Franklin's x ray diffraction patterns were shown to Watson without her knowledge, which 4 weeks later led to construction of the model by Watson and Crick
- Franklin's contribution was never formally acknowledged
- Franklin died in 1958 of ovarian cancer likely the result of her work with X-rays. Nobel prizes are never awarded posthumously

Different DNA Structures

- X ray diffraction studies show that DNA can exist in different forms
    - A-DNA is the dehydrated form, and so it is not usually found in cells, it is a right handed helix with 10.9bp/turn, with the bases inclined 13* from the helix axis. Has a deep narrow major groove and a wide and shallow minor groove
    - B-DNA is the hydrated form of DNA, the kind normally found in cells
        - Franklin worked with this
    - Z-DNA is a left handed helix with a zigzag sugar phosphate backbone that gives it its name. It has 12bp/turn, with the bases inclined at 8.8* from the helix axis. Has a deep minor groove and a very shallow major groove. Its existence in living cells has not been proven
- **Diagram on Slide**

DNA in the Cell

- All known cellular DNA is in the B form. A DNA would not be expected because it is dehydrated and cells are aqueous. Some organisms show evidence of Z DNA, but its physiological role, if any, is unknown

RNA Structure

- RNA structure is very similar to that of DNA
    - It is a polymer of ribonucleotides (the sugar is ribose rather than deoxyribose)
    - Three of its bases are the same (A, G, and C) while it contains U rather then T
- Is single stranded, but internal base pairing can produce secondary structure in molecule
- Some viruses use RNA for their genomes. In some it is dsRNA, while in other it is ssRNA. Double stranded RNA is structurally very similar to dsDNA

Denaturation and Renaturation of DNA

- Heating up a DNA sample causes the DNA double helix to denature and unwind
  - Breaking the hydrogen bonds that hold the opposite strands together but leaving the covalent bonds intact
- During heating viscosity decreases and UV absorption increases
  - Hyperchromic shift
- **Diagram on Slide**
- Disassociation kinetics – denaturation as temperature increases
- Reassociation kinetics – going from high temp to low to renature DNA
- If it plateaus multiple times it means that there are repeated sequences in the DNA
- Denatured DNA can reform a double helix if allowed to cool slowly over time
- Most proteins cant be renatured
- DNA is a strong structure
  - May be used for molecular hybridization experiments
    - DNA probes can be synthesized for particular genes or regions of DNA
    - FISH (fluorescent in situ hybridization) is one example

FISH

- Human metaphase chromosomes hybridize with a centromere specific fluorescent probe
- **Diagram on Slide**

Chapter 10 – DNA Replication and Recombination

13-04-14 5:22 PM

DNA Replication

- Once DNA structure was solved, research into how DNA was replicated was the focus of much attention
- Consider human genome
    - 3 * 10^9 (3 billion) base pairs of DNA
    - Error rate of 1 * 10^6
        - 3000 error would be made each time the DNA was replicated
            - In 50 divisions there would be 150 000 errors
            - DNA replication makes some errors, but no where near this number
                - Remember, without changes in DNA sequence, evolution can't occur

DNA Replication is Semiconservative

- Watson and Crick suggested a semiconservative mode of DNA replicated based on their structure
    - Each strand of the template could serve as a template for synthesis of a complement strand
- End with 2 strands that are the same as the beginning and 2 that have the new DNA incorporated as the other strand
- **Diagram on Slide**

Different Theoretical Modes of Replication

- Semiconservative replication – has each newly replicated DNA molecule consisting of one old strand and one new strand
- Conservative replication – will have two DNA molecules, one molecule with both strands composed of old DNA and one molecule with both strands of new DNA
- Dispersive replication – has both strands composed of a mixture of old and new DNA
- **Diagram on Slide**

Proof for Mode of DNA Replication

- The Meselson-Stahl experiment
    - Grew bacterical cells (E.coli) for many generations in a medium supplements with ^15NH4Cl as the only nitrogen source
        - Remember the nitrogenous bases have nitrogen atoms, so the DNA will now be labeled with this "heavy" version of nitrogen
            - N^15 contains one more neutron than naturally occurring N^14

        - Molecules containing N15 are more dense then molecules containing N14
        - Importantly DNA containing N15 can be distinguished from DNA containing N14

Sedimentation Equilibrium Centrifugation

- Samples forced by centrifugation into a density gradient of a heavy metal salt solution
    - Usually CsCl (cesium chloride)
- Molecules of DNA will migrate into the density gradient
    - The DNA will stop moving when it reaches a region of CsCl that matches its own density
    - 15N-DNA will move further down the tube of CsCl compared to 14N-DNA
    - 14N will increase with each division (dispersive)
- **Diagram on Next 4 Slides**

Semiconservative Replication in Eukaryotes

- In 1957 Taylor, Woods and Hughs provided evidence of semiconservative replication in broad bean plants
- Technique used was autoradiography
    - Can locate location of radioactive isotope in a cell
    - A photographic emulsion is placed over a histological section/sample and placed in the dark
        - Radioactivity exposes the emulsion just like would expose a photographic film
        - Exposure of emulsion results in tiny black dots
            - Locations of black dots = location of radioactive isotope in cell

Taylor et al. Experiment

- Labeled DNA with H^3 thymidine for one generation
    - Then placed in "cold" medium (no isotope)
- Arrested cells at metaphase with colchicine over two division cycles
    - Performed autoradiography
        - Mess up microtubules – tons of cells in metaphase
- Results, at first metaphase both chromatids labeled, while in second metaphase only one chromatid labeled…. Semiconservative replication
- **Diagram on Slide –** recombination (sister chromatin exchange)

Origins, Forks and Units of Replication

- Origin of replication
    - Location where DNA replication is initiated
    - Replication is bidirectional from the origin
- Replication fork is where DNA strands are being unwound
    - Initially appear at origin and then 2 forks move apart in opposite directions
- A replicon represents how much DNA is replicated following one initiation event
    - E.g. E. Coli has only one origin of replication (oriC) for its entire 4.6 million base pairs (bp) of circular DNA
- **Diagram on Slide**

DNA Polymerase, the DNA Replicating Enzymes

- First isolation of an enzyme involved in DNA replication was in 1955. Arthur Kornberg won the 1959 Novel Prize in physiology or medicine for the discovery of mechanisms involved in DNA replication (synthesis)

DNA Polymerase I

- Kornberg accomplished in vitro synthesis of E. coli DNA. His reaction mixture included:
    - A DNA fragment (acts as template)
    - Radioactively labeled dNTPs (dATP, dGTP, dTTP, and dCTP)
    - DNA polymerase I (DNA pol I)
- Enzyme originally called the Kornberg enzyme, now known as DNA poly I. Once isolated, its activity could be characterized, showing that the above components required, along with Mg^2+ ions for maximum activity
- Why dNTPs?
    - Already has cold DNA
    - Going to run on a gel with a control (no poly)
    - Cold DNA shows a band
    - New DNA will be radioactive
    - Place X-ray film on gel, expose and see hot DNA

Roles of DNA Polymerase

- All DNA polymerases link dNTPs into DNA chains. Main features of the reaction:

  - An incoming nucleotide is attached by its 5'-phosphate group to the 3'-OH of the growing DNA chain. Energy comes from the release of two phosphate from the dNTP. The DNA chain acts as a primer for the reaction
    - Indication of evolution
      - Every organism does it this way
  - The incoming nucleotide is selected by its ability to hydrogen bond with the complementary base in the template stand. The process is fast and accurate for dsDNA
  - DNA poly synthesize only from 5' to 3'
- **Diagram on Next 2 Slides**

<u>More than 1 DNA Polymerase?</u>

- The enzyme Kornberg isolated was believed to be the only DNA poly in E. coli. However, mutations in this gene (polA1) were not lethal, indicating that other DNA polymerase must exist in E. coli
- Temperature-sensitive mutants are used to study essential genes. At 37* polAex1 strain are normal, and the protein has normal activity in vitro. At 42* the protein lacks 5'-3' exonuclease activity, and bacterial cells with this mutation are dead
  - Conditional mutants
- Additional DNA poly have been isolated, including DNA poly II (1970), DNA poly III (1971), DNA poly IV, and DNA poly V
- Also cold sensitive mutants – function at 37* but not at 24*
- **Table on Next Slide**

<u>Initiation of Replication in E.coli</u>

- Initiator proteins attach. E. coli's initiate protein DnaA (from the dnaA gene) finds the oriC
- DNA helicase (from dnaB) binds initiator proteins on the DNA, and denatures the AT-rich region using ATP as an energy source
- DNA primase (from dnaG) binds helicase to form a primosome, which synthesizes a short (5-10 nucleotide) RNA primer

<u>Events Occurring in the Replication Form in E. coli</u>

- When DNA denatures at the ori, replication forms are formed. DNA replication is usually bidirectional, but only one form is shown here
- Single strand DNA binding proteins (ssBs) binds the ssDNA formed by helicase preventing reannealing

- Primase synthesizes primer on each template strand
- DNA poly III adds nucleotides to the 3' end of the primer, synthesizing a new strand complementary to the template, and displacing the SSBs. DNa is made in opposite direction on the two template strands
- New strand made 5' to 3' in the same direction as movement of the replication fork is leading strand, while new strand is made in the opposite direction is lagging strand. Leading strand only one primer, while lagging strand needs a series of primers

1. Helicase – separate strands
2. Primase – makes RNA
3. DNA ply III – make leading strand
4. Ss binding proteins – prevent reannealing (before helicase when ss regions come in)
5. DNA poly I
6. Ligase – seals nicks in DNA (gaps)
7. Gyrase – topoisomerase – nicks a strand, free bond rotation relieving positive supercoiling

RNA Replication in E.coli

- Helicase denaturing DNA causes tighter winding in other parts of the circular chromosome. Gyrase relieve this tension
- Leading strand is synthesized continuously, while the lagging strand is synthesized discontinuously, in the form of Okazaki fragments. DNA replication is therefore semidiscontinuous
- _Each fragment requires a primer to begin, and is extended by DNA polymerase III
- **Diagram on Slide**

Concurrent Synthesis of Leading and Lagging Strands

- **Diagram on Next 2 Slides**
- Folded over – one complex going the same way

Okazaki Fragments

- Okazaki data show that these fragments are gradually joined together to make a full length dsDNa chromosome. DNA Poly I uses the 3'-OH of the adjacent DNA fragment as a primer, and simultaneously removes the RNA primer while resynthesizing the primer region in the form of DNA. The nick remaining between the two fragments is sealed with DNA ligase
- **Diagram on Slide**

Replication of Circular DNA and Supercoiling

- Some circular chromosomes (e.g. E. coli) are circular throughout the replication, creating a theta-like shape. As the strands separate on one side of the circle, positive supercoils form elsewhere in the molecule. Replication fork moves about 500nt/second, so at 10bp/turn, replication form rotates at 3000 rpm
- Topoisomerases (Gyrase) relieve the supercoils, allowing the DNA strands to continue separating as the replication forks advance
- **Table on Next Slide**

DNA Replication in Eukaryotes

- Very similar to prokaryotes
- Added complexity since DNA is distributed among many chromosomes
- Added complexity since there is much, much more DNA to replicate
    - 3 billion base pairs of DNA in haploid human genome
    - Replication fork in eukaryotes is slower then in prokaryotes
        - Would take many days to replicate the genome

Replication in Drosophila Melanogaster

- The diploid set of chromosomes in D. melanogaster replicates in 3 minuets, faster then E. coli even though there is over 100 times more DNA to replicate
- Unlike E. coli, eukaryotic chromosomes have multiple replication origins
    - Do not all start at the same time
- **Diagram on Slide (Drawn)**

Replicons

- Eukaryotic chromosomes contain multiple origins, at which DNA denatures and replication then proceeds bidirectionally until an adjacent replication fork is encountered. The DNA replicated from a single origin is called a replicon, or replication unit
- In eukaryotes, replicon size is smaller then it is in prokaryotes, replication is slower, and each chromosome contains many replicons. Number and size of replicons vary with cell type

Temporal Ordering of DNA Replication

- Not all origins within a genome initiate DNA synthesis simultaneously. Cell specific patterns of origin activation are observed, so that chromosomal regions are replicated in a predictable order in each cell cycle
- Cant transcribe the DNA as it is being replicated (S and into the next stage cannot occur)

- **Diagram on Slide**

Initiation of Replication

- Eukaryotic origins are generally not well characterized; those of the yeast Saccharomyces cerevisiae are among the best understood
- Chromosomal DNA fragments (about 100 bp) that are able to replicate autonomously when introduced into yeast as extracellular, circular DNA are known as ARSs (autonomously replicating sequences)
- ARSs are yeast replicators that contain conserved regions necessary for the initiation of replication
- Initiator protein in yeast is the multiprotein origin recognition complex (ORC), which binds within the ARS sequence. Other replication proteins join, including one that unwinds DNA within the ARS sequence

Eukaryotic Replication Enzymes

- Enzymes in eukaryotic DNA replication aren't as well characterized as those in prokaryotes. The replication process is similar in both groups: DNA denatures, replication is semiconservative and semidiscontinuous, and primers are required
- 15 DNA polymerases are known in mammalian cells
    - 3 DNA polymerases are used to replicate nuclear DNA. Poly alpha extends the 10-nt RNA primer by about 30 nt. Poly delta and Poly epsilon extend the RNA/DNA primers, one on the leading strand and the other on the lagging strand (it is not clear which synthesizes which)
    - Other DNA polys replicate mitochondrial or chloroplast DNA, or are use in DNA repair

Replicating the Ends of Chromosomes

- When the ends of chromosomes are replicated and the primers are removed from the 5' ends, there is no adjacent DNA strand to serve as a primer, and so a single stranded region is left at the 5' end of the new strand. If the gap is not addressed, chromosomes would become shorter with each round of replication
- Most eukaryotic chromosomes have short, species specific sequences tandemly repeated at their telomeres. Blackburn and Greider have shown that chromosome lengths are maintained by telomerase, which adds telomere repeats without using the cells regular replication machinery
- Cells can only replicate so many times before it will die

    - Essential genes are lost
- **Diagram on Slide**

Telomerase

- In the ciliate Tetrahymena, the telomere repeat sequence is 5'-TTGGGG-3'
- Telomerase, an enzyme containing both protein and RNA, binds to the terminal telomere repeat when it is single stranded, synthesizing a 3-nt sequence, TTG
- The 3' end of the telomerase, RNA containing the sequence AAC, which binds the TTG positioning telomerase to complete its synthesis of the TTGGGG telomere repeat
- Additional rounds of telomerase activity lengthen the chromosome by adding telomere repeats
- There is always going to be a little gap that is going to be lost
- RNA is complementary to the repeat on the telomerase – starts to pair up and gets 3'-OH hangover
- **Diagram on Slide**

Telomeres

- After telomerase adds telomere sequences, chromosomal replication proceeds in the usual way. Any shortening of the chromosome ends is compensated by the addition of the telomere repeats
- If the sequence of the telomerase RNA is mutated, telomeres will correspond to the mutant sequence, rather then the organism's normal telomere sequence. Using an RNA template to make DNA, telomerase functions as reverse transcriptase is called TERT (telomerase reverse transcriptase)
- Telomere length may vary but organisms and cell types have characteristic telomere lengths. Mutants affecting telomere length have been identified and data indicate that telomere length is genetically controlled. Shortening of telomeres eventually leads to cell death, and this may be a factor in the regulation of normal cell death
    - Found in HIV transposome
    - Telomerase enzyme

Figure 10-17

- Telomerase binds to 3' G-rich tail
- Telomeric DNA is synthesized on G-rich tail
- Telomerase is translocated and steps (a) and (b) are repeated
- Telomerase released; primase and DNA polymerase fill gap

- Primer removed; gap sealed by DNA ligase
- **Diagram on Slides**

DNA Recombination

- Holliday model shows how DNA is cut with endonuclease to produce ssNick
    - ssDNAs displaced and joined by ligase to other strand
        - creates hybrid duplexes called heteroduplex DNA
- **Diagram on Slide**

Gene Conversion

- Initially discovered in fungi, an interesting non-Mendelian inheritance pattern was found
    - In Neurospora asci of a Aa individual, 2:2 patterns were usually seen (AAaa), but sometimes a 4:0 patterns was observed (AAAA)
- Gene conversions are the result of DNA repair in heteroduplex regions following crossing-over
    - Mispairing of bases in heteroduplex will be targeted for repair… problem is the repair enzymes don't know which is wt. and which is mutant (mt)
    - Random which way it is repaired (wt to mt or mt to wt)
- Doesn't know which ones right and which ones wrong
    - Crossing over occurs in order, to do this
    - Only in something that has ordered meiosis
- **Diagram on Slide**

Chapter 12 – The Genetic Code and Transcription

13-04-14 5:22 PM

<u>Characteristics of Genetic Code</u>

- Code is linear using mRNA molecules as letters, where the mRNA sequence is derived from DNA
- Each amino acid is specified by a 3 ribonucleotides in the mRNA. These 3 nts are called a codon
- Code is unambiguous, each triplet specifies a single AA
- Code is degenerate with AA's specified by more then one triplet codon
- Code contains start and stop signals that initiate and terminate translation
- There is no internal punctuation. Once the code is read, it continues without any breaks
- Code is nonoverlapping (only starts on one AUG, doesn't restart on another AUG in sequence)
- Code is nearly universal from viruses to humans

<u>The Triplet Code</u>

- Sidney Brenner (1960's) suggested code was a triplet coed
  - With 20 AA there had to be at least 20 different "words" in the code
    - 1 nt codon = 4 words (4^1)
    - 2 nt codon = 16 words (4^2)
    - 3 nt codon = 64 words (4^3)
      - A 3 nucleotide codon was the smallest size to accommodate all 20 AA

<u>Experimental Evidence for Triplet Code</u>

- Francis Crick studied frameshift mutations in T4 phage
  - Additions of 1 or 2 nucleotides by mutagenesis resulted in a defective product following translation
  - However, if 3 nucleotides were added, the function came back
- **Diagram on Slide**

<u>Deciphering the Code</u>

- Nirenberg and Matthaei (1961) developed an in vitro (cell free) translation system and the ability to produce synthetic mRNAs with polynucleotide phosphorylase
  - Others used these tools to further our knowledge of the code
- **Diagram on Slide**

Homopolymers

- Synthesizing homopolymers of U, A, C and G they tested each of these synthetic mRNAs with 20 different mixtures of AA's
  - In each AA mixture, only one AA was radioactive
- **Table on Slide**
- UUU codes for phenyalanine
- AAA codes for lysine
- CCC codes for proline
- GGG – couldn't get it to work

Mixed Copolymers

- Mixtures of different ribonucleotide diphosphates added in the mixtures
  - The relative proportion of each ribonucleotide should predict the relative proportion of the triplet codon in the synthetic mRNA
- E.g. A & C added in a ratio of 1A:5C
  - Chance of AAA = (1/6)^3 = 0.4%
  - Chance of AAC, ACA and CAA = 3 x (1/6)^2 x (5/6) = 6.9%
  - Chance of CCA, CAC and ACC = 3 x (1/6) x (5/6)^2 = 11.6
  - Chance of CCC = (5/6)^3 = 57.9%
- **Diagram on Next Slide**

Triplet Binding Assay

- Nirenberg and Leder (1964) were able to synthesize specific 3 ribonucleotide long RNAs in vitro
  - Ribosomes bond to these small RNAs and a tRNA with the proper AA attached will associate with the ribosome RNA complex
    - Unbound tRNAs are washed away (since they are much smaller) and the bound tRNA AA is released and identified
    - Not making a protein
- **Diagram on Slide**

Repeating Copolymers

- About 50/64 codons were identified, but the remaining codons were difficult to get unambiguous results
- Khorana developed a way to create repeating ribonucleotide patterns in synthetic RNAs
  - E.g. UUCUUCUUCUUCUUC or UCUCUCUCUCUCUCU

- No AA corresponds to stop codons
- **Diagram on Next 3 Slides**

Degeneracy and Wobble

- There is more then one codon for most of the AA's… the code is degenerate
- Francis and Crick observed that there is a pattern in the degeneracy in the 3rd position
    - E.g. CCC, CCU, CCA, and CCG all code for proline
    - Crick reasoned that the first 2 ribonucleotides of the codon are more important then the third attracting the correct tRNA
- Wobbel is the 3rd in the mRNA and the 1st in the tRNA
- tRNA = anticodon
- mRNA = codon
- **Diagram on Slide**

Initiation and Termination

- AUG codes for methionine (the only one that does) and is the initiator codon
    - In bacterial the initial AUG codon specifies a modified methionine called N-formylmethionine
    - All other AUG codons in the message code for the insertion of methionine in the protein
- UAG, UAA, and UGA are termination codons
    - They do NOT code for any AA
    - They are NOT recognized by and tRNA
    - Causes TRANSLATIONAL TERMINATION

The Transcription Process

- RNA synthesis
- Transcription, or gene expression, is regulated by gene regulatory elements associated with each gene
- DNA unwinds in the region next to the gene, due to RNA polymerase in prokaryotes and other proteins in eukaryotes. In both, RNA polymerase catalyzes transcription
- Enzyme for DNA – RNA is RNA poly
- Upstream – promoter
- Downstream – not transcribed may have embedded codes to terminate

RNA Synthesis

- RNA is transcribed 5'-3'. The template DNA strand is read 3'-5', it is complementary DNA, the nontemplate strand, has the same polarity as the RNA
- RNA polymerization is similar to DNA synthesis except:
    - The precursors are NTPs not dNTPs
    - No primer is needed to initiate synthesis
    - Uracil is inserted instead of thymine
- **Diagram on slide**

Transcription

- Occurs in three steps for both prokaryotes and eukaryotes
    - Initiation…. Promoter
    - Elongation…. RNA coding sequence
    - Termination….. transcription stop by terminator
- The process of elongation is highly conserved btwn prokaryotes and eukaryotes, but initiation and termination are somewhere different
- Transcription start and stop – not AUG and stop codons
- **Diagram on slide**

Prokaryotic Transcription

- A prokaryotic gene is a DNA sequence in the chromosome. The gene has 3 regions, each with a function in transcription
    - A promoter sequence that attracts RNA polymerase to begin transcription at a site specified by the promoter
    - The transcribed sequence called the RNA coding sequence. The sequence of this DNA corresponds with the RNA sequence of the transcript
    - A termination region that specifies where transcription will stop
- Transcription initiation site is at +1
- **Diagram on Slide**

E. coli Promoters

- Generally involve 2 DNA sequences, centered at -35 bp and -10 bp upstream from the +1 start site of TRANSCRIPTION
- The common promoter that is used for most transcription has these consensus sequences
    - For the -35 region the consensus is 5'-TTGACA-3'

    - For the -10 region (previously known as a Pribnow box), the consensus is 5’-TATAAT-3’ (easy to pull apart due to only having 2 hydrogen bonds)
- **Diagram on Slide**

Transcription Initiation

- Requires the RNA poly holoenzyme to bind to the promoter DNA sequence. Holoenzyme consists of:
    - Core enzyme of RNA poly, containing 4 polypeptides (2 alpha, one beta, and one epsilon)
    - Sigma factor
- Sigma factor binds the core enzyme, and confers ability to recognize promoters
- RNA poly holoenzyme binds promoter in 2 steps
    - Loosely binds to the -35 sequence of dsDNA (closed promoter complex)
    - Then binds tightly to the -10 sequence, untwisting about 17 bp of DNA at the site, and in position to being transcription (open promoter complex)
- Once it starts the sigma factor isn’t needed and it is released
- **Diagram on slide**

Sigma Factors

- Promoters often deviate from consensus. The associated genes will show different levels of transcription, corresponding with sigma’s ability to recognize sequence
- E. coli has several sigma factors with important roles in gene regulation. Each sigma can bind a molecule of core RNA poly and guide its choice of genes to transcribe
- Most E. coli genes have a sigma 70 promoter, and sigma 70 is usually the most abundant sigma factor in the cell. Other sigma factors may be produced in response to changing conditions
- E. coli has additional sigma factors. Other bacterial species also have multiple sigma factors
- Sigma factor
    - Weather the genes are expressed
    - If not then the genes wont be expressed
    - Binds to TTGACA (sigma 70) and starts it
- **Diagram on slide**

Elongation and Termination of an RNA Chain

- Once initiation is completed RNA synthesis begins, after 8-9 NTPs have been joined in the growing RNA chain, the sigma factor is released and reused for other initiations. Core enzymes completes the transcript
- Core enzyme untwists DNA helix locally, allowing a small region to denature. Newly synthesized RNA forms an RNA-DNA hybrid, but most of the transcript is displaced as the DNA helix reforms. The chain grows at 30-50 nt / second
- DNA snaps back together as it is being transcribed
- **Diagram on slide**

Proofreading

- RNA poly has two types of proof reading
  - Similar to DNA poly editing, newly inserted nucleotide is removed by reversing synthesis reaction
  - Enzyme moves back one or more nucleotides, cleaves RNA, then resumes synthesis in a forward direction

Termination

- Terminator sequences are used to end transcription. In prokaryotes there are two types
  - Rho-independent or type I terminators have twofold symmetry that would allow a hairpin loop to form. The palindrome is followed by a 4-8U residue in the transcript and when these sequences are transcribed, they cause termination
    - Slow down poly and push it off
    - Bifold symmetry forms hairpin then hits poly U sequence which cause RNA to fall off DNA
    - Hit poly U and UTP concentration decreases and slows down
  - Rho-dependent or type II terminators lack the poly U region, and many also lack the palindrome. The protein is required for termination. It has two domains, one binding RNA and the other binding ATP. ATP hydrolysis provides energy for the protein to move along the transcript and destabilize the RNA-DNA hybrid at the termination region
- **Diagram on Slide**

Transcription in Eukaryotes

- Prokaryotes contain only one RNA poly, which transcribes all RNA for the cell. Eukaryotes have three different poly's, each transcribing a different class of RNA. Processing of transcripts is also more complex in eukaryotes
- Eukaryotic RNA poly's
    - RNA poly I, located in the nucleolus, transcribes the 3 major rRNAs (28S, 18S and 5.8S)
    - RNA poly II, located in the nucleoplasm, transcribes mRNA and some snRNAs
    - RNA poly III, located in the nucleoplasm, transcribes tRNAs, 5S rRNA, and the remaining snRNAs

Eukaryotic RNA Polymerase

- Harder to study then the prokaryotic counterpart, because they are present at low concentrations
- All known eukaryotics RNA polys have multiple subunits. An example is yeast RNA poly II with 12 subunits, 5 of which are also in its RNA poly III
    - Central core is the same but some accessory subunits are different

Transcription of Protein-Coding Genes by RNA Poly II

- When protein coding genes are first transcribed by RNA poly II, the product is a precursor-mRNA (pre mRNA). The pre-mRNA will be modified to produce a mature mRNA
- Promoters for protein coding genes are analyzed in two ways
    - Directed mutation
    - Comparison of sequences from known genes

Eukaryotic Promoter Analysis

- Core promoter elements are located near the transcription start site and specify where transcription begins. Examples include
    - The initiator element (inr), a pyrimidine rich sequence that spans the transcription site
    - The TATA box (aka a TATA element or Goldberg-Hogness box) at -30l its full sequence is TATAAAA. This element aids in local DNA denaturation, and sets the start point for transcription
- Promoter proximal elements are required for high levels of transcription. They are further upstream from the start site, at positions btwn -50 and -200. These elements generally function in either orientation. Examples include:
    - The CAAT box, located at about -75
    - The GC box, consensus sequence GGGCGG, located at about -90

Eukaryotic Promoter Elements

- Various combos of core and proximal elements are found near different genes. Promoter proximal elements are key to gene expression
- Activators, proteins important in transcription regulation, are organized by promoter proximal elements
- Housekeeping (used in all cell types for basic cellular functions) genes have common promoter proximal elements and are recognized by activator proteins found in cells
    - Actin
    - Glucose-6-phosphate dehydrogenase
- Genes expressed only in some cell types or at particular times have promoter proximal elements recognized by activator proteins found only in specific cell types or times

Enhancers

- Another cis-acting element. They are required for maximal transcription of a gene
- Usually upstream of the transcription initiation site, but may also be downstream. They may modulate from a distance of thousand of base pairs away from the initiation site
- Contain short sequence elements, some similar to promoter sequences
- Activators bind these sequences and other protein complexes form, bringing the enhancer complex close to the promoter and increasing transcription
- **IMPORTANT Diagram on slide**

Eukaryotic Transcription Initiation

- Transcription initiation requires assembly of RNA poly II and binding of general transcription factors (GTFs) on the core promoter
    - GTFs are needed for initiation by all 3 RNA poly's
    - GTFs are numbered to match their corresponding RNA poly, and lettered in the order of discovery (e.g. TFIID was the 4th GTF discovered that works with RNA poly II)
- Sequence of binding is not completely understood
    - Binding of GTFs and RNA poly II occurs in a set of order in vitro experiments to produce the complete transcription initiation complex or preinitiation complex (PIC)
    - The situation is less clear in vivo some data indicate that initiation complex forms before binding promoter

    - Transcription for eukaryotes is also complicated by the nucleosome organization of the chromosome
- **IMPORTANT Diagram on Slide**

Eukaryotic mRNAs

- Eukaryotic mRNAs have 3 main parts
    - The 5’ leader sequence, or 5’ untranslated region (5’ UTR), varies in length
    - The coding sequence, which specifies the AA sequence of the protein that will be produced during translation. It varies in length according to the size of the protein that it encodes
    - The trailer sequence, or 3’ untranslated region (3’ UTR), also varies in length and contains information influencing the stability of the mRNA
- Eukaryotes and prokaryotes produce mRNAs somewhat differently
    - Prokaryotes use the RNA transcript as mRNA without modification. Transcription and translation are coupled in the cytoplasm. Messages may be polycistronic
    - Eukaryotes modify pre-mRNA into mRNA by RNA processing. The processed mRNA migrates from nucleus to cytoplasm before translation. Messages are always monocistronic
- **Diagram on Slide**

Prokaryotic and Eukaryotic mRNA

- Prokaryotic
    - No nucleus
    - May move on before poly is done
- Eukaryotic
    - 5’ cap and poly A tail
    - produce introns that will be spliced out
    - mRNA is transported
- **Diagram on Slide**

Eukaryotic 5’ Modifications to mRNA

- The newly made 5’ end of the mRNA is modified by 5’ capping
- A capping enzyme adds a guanine, usually 7-methyl guanosine (m7G), to the 5’ end using an unusual 5’-to-5’ linkage. Sugars of the 2 adjacent nt are also methylated. The cap is used for ribosome binding to the mRNA during translation initiation

- No nucleases that can degrade it
- **Diagram on Slide**

Eukaryotic 3' Modifications to mRNA

- The 3' end of mRNA is marked by a poly(A) tail
- Transcription of mRNA continues through the poly(A) consensus sequence (AAUAAA)
- Proteins bind and cleave RNA. These include
    - CPSF (cleavage and polyadenylation specificity factor)
    - CstF (cleavage stimulation factor)
    - Two cleavage factor proteins (CFI and CFII)
- After cleavage, the enzyme poly (A) polymerase (PAP), adds A nucleotides to the 3' end of the RNA, using ATP as a substrate. PAP is bound to CPSF during this process
- PABII (poly (A) binding protein II) binds the poly (A) tail as it is produced
- Poly (A) sequence not recognized in the gene
- **Diagram on Slide**

Introns

- Eukaryotic pre-RNAs often have introns (intervening sequences) between the exons (expressed sequences) that are removed during RNA processing
- Discovered in 1977 by Richard Roberts, Philip Sharp, and Susan Berger
- Removal of introns is necessary for mRNA maturation, as hnRNA (heterogeneous nuclear RNA) becomes functional mRNA
- In Philip Leder's lab (1978) it was shown that the mouse beta-globin pre-mRNA (part of the cells hnRNA) is collinear with the gene that encodes it, while the mature beta-globin mRNA is shorter then the gene. The missing RNA was an intron that was removed during RNA processing
- Introns are found in most eukaryotic genes that encode proteins, and have also been found in bacteriophage genes

Diagram on slide

Processing of Pre-mRNA to Mature mRNA

- Steps in mRNA processing include
    - Transcription of the gene by RNA poly II
    - Addition of the 5' cap
    - Addition of the poly (A) tail
    - Splicing to remove introns

- **Diagram on slide**

Introns and Spliceosome

- Introns typically begin with 5' –GU and end with AG-3', but mRNA splicing signals involve more then just these two small sequences
- Spliceosomes are small nuclear ribonucleoprotein particles (snRNPs) associated with pre-mRNAs
- Spliceosome principal snRNAs are U1, U2, U4, U5, and U6
    - Each snRNA is associated with several proteins
    - U4 and U6 are part of the same snRNP, others are in their own
    - Each snRNP type is abundance (at east 10^5 copies per nucleus)

Spliceosome

- The steps of splicing are outlined in the figure to the right
    - U1 snRNP binds the 5' splice junction of the intron, as a result of base pairing of the U1 snRNA to the intron RNA
    - U2 snRNP binds the branch point sequence upstream of the 3' splice junction
    - U4/U6 and U5 snRNPs interact, then bind the U1 and U2 snRNPs, creating a loop in the intron
    - U4 snRNP disassociates from the complex, forming the active spliceosome
    - The spliceosome cleaves the intron at the 5' splice junction, freeing it from exon 1. The free 5' end of the intron bonds to a specific nucleotide (usually A) in the branch point sequence to form an RNA lariat
    - The spliceosome cleaves the intron at the 3' junction, liberating the intron lariat. Exons 1 and 2 are ligated and the snRNPs are released
- **Diagram on slide**

Self-Splicing Introns

- The rRNA genes of most species do not contain introns
- Some species of the protozoan Tetrahymena have a 413 bp intron in their 28S rRNA sequence. Tom Cech (1982) showed that splicing of this intron (called a group I intron) is protein independent. The intron self-splices by folding into a secondary structure that catalyzes its own excision
- Group I introns are rare, but examples occur in
    - Nuclear rRNA genes
    - Some mitochondrial mRNA genes

    - Some tRNA and mRNA genes in bacteriophages

<u>Splicing of Group I Introns</u>

- The steps in self-splicing group I intron in Tetrahymena include
    - The pre-rRNA is cleaved at the 5’ splice junction and guanosine is added to the 5’ end of the intron
    - The intron is cleaved at the 3’ splice junction
    - The two exons are joined together
    - The excised intron forms a lariat structure, which is cleaved to produce a circular RNA and a short linear piece of RNA
- Self splicing is not an enzyme activity, because the RNA is not regenerated in its original form at the end of the reaction. However, the discovery of ribozymes (catalytic RNAs) has significantly altered our view of the biochemistry involved in the origin of life

Chapter 13 – Translation and Proteins

<u>Characteristics of the Genetic Code</u>

- It is a triple code. Each 3 nucleotide codon in the mRNA specifies one AA in the polypeptide
- It is comma free. The mRNA is read continuously, three bases at a time, without skipping any bases
- It is nonoverlapping. Each nucleotide is part of only one codon, and is read only once during translation
- **Diagram on Slide**
- It is almost universal. In nearly all organisms studied, most codons have the same AA meaning. Examples of minor code differences include the protozoan Tetrahymena and mitochondria of some organisms
- It is degenerate. Of 20 AA, 18 are encoded by more than one codon. Met (AUG) and Trp (UGG) are the exceptions; all other AA correspond to a set of two or more codons. Codon sets often show a pattern in their sequences; variation at the 3rd position is most common
- The code has a start and stop signal. AUG is the usual start signal for protein synthesis and defines the open reading frame. Stop signals are codons with no corresponding tRNA, the nonsense or chain terminating codons. There are generally 3 stop codons: UAG (amber), UAA (ochre) and UGA (opal)
- **Diagram on Slide**

<u>Translation: The Process of Protein Synthesis</u>

- Ribosomes translate the genetic message of mRNA into protein
- The mRNA is translated 5’-3’, producing a corresponding N-terminal to C-terminal polypeptide
- AA’s bound to tRNAs are inserted in the proper sequence due to:
    - Specific binding of each AA to it’s tRNA
    - Specific base pairing btwn the mRNA codon and tRNA anticodon

<u>The mRNA Codon Recognizes the tRNA Anticodon</u>

- tRNA.Cys normally carries the AA cysteine. Ehrestein, Weisblum, and Benzer attached cysteine to tRNA.Cys in vitro (making Cys-tRNa.Cys) and then chemically altered it to alanine (making Ala-tRNA.Cys)
- When used for in vitro synthesis of hemoglobin the tRNA inserted alanine at sites where cysteine was expected
- They concluded that the specificity of codon recognition lie in the tRNA molecule, and not in the AA it carries
- **Diagram on Next 2 Slides**

Adding an AA to tRNA

- Aminoacyl-tRNA synthetase attaches AA's to their specific tRNA molecules. The charging process (aminoacylation) prodces a charged tRNA (aminoacyl-tRNA), using energy from ATP hydrolysis
- There are 20 different aminoacyl-tRNA synthetase enzymes, one for each AA, and each recognizes the structure of the specific tRNA(s) it charges
- **Diagram on Slide**

Translation

- Protein synthesis is similar in prokaryotes and eukaryotes. However, some significant differences do occur
- In both it is divided into 3 stages
  - Initiation
  - Elongation
  - Termination
- Initiation of translation requires:
  - An mRNA
  - A ribosome
  - A specific initiator tRNA
  - Initiation factors
  - Mg^2+ (magnesium ions)
- Prokaryotes
  - Large subunit – 23S rRNA, 31 proteins, 5S rRNA
  - Small subunit – 16s rRNA, 21 proteins
- Eukaryotes
  - Large subunit – 28S rRNA, 49 proteins, 5S and 5.8S rRNA
  - Small subunit – 18S rRNA, 33 proteins
- **Diagram on Slide**

Initiation of Translation

- Prokaryotic translation begins with binding of the 30S ribosomal subunit to mRNA near the AUG codon at the Shine-Dalgarno sequence. The 30S comes to the mRNA bound to:
  - All three initiation factors, IF1 IF2 and IF3
  - GTP
  - Magnesium ions

- **Diagram on Slide**

Ribosome Binding to mRNA Requires more than the AUG

- RNase protection experiments have shown that the ribosome binds at a ribosome binding site, where it is oriented to the correct reading frame for protein synthesis
- The AUG is clearly identified in these studies
- An additional sequence 8-12 nucleotides upstream from the AUG is commonly involved. Discovered by Shine and Delgarno, these purine rich sequences (e.g. AGGAG) are complementary to the 3' end of the 16S rRNA
- Complementarity between the Shine-Delgarno sequence and the 3' end of 16S rRNA appears to be important in ribosome binding to the mRNA
- At initiation codon fmet binds
- **Diagram on Slide**

Initiation of Translation

- Initiator tRNA binds the AUG to which the 30S subunit is bound. AUG universally encode methionine. Newly made proteins begin with Met, which is often subsequently removed
    - Initiation methionine in prokaryotes is formylmethionine (fMet). It is carried by a specific tRNA (with the anticodon 5'-CAU-3')
- Methionine at sites other then the beginning of the polypeptide are inserted by tRNA.Met (a different tRNA), which is charged by the same aminoacyl-tRNA synthetase as tRNA.fMet
- **Diagram on Slide**

Main Differences in Eukaryotic and Prokaryotic Translation

- Initiator methionine is not modified. As in prokaryotes, it is attached to a special tRNA
- Ribosome binding involves the 5' cap, rather then the shine delgarno sequence
    - Eukaryotic initiator factor (eIf-4F) is a multimer of proteins, including the cap binding protein (CBP). It binds the 5' mRNA cap
    - Then the 40S subunit, complexed with initiator Met-tRNA, several eIFs, and GTP binds the cap complex, along with other eIFs
    - The initiator complex scans the mRNA for a Kozak sequence that includes the AUG start codon. This is usually the first AUG in the transcript
    - When the start codon is located, 40S binds, and then 60S binds, displacing the eIFs and creating an 80S initiation complex with initiator Met-tRNA in the ribosome's P site

- The eukaryotic mRNA's 3' poly (A) tail also interacts with the 5' cap. Poly (A) binding protein (PABP) binds the poly (A) and also binds a protein in eIF-4F on the cap, circularizing the mRNA and stimulating translation

Elongation of the Polypeptide Chain

- Elongation of the AA chain has 3 steps
    - Binding of aminoacyl-tRNA to the ribosome
    - Formation of a peptide bond
    - Translocation of the ribosome to the next codon

Steps in Elongation During Translation

- Second charged tRNA has entered A site, facilitated by EF-Tu; first elongation step commences
- Peptide bond formal uncharged tRNA move to the E site and subsequently out of the ribosome; the mRNA has been translocated 3 bases to the left, causing the tRNA bearing the dipeptide to shift into the P site
- The first elongation step is complete, facilitated by EF-G. Te third charged tRNA is ready to enter the A site
- Third charged tRNA has entered A site, facilitated by EF-Tu; second elongation step begins
- Tripeptide formed; second elongation step completed; uncharged tRNA moves to E site
- Polypeptide chain synthesized and exiting ribosome
- **Diagram on Slides**

Peptide Bond Formation

- The 2 aminoacyl-tRNAs are positioned by the ribosome for peptide bond formation which occurs in 2 steps
    - In the P site, the bond btwn the AA and its tRNA is cleaved
    - Peptidyl transferase forms a peptide bond between the non free AA in the P site and the AA attached to the tRNA in the A site. Experiments indicate that the 23S rRNA is most likely the catalyst for peptide bond formation
- The tRNA in the A site now has the growing polypeptide chain attached to it
- **Diagram on Slide**

Polyribosomes

- In both prokaryotes and eukaryotes, simultaneous translation occurs. New ribosomes may initiate as soon as the previous ribosome has moved away from the initiation site, creating a polyribosome (polysome); an average mRNA might have 8-10 ribosomes

Termination of Translation

- Signaled by a stop codon (UAA, UAG, UGA), which has no corresponding tRNA
- Release factors (RF) assist the ribosome in recognizing the stop codon and terminating translation
- In E. coli
    - RF1 recognizes UAA and UAG
    - RF2 recognizes UAA and UGA
    - RF3 stimulates termination
- In eukaryotes, eRF is the only termination factor, recognizing all 3 stop codons and stimulating termination
- **Diagram on Slide**

Termination

- Peptidyl transferase releases the polypeptide from the tRNA in the ribosomal P site
- The tRNA is released from the ribosome
- The 2 ribosomal subunits and RF dissociate from the mRNA
- **Diagram on Slide**

Chemical Structure of Proteins

- Proteins are built from AA's held together by peptide bonds. Te AA confer shape and properties to the protein
- Two or more polypeptide chains may associate to form a protein complex. Each cell type has characteristic proteins that are associated with its function

Amino Acid Structure

- All AA except for proline have a common structure
- The alpha carbon is bonded to
    - An amino group (NH3), which is usually charged at cellular pH (NH3+)
    - A carboxyl group (-COOH) which is also usually charged at cellular pH (-COO^-)
    - A hydrogen atoms (H)

  - An R group which is different for each AA and confers distinctive properties. The R group in an AA chain give polypeptides their structural and functional properties
- **Diagram on Slide**

<u>Amino Acids</u>

- There are 20 AA used in biological porteins. They are divided into subgroups according to the properties of their R groups (acidic, basic, neutral and polar, or neutral and nonpolar)
- **Diagram on Slide**

<u>Peptide Bond</u>

- Polypeptides are chains of AA's joined by covalent peptide bonds. A peptide bond forms btwn the carboxyl group of one AA and the amino group of another
- Polypeptides are unbranched, and have a free amino group at one end (the N terminus) and a carbonyl group at the other (the C terminus). The N-terminal end defines the beginning of the polypeptide
- N terminus is usually thymine
- Linear structure
- **Diagram on Slide**

<u>Molecular Structure of Proteins</u>

- Proteins have up to 4 levels of organization
  - Primary structure is the AA sequence of the polypeptide. This is determined by the nucleotide sequence of the corresponding gene
  - Secondary structure is folding and twisting of regions within a polypeptide, resulting from electrostatic attractions (with R groups) and or hydrogen bonding. Common examples are alpha helix and beta pleated sheet
  - Tertiary structure is the 3D shape of a single polypeptide chain, often called its conformation. Arises from interactions btwn R groups on the AA of the polypeptide, and thus relates to primary structure
  - Quaternary structure occurs in multiple subunit proteins, as a result of the association of polypeptide chains. Hemoglobin is an example with two 141 AA alpha polypeptides and 2 146 AA beta polypeptides (each associated with a heme group)

Protein Folding: Chaperones

- More then AA sequence alone determines the folding of a polypeptide into a functional protein. Cell biology experiments show that proteins in the molecular chaperone family assist other proteins in folding
    - Initially discovered in Drosophila as heat-shock proteins
- Misfolded proteins are tagged with ubiquitin proteins which target the misfolded proteins for degradation by the proteasome
- **Diagram on Slide**

Prion Diseases and Protein Misfolding

- Bovine spongiofrom encephalopathy (cows), scrapie (sheep) and Creutzfeldt-Jakob disease in humans
    - Normal cellular version (PrP^c) is synthesized in neurons and found in brain tissue of all adult animals
        - Folds into an alpha helix
    - Above diseases caused by Prions, which are aggregates of misfolded protein called PrP^sc
        - Folds into beta pleated sheet
            - Misfolded version creates a chain reaction causing properly folded PrP^c to misfold as well

Gene Control of Enzyme Structure

- These discoveries are the foundation of modern molecular genetics
    - Genes encode proteins, including enzymes
    - Genes work in sets to accomplish biochemical pathways
    - Genes often work in cooperation with other genes
    - Genes also encode for RNAs (tRNA and rRNA)

Inborn Errors of Metabolism

- Alkaptonuria is a human trait characterized by urine blackening on exposure to air and arthritis in later life
- Archibald Garrod and Willion Bateson (1902) concluded that it is genetically determined because
    - Families with it often have several affected memebers
    - Is much more common in first cousin marriages then marriages with unrelated partners

- Garrod showed that it results from homogentisic acid (HA) in the urine. HA is absent from normal urine. Garrod reasoned that normal people metabolize HA, but those with it do not because they lack the necessary enzyme. He termed this an inborn error of metabolism
- Responsible mutation is recessive. The gene was later shown to be on chromosome 3
- Garrod's work was the first evidence of a specific relationship btwn genes and enzymes. With the insight that a mutation can block a human metabolic pathway by changing an enzyme, causing a detectable buildup of that enzymes substrate, he found a similar relationship in 3 other human diseases
- **Diagram on Slide**

PKU

- Commonly caused by a mutation on chromo 12 in the phenylalanine hydrolase gene, preventing the conversion of phenylalanine into tyrosine
- Phenylalanine is an essential AA, but in excess it is harmful, and so it is normally converted to tyrosine. Excess phenylalanine affects the CNS, causing mental retardation, slow growth, and early death
- PKU effect is pleiotropic. Some symptoms result from excess phenylalanine. Others result form inability to make tyrosine; these influence fair skin and blue eyes (even with brown eye gene) and low adrenaline levels
- **Diagram on Slide**

Living with PKU

- Diet is used to manage PKU by providing just enough phenylalanine for protein synthesis, but not enough that accumulates. TO be effective, the special diet must commence in the first 2 months after birth, continues at least throughout childhood, and is resumed before pregnancy in PKU women to avoid phenylalanine level that would affect the fetus
- Nutrasweet is aspartame, which breaks down to aspartic acid and phenalanine, with serious consequences for PKU

PKU Screening

- All US and Canadian newborns are screened for PKU using the Guthrie test
    - A drop of blood on filter paper is placed on a solid media containing beta-2-thienylalanine and the bacterium Bacillus subtilis
    - Normally beta-2-thienylalaine inhibits growth of Bacillus subtilis

  - Phenylalanine allows Bacillus sutyilis to grow in the presence of beta-2-thienylalanine, so bacterial growth indicates high phenylalanine levels in the blood and the possibility that the infant has PKU

Albinism

- Classic albinism results from an autosomal recessive mutation in the gene for tyrosinase. Tyrosinase is used to convert tyrosine to DOPA in the melanin pathway. Without melanin, individuals have white skin and hair, and red eyes due to lack of pigmentation in the iris
- Two other forms of albinism are known, resulting from defects in other genes in the melanin pathway. A cross btwn parents with different forms of albinism can produce normal children
- 2 homo parents
  - all progeny normal, due to complementation
  - due to parents having mutation in 2 different pathways that affect melanin

One Gene – One Enzyme Hypothesis

- genes act by regulating definite chemical events
- George Beadle and Edward Tatum (1942) showed a direct relationship btwn genes and enzymes in haploid fungus Neurospora crassa. This lead to their one gene- one enzyme hypothesis, and a share of the 1958 Nobel Prize in Physiology or Medicine

Life Cycle of N. crassa (Orange Bread Mold)

- Propagates asexually by dispersal of bits of mycelium or by conidia
- It also propagates sexually by means of two mating types a and alpha
- Wild type needs only simple minimal media with
  - Inorganic salts (including nitrogen source)
  - An organic carbon source (such as glucose or sucrose)
  - Biotin (a vitamin)
- To grow on minimal media, wild type synthesizes all organic molecules it needs for growth (prototrophic). An auxotrophic mutant unable to make a needed nutrient will grow only if that nutrient is provided as a supplement in its medium
- **Diagram on Slide**

Beadle and Tatum's Experiments

- Auxotrophic mutants were created by mutating with x-rays and then crossing with a wild type strain. The cross ensured that effects were due to inheritance, rather then due to damage from radiation. In their experiment
    - One progeny spore per ascus was germinated in a complete medium so that growth would occur regardless of nutritional mutations. The cells were transferred to minimal media, where auxotrophs wont grow.
    - Each mutant was then tested on an array of minimal media, each with a different single supplement to determine the type of nutritional mutation
    - In this example it was a mutation in the AA pathway
        - Pathway was for tryptophan bc it grows in that one and not the others when supplemented with the AA
- **Diagram on Slide**

Many Genes Interact in Cells

- Beadle and Tatum reasoned that metabolism proceeded by a series of reactions, each catalyzed by an enzyme, and organized into pathways. Their analysis of methionine biosynthesis is an example
    - They started with a set of methionine auxotrophs and found that 4 genes were involved, met-2+, met-3+, met-5+, and met-8+
    - They checked each mutant on a series of minimal media, each supplemented with a different chemical believed to be involved in the pathway. They expected growth if providing a chemical used after the metabolic block, so the earlier mutant gene functions in the pathways, more supplements will support growth
- They were able to deduce the pathway of methionine synthesis, and to correlate mutations with enzymes used in the pathway
- **Diagram on Slide**

Beadle and Tatum's Conclusion

- Famous conclusion from this experiment is that one gene codes one enzyme
- Later work showed that some proteins consists of more then one polypeptide, and that not all proteins are enzymes. The principle is now usually stated as one gene – one polypeptide
- End product is required for growth
- Only the product is in the cell not the substrates (they can be toxic)

Gene Control and Protein Structure

- Genes also make proteins that are not enzymes. Structural proteins, such as hemoglobin, are often abundant, making them easier to isolate and purify

Sickle-Cell Anemia

- J. Herrick (1910) first described sickle-cell anemia, finding that RBC's change shape (form a sickle) under low oxygen tension
    - Sickled RBC's are fragile, hence anemia
    - They are less flexible than normal RBC's, and form blocks in capillaries, resulting in tissue damage downstream
    - Effects are pleiotropic, including damage to extremities, heart, lungs, brain, kidneys, GI tract, muscles, and joints. Results include heart failure, pneumonia, paralysis kidney failure, abdominal pain, and rheumatism
    - Heterozygous individuals have sickle-cell trait although it is a much milder form of the disease
- E.A Beet and J.V. Neel independently proposed (1949) that sickle-cell trait and disease were the result of a single mutant allele
- Lineus Pauling and coworkers (1949) used electrophoresis and showed
    - Hemoglobin from individuals with sickle-cell anemia (Hb^s) has altered mobility compared with normal hemoglobin (Hb^a)
    - Hemoglobin from individuals with sickle-cell trait show equal amounts of Hb-A and Hb-S indicating that hetero make both forms of hemoglobin
    - Therefore, the sickle-cell mutation changes the form of its corresponding protein, and protein structure is controlled by genes
- **Diagram on Slide**
- Proteins – charged (-,+, neutral), not all the same
    - Due to R-groups

Hemoglobin

- Formed by 4 polypeptide chains, 2 molecules of the alpha polypeptide and 2 of the beta polypeptide, each associated with a heme group
- V.M. Ingram (1956) found that the 6th AA of the beta chain in sickle-cell hemoglobin is valine (no electrical charge) rather then the negatively charged glutamic acid in the beta chain of normal hemoglobin (Glu6Val)
- **Diagram on Slide**

Outline of Sickle-Cell Anemia Trait

- Wild type beta chain allele is beta-A, which is codominant with mutant sickle chain allele beta-S
- Hemoglobin of beta-A/beta-A individuals has normal beta subunits, while hemoglobin of those with the genotype beta-S/beta-S has beta subunits that sickle at low oxygen tension
- Hemoglobin of beta-A/beta-S individuals is ½ normal and ½ sickling form. (The 2 beta chains of an individual hemoglobin molecule will be of the same type, rather than mixed). These heterozygotes may experience sickle-cell symptoms after a sharp drop in the oxygen content of their environment

Other Hemoglobin Mutants

- Screening of hemoglobin for altered electrophoretic mobility has identified over 200 hemoglobin mutants, showing a variety of AA substitutions in both the alpha and beta chains. Each appears to derive from a single AA change
- Most effects are not as severe as those seen in sickle-cell anemia
- **Diagram on Slide**

**** Sickle-cell Allele Confers Resistance to Malaria****

Chapter 14 – Gene Mutation, Transposition & DNA Repair13-04-14 5:22 PM

Defining Mutations

- A mutation is a change in a DNA base pair or a chromosome
    - Somatic mutations affect only the individual in which they arise
    - Germ line mutations alter gametes, affecting the next generation
- Mutations are quantified in 2 different ways:
    - Mutation rate is the probability of a particular kind of mutation as a function of time (e.g. number per gene per generation)
    - Mutation frequency is the number of times a particular mutation occurs in proportion to the number of cells or individuals in a population (e.g. number per 100 000 organisms)

Types of Point Mutations

- There are 2 general categories of point mutations: base pair substitutions and base pair deletions or insertions
- A base pair substitution replaces one base pair with another, there are 2 types
    - Transitions convert a purine-pyrimidine pair (e.g. AT to GC or TA to CG)
    - Transversion convert a purine-pyrimidine pair to a pyrimidine-purine pair (e.g. AT to TA, or AT to GC)
- **Diagram on Slide**

Base Pair Substitutions and their Effects

- Base pair substitutions in ORF's are also defined by their effect on the protein sequence. Effects vary from none to severe
    - Missense mutations have a base pair change resulting in a different mRNA codon, and therefore a different AA in the protein
    - Nonsense mutations change a codon in the ORF to a stop codon (nonsense) resulting in premature termination of translation and a truncated (often nonfunctional) protein
- **Diagram on Slide**

Do All Mutations Lead to Changes in the Phenotype

- Phenotypic effects may or may not occur, depending on the specific AA change
    - Neutral mutations change a codon in the ORF, but the resulting AA substitution produces no detectable change in the function of the protein (e.g. AAA to AGA substitutes arginine for lysine. The AA have similar properties so the proteins function may not be altered

    - Silent mutations occur when the mutant codon encodes for the same AA as the wild type gene, so that no change occurs in the protein produced (e.g. AAA and AAG both encode for lysine, so this transition would be silent)
- **Diagram on Slide**

Frameshift Mutations

- Deletions and insertions can change the reading frame of the mRNA downstream of the mutation, resulting in a frameshift mutation
    - When the reading frame is shifted, incorrect AA are usually incorporated
    - Frameshifts may bring stop codons into the reading frame, creating a shortened protein
    - Frameshifts may also result in read-through of stop codons, resulting in a longer protein
- Frameshift mutations result from insertions or deletions when the number of affected base pairs is not divisible by 3
- **Diagram on Slide**

Reverse Mutations and Suppressor Mutations

- Point mutations are divided into 2 classes based on their effect on the phenotype
    - Forward mutations change the genotype from wild type to mutant
    - Reverse mutations (reversion or back mutations) change the genotype from mutant to wild type or partially wild type
        - A reversion to the wild type AA in the affected protein is a true reversion
        - A reversion to some other AA that fully or partially restores protein function is a partial reversion

Review

- Mutations
    - Point mutations
        - Transitions (purine to purine or pyrimidine to pyrimidine)
        - Transversions (purine to pyrimidine or pyrimidine to purine)
            - Missense – different AA
            - Nonsense – stop codons
            - Neutral – similar AA, no change in protein function
            - Silent – same AA
            - Frameshift – insertion or deletion alters reading frame

        - Suppressors
            - Intragenic or intergenic

Spontaneous Mutations

- Most mutations are spontaneous, rather then induced by a mutagen
- All types of point mutations can occur spontaneously, during S, G1, and G2 phases of the cell cycle, or by the movement of transposons
- The spontaneous mutation rate in eukaryotes is btwn 10^-4 and 10^-6 per gene per generation, and in bacteria and pages btwn 10^-5 and 10^-7 per gene per generation
- Genetic constitution of the organism affects its mutation rate
    - In Drosophila, males and females of the same strain have similar mutation rates
    - Flies of different strains, however, may have different mutation rates
    - Even certain genes within the same organism may have a different mutation rate than other genes along the same chromosome
- Many spontaneous errors are corrected by the cellular repair systems, and so do not become fixed DNA

DNA Replication Errors

- DNA replication errors can be either point mutations, or small insertions or deletions
- Base pair substitution mutations can result from “wobble” pairing. Bases are normally in the keto form, but sometimes undergo tautomeric shift to form the abnormal (enol) form. The enol form can hydrogen bond with an incorrect partner due to a different spatial positioning of the atoms involved in H-bonding
- **Diagram on Slide**

Production of Mismatch Caused by Wobble Base Pairing

- An example is a GC to AT transition
    - During DNA replication, the enol form of the G could wobble pair with a normal T, producing a GT pair
    - In the next round of replication, G is likely to be back in its keto form, and so both G and A are likely to pair normally, producing one progeny DNA with a GC pair, and another with an AT pair
    - GT pairs are targets for correction by proof reading during replication, and by other repair systems. Only mismatches uncorrected before the next round of replication lead to mutations
- **Diagram on Slide**

DNA Replication Errors

- Additions and deletions can occur spontaneously during replication
    - DNA loops out from the template strand, generally in a run of the same base
    - DNA poly skips looped out bases, creating a deletion mutation
- If DNA poly adds untemplated base(s), new DNA looping occurs, resulting in additional mutations
- Insertions and deletions in structural genes generate frameshift mutations (if they are not multiples of 3)
- **Diagram on slide**

Spontaneous Chemical Changes

- Spontaneous chemical changes include depurination and deamination of particular bases, creating lesions in the DNA
- Depurination removes the purine (A or G) from DNA by breaking the bond with its deoxyribose in the backbone
    - Is common
    - If not repaired before the next round of replication, DNA poly may stall or dissociate
- **Diagram on slide**

Deamination

- Removes an AA group from a base (e.g. cytosine to uracil)
- Uracil is an abnormal base in DNA, and it will usually be repaired
- If uracil is not replaced, it will pair with an A during replication, resulting in a CG-to-TA transition
- Both prokaryotic and eukaryotic DNA have small amounts of 5-methylcytosine (5^mC) in place of the normal cytosine
    - Deamination of the 5mC produces T
    - T is a normal nucleotide in DNA, so it is not detected by repair mechanisms
    - Deamination of 5^mC resulting in CG-to-TA transitions
    - Location of 5^mC in the chromosome are often detected as mutational hot spots
- **Diagram on Slide**

Induced Mutations: X-Rays

- Radiation induces mutations

- X-rays are an example of ionizing radiation, which penetrates tissue and collides with molecules, knocking electrons out of orbits and creating free radicals
    - Ions can break covalent bonds, including those in the DNA sugar-phosphate backbone
    - Ionizing radiation is the leading cause of human gross chromosomal mutations
    - Ionizing radiation kills cells at high doses, and lower doses produce point mutations
    - Ionizing radiation has a cumulative effect. A particular dose of radiation results in the same number of mutations whether it is received over a short or a long period of time

Induced Mutations: UV Light

- UV light causes photochemical changes in the DNA
- UV is not energetic enough to induce ionization
- UV has lower-energy wavelengths than do X-rays, and so has limited penetrating power
- However, UV light in the 254-260 nm range is strongly absorbed by purine and pyrimidine's, forming abnormal chemical bonds
    - A common effect is dimer formation btwn adjacent pyrimidine's, commonly thymine's (designated T^T)
    - C^C, C^T, and T^C dimers also occur, but at a lower frequency. Any pyrimidine dimer can cause problems during DNA replication
    - Most pyrimidine dimers are repaired, because they produce a bulge in the DNA helix. If enough are unrepaired, cell death may result
- **Diagram on slide**

Chemical Mutagens: Base Analogs

- May be naturally occurring or synthetic. They form different groups based on their mechanisms of action
- Base analogs depend upon replication, which incorporates a base with alternate states (tautomers) that allow it to base-pair in alternate ways, depending on its state
    - Analogs are similar to normal nitrogen bases, and so are incorporated into DNA readily
    - Once in the DNA, a shift in the analog's form will cause incorrect base-pairing during replication, leading to mutation

Base Analog: 5BU

- 5BU is an example, it has a bromine residue instead of the methyl group of thymine
- Normally 5BU resembles thymine, incorporated into DNA during replication

- In its rare state, 5BU pairs only with guanine, resulting in a TA-to-GC transition mutation
- If 5BU is incorporated in its rare form, the switch to its normal state resulting in a CG-to-TA transition. Thus, 5BU induced mutations may be reverted by another exposure to 5BU
- Only those base analogs that cause base pair changes are mutagents (e.g. AZT is a stable analog that does not shift)
- **Diagram on slide**

<u>Chemical Mutagens: Base Modifying Agents</u>

- Base modifying agents can induce mutations at any stage of the cell cycle. They work by modifying the chemical structure and properties of the bases. There are 3 types.
- Deaminating agents remove amino groups. An example is nitrous acid, which deaminate G, C, and A
    - Deaminates guanine to produce xanthine, which has the same base pairing as G. No mutations result
    - Delaminates cytosine to produce uracil, which produces a CG-to-TA transition
    - Delaminates adenine to produce hypoxanthine, which pairs with cytosine, causing AT-to-GC transition
    - Mutations induced by it can be reverted with a second treatment
- **Diagram on slide**

<u>Chemical Mutagens: Hydroxylating Agents</u>

- Include hydroxylamine
- Specifically modifies C with a hydroxyl group (OH), so that it pairs only with A instead of G
- Produces only CG-to-TA transitions, and so revertants do not occur with a second treatment
- Mutants, however, can be reverted by agents that do cause TA-to-Cg transitions (e.g. 5BU and nitrous acid)
- **Diagram on slide**

<u>Chemical Mutagens: Alkylating Agents</u>

- Are a diverse group that add alkyl groups to bases. Usually alkylation occurs at the 6-oxygen of G, producing O6^-alkylguanine
- An example is methylmethane sulfonate which methylate's G to produce this
    - Pairs with T rather then C, causing GC-to-AT transitions
- **Diagram on Slide**

Chemical Mutagens: Intercalating Agents

- Insert themselves btwn adjacent bases in dsDNA. They are generally thin, plate like hydrophobic molecules
- At replication, a template that contains an intercalated agent will cause insertion of a random extra base
- The base-pair addition is complete after another round of replication, during which the intercalating agent is lost
- If an alternating agent inserts into new DNA in place of a normal base, the next round of replication will result in a deletion mutation
- Point deletions and insertions in ORFs result in frameshift mutations. These mutations show reversion with a second treatment
- **Diagram on slide**

Chemical Mutagens in the Environment

- A wide variety of chemicals exist in our environment, and many can have mutagenic effects that can lead to genetic diseases and cancer
    - Drugs, e.g. Thalidomide (teratogen)
    - Cosmetics, e.g. acrylamide and ethylene oxide
    - Food additives, e.g. dyes and preservatives
    - Pesticides, e.g. DDT
    - Industrial compounds , e.g. PCBs and dioxins
    - Chemical warfare agents such as mustard gas

Blocking the Use of Thalidomide in the US

- A Canadian born pharmacologist by the name of Frances Oldham Kelsey, while working with the FDA, blocked the use of thalidomide in the US
- Under extreme pressure from Merrel pharmaceuticals, she persevered and insisted that proper testing should be performed
- She received the Distinguished Federal Civilian Service Award, given to her by JFK in 1962
    - In Germany, where the drug was developed approx.. 3500 babies were affected, 1000 in the UK and 115 in Canada, and a small number in the US

How to Determine if a Substance is a Mutagen: The Ames Test

- The Ames test is a simple and inexpensive screen for potential carcinogens. It assays the reversion rate of mutant strains of Salmonella typhimurium back to wt
    - Note: all carcinogens are mutagens, but not all mutagens are carcinogens

- The Ames test is used routinely to screen industrial and agricultural chemicals, and shows a good, but not perfect, correlation btwn mutagens and carcinogens
- The test can be made quantitative to produce a dose-response curve, allowing comparison of relative mutagenicity of different chemicals

The Ames Test

- Histidine (his) auxotrophs are tested for reversion in the presence of the chemical by plating on media lacking the AA histidine
  - Different his tester strains are available to test for base-substitution and frameshift mutations
- Liver enzymes (the S9 extract) are mixed with the test chemical to determine whether the liver's detoxification pathways convert it to a mutagenic form
- More revertants in the region of the test chemical than in the untreated control indicate that it may be a mutagen, and further tests are indicated
- **Diagram on slide**

Detecting Mutations

- Generally, mutation in haploid organisms is readily detected, while recessive mutations in diploid organisms are more difficult to characterize. The problem is compounded in humans, where controlled crosses cannot be done
- For some organisms, especially microorganisms, selection and screening procedures exist
- Visible mutations – some mutations affect the appearance of an organism (e.g. Drosophila eye or wing shape, coat color in animals, colony size in yeast, plaque morphology of phages)
- Nutritional mutations
  - Autotrophic mutants are easily detected for microorganisms that normally can grow on minimal medium, using methods that have been developed for selection and screening
  - An example is replica plating. Cells are first grown on supplemented medium, and then the colonies transferred to minimal medium, as well as to a control plate of supplemented medium. Colonies that grow on supplemented, but not minimal, media are selected for further study
- **Diagram on slide**
- Conditional mutations
  - Mutations in some genes (e.g. DNA and RNA poly) are easily lethal, so these genes are studied by isolating conditional mutations

  - Heat sensitivity is a common conditional mutations, in which a normal protein is produced at permissive temperature, and a nonfunctional protein results at the non-permissive temperature. Screening is generally done by replica plating and incubation at different temperatures
- Resistance mutations
  - Microorganisms such as E. coli and yeast are easily screened for resistance to viruses, chemical, or drugs, because resistant cells will grow when wt cells will not

Repair of DNA Damage

- Both prokaryotes and eukaryotes have enzyme based DNA repair systems that prevent mutations and even death from DNA damage
- Repair systems are grouped by their repair mechanisms. Some directly reverse damage, while others excise the damaged area and then repair the gap

Direct Reversal of DNA Damage

- DNA poly proofreading corrects most of the incorrect nucleotide insertions that occur during DNA synthesis, which stalls until the wrong nucleotide is replaced with a correct one
  - The role of 3'-5' exonuclease activity is illustrated by mutator mutations in E. coli, which confer a much higher mutations rate on the cells that carry them
  - The mutD gene, encoding the e subunit of DNA poly III, is an example. Cells mutant in mutD are defective in proofreading
- UV-induced pyrimidine dimers are repaired using photoreactivation (light repair)
  - Near UN light (320-370nm) activates photolyase (product of the phr gene) to split the dimer
  - Photolysases are found in prokaryotes and some simple eukaryotes, but not in humans
- Damage by alkylation (usually methyl or ethyl groups) can be removed by specific DNA repair enzymes
  - For example O^6-methylguanine methyltransferase (from the ada gene) recognizes it in DNA and removes the methyl group, restoring the base to its original form
  - A similar system repairs alkylated thymine
- Note: mutations of repair enzyme genes increase the organisms rate or spontaneous mutations

Base Excision Repair

- Uses a repair glycosylase enzyme to recognize and remove damaged bases
  - Bond btwn base and deoxyribose is cleaved by the glycosylase

  - Other enzymes cleave sugar-phosphate from backbone, leaving a gap in the DNA
  - Repair DNA poly and DNA ligase use the opposite strand as template to fill in the gap
- **Diagram on slide**

Repair Involving Excisions of Nucleotides

- A repair system that does not require light was discovered in 1964. It is called dark repair, the excision repair system, or, more recently, the nucleotide excision repair (NER) system
  - In E. coli NER corrects pyrimidine dimers and other damage induced distortions of the DNA helix
  - The proteins required are UvrA, UvrB, UvrC, and UvrD (encoded by genes of the same name)
  - A complex of two UvrA and one UvrB proteins slides along the DNA, Wen it encounters a helix distortion, the UvrA subunits dissociate, and a UvrC bind the UvrB at the lesion
  - When UvrBC forms, the UvrC cuts 4-5 nucleotides from the lesion on the 3' side, and 8 nucleotides away on the 5' side. Then UvrB is released and UvrD binds the 5' cut end
- UvrD is a helicase that unwinds the region btwn the cuts, releasing the short ssDNA, while the DNA poly I fills that gap and DNA ligase seals that backbone
- In yeast and mammalian systems, about 12 genes encode proteins involved in excision repair
- A large portion is removed
- Pull helix up and causes a kink that is recognized as the place that needs fixing
- **Diagram on Slide**

Methyl-Directed Mismatch Repair

- Methyl-directed mismatch repair recognizes mismatched base pairs, excises the incorrect bases and then carries out repair synthesis
- In E. coli, initial stages involve products of the mutS, mutL, and mutH genes
  - MutS binds the mismatch
  - MutL and MutH bind unmethylated GATC sequences (site of methylation in E. coli) and bring the GATC close to the mismatch by binding MutS
  - MutH then nicks the unmethylated GATC site, the mistmatch is removed by an exonuclease and the gap is repaired by DNA poly III and ligase
- Only works right after DNA replication
  - Inserted wrong base

    - No proofreading occurred before
- Old will be methylated (on outside)
- New will not be methylated
    - This is how you tell which base is right or wrong
- **Diagram on slide – focus on how it works**

Translesion DNA Synthesis

- Translesion DNA synthesis is used to allow a cell to survive when specific base pairing cannot occur. Survival is usually at the cost of incurring new mutations
- Bacteria use a system called SOS. In E. coli SOS is controlled by 2 genes, lexA and recA. Mutants in either of these genes have their SOS response permanently turned on
    - When no DNA damage is present. LexA represses transcription of about 17 genes with products involved in various types of DNA repair
    - Sufficient DNA damage activates the RecA protein, which stimulates LexA to autocleave, removing repression of the DNA repair gene
    - After damage is repaired, RecA is inactivated, and newly synthesized LexA again represses the DNA repair genes
- The SOS system is mutagenic bypass synthesis system
    - DNA poly for translesion synthesis is made in SOS response. It replicates over and past the lesion
    - Nucleotides are incorporated at the lesion that may not match wt, leading mutations
    - This is better than the alternative, death due to incompletely replicated DNA

Human Genetic Diseases Resulting from DNA Replication and Repair Mutations

- Many human genetic disorders result from gene mutation
- Some human genetic diseases result from defects in DNA replication or repair. Xeroderma pigmentosum is an example
    - The disorder occurs in homozygotes for a mutation in a repair gene
    - Affected individuals are photosensitive, and portions of skin exposed to light show intense pigmentation and warty growths that may become malignant
    - The defect is in nucleotide excision repair, and the inability to repair radiation damage to DNA often results in malignancies

Werner Syndrome

- Rapid aging after puberty
- Defect in a human DNA helicase (likely involved in DNA repair)

- Mutations accumulate leading to apoptosis of the cells

Mutations in ABO Blood Groups

- Sequencing of the human genome has given us insight into how the ABO blood group arose
    - The I^A and I^B alleles encode glycosyltransferases that differe in 4 nucleotides. These changes lead to 2 forms of the glycosyltransferase
    - Type O individuals have a mutation early in the glycosyltransferase gene causing a frameshift mutation. No functional product produced

General Features of Transposable Elements

- Transposable elements are divided into 2 classes on the basis of their mechanism for movement
    - Some encode proteins that move the DNA directly to a new position or replicate the DNA to produce a new element that integrates elsewhere. This type is found in both prokaryotes and eukaryotes (insertion sequences and transposons)
    - Others are related to retroviruses, and encode reverse transcriptase for making DNA copies of their RNA transcripts, which then integrate at new sites. This type is found only in eukaryotes (retroposons)

General Features of Transposons

- Transposition is nonhomologous recombination, with insertion into DNA that has no sequence homology with the transposon
    - In prokaryotes, transposition can be into the cells chromosome, a plasmid, or a phage chromosome
    - In eukaryotes, insertion can be into the same or a different chromosome
- Transposable elements can cause genetic changes, and have been involved in the evolution of both prokaryotic and eukaryotic genomes
- Transposons may
    - Insert into genes
    - Increase or decrease gene expression by insertion into regulatory sequences
    - Produce chromosomal mutations through the mechanics of transposition
- Transposition frequency is low, as a high rate is likely to kill the cell

Transposable Elements in Prokaryotes

- Insertion sequence (IS) elements
- Transposons (Tn)

- **Diagram on slide**

Insertion Sequences

- IS elements are the simplest transposable elements found in prokaryotes, encoding only genes for mobilization and insertion of its DNA. IS elements are commonly found in bacterial chromsomes and plasmids
- IS elements were first identified in E. coli's galactose operon, where some mutations were shown to result from insertion of a DNA sequence now called IS1
- Prokaryotic IS elements range in size from 768bp to over 5 thousand bp. Known E. coli elements include IS1 which is 768 bp long and present in 4-19 copies on the E. coli chromosome, IS2 and IS1OR
- The ends of all sequenced IS elements show inverted terminal repeats (IRs) of 9-14 bp. IS1 has 23 bp of nearly identical sequences
- **Diagram on slide**

How an IS Element Transposes

- Transposition requires transposase, an enzyme encoded by the IS element
- Transposases recognizes the IR sequence to initiate transposition of a frequency characteristic for each IS element (range is 10^5 to 10^7 per generation)
- IS elements insert into the chromosome without sequence homology (illegitimate recombination) at target sites
    - A staggered cut is made in the target sidte, and the IS element is inserted
    - DNA poly and ligase fill in gaps, producing small direct repeats of the target site flanking the IS element (target site duplication)
- Makes a staggered patter
- Acts like a dimer
- Makes small repeat in DNA
- Makes a foot print due to repeat
- **Diagram on slide**

Composite Transposons

- Transposons are similar to IS elements, but carry additional genes, and have a more complex structure. There are 2 types of prokaryotic transposons
- Composite transposons carry genes (e.g. antibiotic resistance) flanked on both sides by IS elements (IS modules)
    - The IS elements are of the same type, and called ISL (left) and ISR (right)

  - ISL and ISR may be direct or inverted orientation to each other
  - Transposition of composite transposons restuls from the IS elements, which supply transposase and its recognition signals, the IRs
- Tn10's transposition is rare (10^-7/molecule/generation)
- Transposons, like IS elements, produce target site duplications (e.g. 9 bp duplication for Tn10)
- **Diagram on slide**

Noncomposite Transposons

- Also carry genes (e.g. drug resistance) but do not terminate with IS elements
  - Transposition proteins are encoded in the central region
  - The ends are repeated sequences (but not IS elements)
  - Noncomposite transposons cause target site duplications (like composite transposons)
  - An example is Tn3, which produces a 5 bp duplication upon insertion
- **Diagram on slide**

Model for Transposition

- Cointegration is an example of the replicative transposition that occurs with Tn3 and its relatives
  - Donor DNA containing the Tn fuses with recipient DNA
  - The Tn is duplicated with one copy at each donor-recipient DNA junction, producing a cointegrate
  - The coinegrate is resolved into 2 products, each with one copy of the Tn
- Conservative (nonreplicative) transposition is used by Tn10, for example. The Tn is lost from its original position when it transposes
- **Diagram on slide**

Mutations Induced by Insertion Sequences and Transposons

- Insertion into a gene disrupts it
- Gene expression is changed by adjacent promoters within the IS or Tn
- Deletions and insertions occur
- Crossing-over results from duplicated IS and Tn sequences in the genome

Transposable Elements in Eukaryotes

- Rhodes (1930's), working with sweet corn, observed interactions btwn two genes

  - A gene for purple seed color, the A1 locus. Homozygous mutants (a/a) have colorless seeds
  - A gene on a different chromosome, Dt (dotted) that causes seeds with genotype a/a Dt/- to have purple dotes
- Dt appears to mutate the a allele back to the A1 wt in regions of the seed, producing a dotted phenotype
- The effect of the Dt allele is dose dependent
  - One dose gave an average of 7.2 dots per seed
  - Two doses gave an average of 22.2 dots per seed
  - Three doses gave an average of 121.9 dots per seed
- Concluded that Dt was a mutator gene
- McClintock (40's-50's), working with corn, proposed the existence of "controlling elements" that regulate other genes and are mobile in the genome
- The genes studied by both Rhoades and McClntock have turned out to be transposable elements, and many others have been identified in various eukaryotes
- The most studied are transposons of yeast, drosophila, corn, and humans
- Their structure is very similar to that of prokaryotic transposable elements
- Eukaryotic transposable elements have genes for transposition and integration at a number of sites, as well as a variety of other genes
- Random insertion results from nonhomologous recombination and means that any chromosomal gene may be regulated by a transposon

Effects of Transposition

- The result of transposon insertion into a chromosome will depend on the properties of the transposon, with possible effects including
  - Chromosome mutations such as duplication, deletions, inversions, translocations, or breakage
  - Disruption of genes to produce a null mutation (gene is nonfunctional)
  - Activation or repression of adjacent genes by disrupting a cellular promoter, or by action of transposon promoters

Chapter 16 – Cell Cycle Regulation and Cancer

13-04-14 5:22 PM

Cancer is a Genetic Disease

- Single nucleotide substitutions to large scale rearrangements of the genome are associated with cancer
- Unlike other diseases, cancer is attributed to mutations that arise predominantly in somatic cells
    - Rarely the result of a single mutation, rather an accumulation of mutations (sometimes 6-10)
- **Diagram on slide**

What is Cancer?

- Abnormal cell growth (cell proliferation)
    - Benign tumor (operable)
- Abnormal restraints that keep cells from spreading and invading other parts of the body (metastasis)
    - Malignant tumor (not operable)
- Cancer cells have more metabolic activity so they give off more heat in a heat scan
- **Diagram on slide**

Clonal Origin of Cancer Cells

- Malignant tumor
    - Billions of cells
    - Originated from one common ancestral cell
        - Important for its causes, diagnosis and prognosis
- Burkitt's Lymphoma
    - Reciprocal translocations btwn chromosome 8 (breakpoint near, c-myc gene) and chromosomes 2,14, or 22 near immunoglobulin genes
        - In every patient, all the lymphoma cells have the same breakpoints
        - Each patient will have their own unique breakpoint

Cancer is a Multistep Process

- Cancer is a genetic disease caused by mutations
    - If a single mutation could lead to cancer
        - Cancer would be more prevalent
        - Phenomenon of age-related cancers
        - Delay upon exposure to carcinogens

    - Radiation exposure delay for cancer of 5-8 years following atomic bombing of Hiroshima and Nagasaki

Delay of Onset

- The survivors of the atomic bombs were dropped on Hiroshima and Nagasaki have been of long term studies on the effects of ionizing radiation on cancer incidence. The frequencies of different types of cancer in these patients ranged btwn the decades. At high levels in the 1950, of leukemia and cancer of the lung and thyroid gland were observed. The 60's and 70's brought high levels of breast cancer and salivary glands. In the 1980's, rates of colon cancer was especially high
- Why do you think different types of cancers varies rom one time to another?
  - Right after the bomb they are breathing in the radiation – lung and leukemia
  - Colon cancer – radiation has settled into the ground and now all their food has a slight contamination along with water supply so it is getting into the digestion system and it is a slow build up

Defects in Genome Stability and DNA Repair

- Cancer cells show higher rates of mutation, chromosomal abnormalities and genomic instability
  - Mutator phenotype
- Chronic myelogenous leukemia (CML)
  - Philadelphia chromosome
  - C-ABL gene fused to BCR gene
    - BCR-ALB fusion product is an abnormal signal transduction molecule stimulating cell proliferation in absence of growth signals
- Normal chromo 9 and 22 goes to 9 with a bit of 22 and 22 has Bcr Apl (which looses control of replication)**
- **Diagram on slide**

DNA Repair Defects

- Xeroderma Pigmentosum
  - Defective in nucleotide excision repair
  - Sensitive to UV light
  - Often develop skin cancer
- Hereditary Nonpolyposis Colorectal Cancer (HNPCC)
  - Defect in DNA mismatch repair

    - Increased risk of developing colon, ovary, uterine, and kidney cancers
- **Diagram on slide**

The Cell Cycle and Signal Transduction

- G0 – decision to reenter cell cycle/ exit cell cycle (quiescence)
- G1/S checkpoint – cell monitors size and DNA integrity
- G2/M checkpoint – Cell monitors DNA synthesis and damage
- M checkpoint – cell monitors spindle formation and attachment to kinetochores
- **Diagram on Slide**

Cell Cycle Regulation

- In addition to regulation at the checkpoints, cells control progression through cell cycle with:
    - Cyclins and CDKs (cyclin-dependent kinases)
        - The cell synthesizes and destroys cyclins in a precise pattern during cell cycle
        - Cyclins bind to their specific CDK which activates the CDK/cyclin complex
            - Selectively phosphorylates and activates other proteins to bring about advance in cell cycle

Cyclin Activity During Cell Cycle

- **Diagram on slide**
- G1/S – D2
- G1 - E
- S/G2 – A
- G2/M - B

Apoptosis

- If DNA damage cant be repaired, then the cell will activate programmed cell death
- Genetically controlled event leading to cell suicide
    - Nuclear DNA becomes fragmented
    - Internal cellular structures disrupted
    - Cell dissolves into small spherical bodies
        - Carried out by caspases (proteases)

Many Cancer-Causing Genes Disrupt Cell Cycle Control

- Two types of genes are mutated or mis-expressed in cancer cells, proto-oncogenes and tumor suppressors

- Proto-oncogenes. Important for normal cell function, especially cell growth and division
    - Transcription factors stimulated other genes
    - Signal transduction molecules that stimulate cell division
    - Cell cycle regulators that move the cell through the cell cycle
- Quiescent cell repress proto-oncogenes cannot be controlled
- In cancer cells, one of more proto-oncogenes cannot be controlled
    - Continually in an “ON” state which can constantly stimulate the cell to divide

__Tumor Suppressor Genes__

- Products normally regulate cell cycle checkpoints and initiate apoptosis
    - In normal cells they stop progression through cell cycle in response to DNA damage or from external environment
    - If tumor suppressor genes are mutated
        - Cells are unable to respond normally to cell-cycle checkpoints or unable to undergo apoptosis
        - Leads to further mutations and failure of cell to leave the cell cycle
        - Have to have both in order to have it full blown

__The Ras Proto-oncogene__

- Mutations in Ras convert the proto-oncogene to an oncogene preventing Ras from hydrolyzing GTP to GDP
    - Keeps Ras in its “ON” state permanently
    - Constantly stimulating cell division
    - Found mutated in 40% of human cancer
- 1. Growth factor binds to cell surface
- 2. Ras transiently exchanges GTP for GDP (to become active)
- 3. Ras sends signals to cascades of activated proteins
- 4. Signal transduction proteins activate transcription factors (express genes necessary for cellular division
- 5. Activation or repression of gene transcription
- Even without the presence of a growth factor the Ras will be turned on and thus constantly stimulating cell division
- They become oncogenic when one allele is mutated
- **Diagram on slide**

The p53 Tumor Suppressor Gene

- Most frequently mutated gene in human cancers (more then 50% of all cancers)
- P53 is a nuclear protein that acts as a transcription factor that represses or stimulates more then 50 different genes
- If a cell lacks p53, they are unable to arrest at cell-cycle checkpoints, or enter apoptosis in response to DNA damage
  - "Guardian of the Genome"
- **Diagram on slide**

The RB1 Tumor Suppressor Gene

- Retinoblastoma protein (pRB) controls the G1/S cell cycle checkpoint
- In G0, pRB binds transcription factors, blocking their function
  - Upon stimulation to enter G1. CDK4/cyclin D1 phosphorylate pRB releasing the transcription factors
- **Diagram on Slide (Go Over)**
- Retinoblastoma – highly seen in young children (cancer of the eyes)
  - Due to recombination from heterozygous to having a mitotic crossing over event
  - 50:50 chance of how they sit on the mitotic plate – 50% chance of having it

Cell Adhesion and Cancer Cells

- The features of metastasis require cells to leave the site of the primary tumor and invade other tissues
  - Must be able to dissociate from tumor
  - Must be able to digest components of basal lamina and extracellular matrix which normally contains tissues
- Alterations in genes encoding cell-adhesion molecules and proteinases
  - E.g. high levels of metalloproteinases are found in highly malignant tumors

Multistep Development of Colon Cancer

- **Diagram on slide**
- normal colon epithelium (APC mutation) – proliferating epithelium – benign adenoma – intermediate adenoma (Ras mutation) – late adenoma with villi (DCC mutation) – cancerous adenoma (P53 mutation) – colon cancer

Viruses Contributing to Cancer**

- **Diagram on slide**

Viruses that Contribute to Human Cancers

- Papillomaviruses (HPV) – sexually transmitted
  - Can cause cervical cancer
- Human T-cell leukemia virus
- Hepatitis B virus
- Epstein-Barr virus

Environmental Agents Contribute to Human Cancers

- Any substance that damages DNA is potentially carcinogenic
- Natural and man-made carcinogens are everywhere
  - Chemicals, radiation, some viruses and chronic infections
  - One of the most potent is cigarette smoke
    - 30% of cancer deaths can be attributed to exposure to cigarette smoke

Chapter 17 – Recombinant DNA Technology and Gene Cloning 13-04-14 5:2

Recombinant DNA

- A DNA molecule formed by joining together DNA from different biological sources
- Cloning can produce millions of copies of this recombinant DNA molecule
  - The sequence can now be manipulated
    - Change the sequence
    - Produce a protein
    - Produce novel proteins

Steps in Cloning DNA

1. Purify DNA
2. Cut DNA with restriction enzyme (endonucleases)
3. Join DNA to other DNAs that serve as vectors or carrier molecules (e.g. plasmids)
4. Recombinant DNA transferred to a host cell
5. Allow host cell to grow and divide, thereby replication the recombinant DNA molecule as well
6. Purify cloned DNA from host cells for analysis
7. Cloned DNA may also be transcribed to produce the gene product for research or commercial use
   - In bacteria – immune system cuts invading DNA (endonuclease)

- **Diagram on Slide**
- Know how to name
- Given sequence say which is a restriction enzyme
  - Same forward on one strand and backwards on the other (ATGCAT – A and T are complementary)

Vectors (Plasmids)

- Fragments of DNA can't be introduced into a host cell directly
  - They don't have the necessary sequences for replication, copy number and selectable markers

Vectors carry out 3 important functions

- Replicated independently along with DNA fragment it caries once in the host cell (it has an origin of replication)
- Contain several restriction sites for cloning
- Carries a selectable marker (usually antibiotic resistance)
  - Identifies host cells that contain vector

        - Maintains the vector in the host cell

Plasmids

- Naturally occurring, extrachromosomal and self replicating dsDNA molecules found in certain bacteria
- Over the years plasmids have been extensively modified to serve as cloning vectors for a variety of application
- **Diagram on slide**

pUC 18 Plasmid

- Small 2686 bp
- Origin of replication, can produce up to 500 copies per cell
- Large number of restriction enzyme sites for cloning, especially polylinker sites
- Selectable marker for ampicillin resistance (beta- lactamase)
- Polylinker in the lacZ gene for blue-white color selection
    - Active lacZ gene product cleaves Xgal into blue product
    - Inactive lacZ gene product can't cleave, so cells are white
- **Diagram on slide**

Cloning into Other Microorganisms

- Yeast artificial chromosomes (YACs)
    - Large fragments can be cloned (up to a million bp)
    - Useful in human genome project
    - Essentially turn into a linear chromosome
- **Diagram on slide**

For Expression of Eukaryotic Genes, Yeast is the Host of Choice

- Eukaryotic proteins may need to be modified to have optimal activity (e.g. through the ER or golgi)
    - Bacteria lack these organelles
- **Chart on Slide**

The Polymerase Chain Reaction (PCR)

- In 1986 a technique was developed that accelerated recombinant DNA methodology
- PCR is a rapid method of copying DNA in vitro
    - No need for host cell to amplify DNA

    - Method of choice in most cloning protocols
    - Produce billions of copies of a specific DNA sequence in a matter of hours
- **Diagram WITH STEPS on slide**
- 3 steps
    - Denaturation at 95*
        - Separates dsDNA into ssDNA
    - Annealing at 50-70*
        - ssDNA primers base pair with specific sequences on target DNA (annealing)
        - Temperature veries and depends on the Tm of primer
    - Extension at 72*
        - Optimal temperature of TAQ polymerase activity
        - Extends primers from 5’ to 3’
- These 3 steps comprise 1 cycle. A typical PCR protocol calls for 25-30 cycles
- The amount of target DNA will double with each cycle
- Important enzyme used is TAQ polymerase
    - Isolated DNA polymerase from Thermophilus aquactius, a bacterial strain isolated from hot spring
    - Its enzymes are stable at high temps, so it wont be denatured with repeating denaturing temps of 95* during PCR cycling

cDNA Library

- Contains DNA copied from mRNA in a cell population at a given time
    - Remember that not all genes are expressed at same time, so a cDNA library will only have a subset of genes from the entire genome
        - E.g. we could create a cDNA library for liver cells and another for skin cells (some mRNAs would be present in both, while some other in only skin and others only in liver)
- Stands for complementary DNA
    - We can only clone dsDNA into vectors for cloning into a library
    - Challenge is to make a dsDNA copy of mRNA
- To make the cDNA library the enzyme reverse transcriptase will be used
    - Make DNA copy of mRNA
- Use poly T primer to anneal to poly A tail
- *enzymes and substrates
- **Diagram on slide**

<u>Retrieval of Specific Clones from Libraries</u>

- A probe is used to screen a library to recover a specific gene
    - The probe can be a nucleic acid probe or an antibody
- For nucleic acid probes, the probe is make single stranded (usually by heating to 95*) and labeled with radioactivity, such as 32P
    - ssDNA probe will anneal with complementary sequence in library. Location of binding will be seen with location of radioactivity
- Has to be able to hybridize, so probe and DNA have to be single stranded
- **Diagram on slide**

<u>Analysis of Cloned Sequences</u>

- Restriction mapping
    - Map out the location of restriction sites within sequence
        - Valuable info for eventual sequencing and subcloning of DNA fragments
    - Serve as marker within the DNA
- **Diagram on slide – WILL BE ON FINAL**

<u>Nucleic Acid Blotting</u>

- Following agarose gel electrophoresis, we can identify DNA fragments (usually those from restriction digests) that bind to a specific probe (nucleic acid)
- Southern blot
    - DNA on gel, probed with nucleic acid probe (ssDNA or ssRNA)
- Northern blot
    - RNA on gel, probed with nucleic acid probe
- Western blot
    - Protein on gel, probed with an antibody probe
- **Diagram on slide**
- Need to get DNA onto a membrane from the gel – blotting
- Uses a buffer to get DNA to transfer to the membrane
- Then you can probe

<u>DNA Sequencing</u>

- The most common method of sequencing DNA is the Sanger method
    - Purified DNA is heat denatured to make single stranded and distributed to 4 tubes
    - A single primer is added to anneal to 3’ end of each strand

  - DNA is synthesized with DNA polymerase and dNTPs…. Although, a ddNTP is added (a different on in each of the 4 tubes, ddATP, ddCTP, ddTTP, ddGTP)
    - Terminates replication – are labeled so you can see it
- **Diagram on slide**

DNA Sequencing Gel

- Read sequence from bottom up
  - E.g. 5’-CGCTTTCATGTCA…..
  - 5’ end is closer to the original primer
  - think about how the reaction would change by altering the proportion of the ddNTP in each tube

Automated DNA Sequencing with Fluorescent Dyes

- **Diagram on slide**

Chapter 18 - Genomics, Bioinformatics and Proteomics 13-04-14 5:22 PM

Sequencing Genomes

- Structural genomics
    - Sequencing genomes
    - Comparing sequences to other sequences to identify genes and regulatory sequences
- Methods of sequencing
    - Map based approach
        - All sequencing done in ordered fashion of overlapping clones
        - Human genome project
    - Shot gun
        - Sequencing done in random fashion and then ordered later on
        - Celera Genomics corp.
- **NEXT 3 DIAGRAM VERY IMPORTANT**

Looking for Similarities

- BLAST (basic local alignment search tool)
    - Compares two sequences at a time for best possible alignment
    - Shows similarities, differences and common domains btwn genes
- **Diagram on slide**

Human Genome Project

- Human genome project
    - Initiated in 1990
        - Proposed completion date, 15 years later
    - Under direction of Dr. Francis Colins
    - Sequences 3 billion base pairs
        - All 22 autosomes, X and Y
- Celera Genomics
    - Started project in 1998
    - Under direction of Craig Venter
    - Proposed sequencing genome in 2 years
        - Using their shot gun approach
- In 2001 both groups co-published the results
    - Some work still was needed to fill in some gaps
        - Centromeres, telomeres, and repetitive sequences to verify sequences (approx.. 600 000 errors present)
- **Table on slide (Know 1st 5)**

Genes Identified in the Human Genome

- ~42% of genes function not known
- **Diagram on slide**

Representative Genomes Sequenced for Many Other Species

- Over 400 prokaryotic and eukaryotic genomes have been sequenced (model organisms)
  - Yeast
  - Bacteria (E.coli and many others)
  - Nematodes
  - Plants
  - Mice (best human model)
  - Fish (zebra fish)
  - Fruit flies

E. coli Genome

- ~20% of the genes don't know the function of
- **Diagram on Slide**

Evolution of Multigene Families

- Genes share similar but, not identical DNA sequences
  - Usually have a similar function
  - Group of multigene families – a superfamily
- Globin gene superfamily
  - Myoglobin
    - Oxygen carrying protein found in muscle
  - Alpha globin subfamily
    - 3 genes
      - Zeta (expressed only in embryogenesis)
      - Alpha 1 (expressed only during fetal stage)
      - Alpha 2 (expressed only in adult)

Beta Globin Subfamily

- Five genes over 60kb of DNA
  - Order of genes parallels their expression during development (as does alpha-globin)
  - Three expressed before birth (epsilon, gamma-a, and gamma-g)

  - Delta and beta expressed after birth and throughout life
  - A single pseudogene (phi-beta 1) is also present in this cluster
    - Similar DNA sequence to other globins but not expressed

<u>Comparison of Amino Acid Sequences of Globins</u>

- Blue are conserved btwn the 2
- ~50% are the same
- patch that are identical have the same vital function that gives the protein its characteristic (e.g. sugar binding)
- **Diagram on slide**

<u>Organization of Globin Genes</u>

- 2 pseudo genes are for alpha
- when you have duplications (mutations) of a gene you allow for changes to occur, but don't affect the organism (novel, same similar function)
- due to different selection pressures
- **Diagram on slide + Evolution of Globin Genes**
- Mutations can be the driving force of evolution

<u>The Transcriptome</u>

- Looking at whether the gene is transcribed and when/where
- Make short sequences to recognize every individual gene (~20 nucleotides)
- **Diagram and STEPS on slide**

<u>Proteomics: 2D Gel Electrophoresis</u>

- Separate based on charge (in pH)
- Push gel out and put on SDS gel
- Protein separate based on molecular weight
- In SDS (-ve) solution
  - Each band may not be a single protein
  - Need to do 2D
- To differentiate between 2 (e.g. cancers)
- Even though it may be transcribed doesn't mean it will be translated
- **Diagram on slide**

<u>Proteomic Analysis</u>

- Cells grown in vitro and proteins are extracted
- Run on a 2D gel
- Compare the 2 gels (one will express a certain protein and one wont)
- Cut out protein
- Mass spectrometry and comparison to protein databases to identify the protein in the spot

Chapter 19 – Applications of Genetic Engineering 13-04-14 5:22 PM

Genetically Engineered Organisms Synthesize Biological & Pharmaceutical Products

- Insulin in bacteria
    - First human gene product manufactured with recombinant DNA technology in 1982
    - Previously isolated from pancreas of cows and pigs in slaughterhouses
        - Originally isolated from dog pancreas in Banting and Best
- **Diagram on slide**

Recombinant Approaches for Vaccine Production

- Traditional vaccines
    - Inactivated vaccines
        - Prepared from killed samples of bacteria and viruses
    - Attenuated vaccines
        - Weakened or live pathogen which cant reproduce causing a mild form of disease
        - Better immune response
- Subunit vaccines
    - Cloned surface proteins of pathogenic bacteria and viruses
        - Protein acts as an antigen and stimulates the immune system to make antibodies against the pathogen
        - E.g. Gardasil is a subunit vaccine against HPV

Plants Expressing Recombinant Proteins: Edible Vaccine

- Gene from human pathogen is inserted into a vector
- Vector is introduced to plant cells
- Leaf segments sprout into whole plants carrying gene from human pathogen
- Eating plant triggers immune response to pathogen
- Vaccines have to keep cold so that they don't degrade over time
- In some places this is hard to do so this approach is used

Selective Breeding

- Farmers have been selectively breeding crops for thousands of years
- Maize yields have increased 4 fold over last 60 years (increase in kernel size)

Herbicide Resistance in Crop Plants

- Glycophosphate is a herbicide that kills all plants at low concentrations, is biodegradable and is non-toxic to humans

  - Active herbicide is round up
  - Kills plants by inhibiting chloroplast enzyme ESPS synthase
    - Necessary for AA biosynthesis in both plants and bacteria
- A glycophosphate resistant gene was introduced into certain crop plants
  - Now 75% of soybean and cotton plants are Roundup ready
- **Diagram on slide**

<u>Other Examples of Trangenics</u>

- Golden rice
  - Introduction of vitamin A into rice
  - Increases nutritional value of a staple for many in 3rd world countries
    - Could dramatically reduced blindness
- Transgenic Atlantic salmon
  - Chinook salmon growth hormone adjacent to constitutive promoter
  - Grow 400-600 percent faster

<u>Sometimes, Things Just Don't go Your Way!</u>

- Leptin
  - A 167 AA peptide hormone found in adipose tissue
  - Low levels associated with obesity in mouse models
  - A known out mouse was created that could not product leptin (became fat)
    - Injection with leptin can reverse obesity in mice
    - Unfortunately, the same is not true in humans
- Increase in fat – inject leptin – decrease in food intake
- Does work for some humans – obesity was due to a problem with the leptin gene

<u>Uses for DNA Technology: Allele-Specific Oligonucleotides (ASOs)</u>

- A way to test for different diseases in the human population
- Detect single nucleotide polymorphisms (SNPs) in many disease causing alleles
  - E.g. beta-globin allele variant that causes sickle cell anemia
- Since only a small amount of DNA is required, this technique can also be used for preimplantation genetic diagnosis
- Typically only a single nucleotide change
- Make oligonucleotides (10 nucleotides long) to differentiate the 2 (wt and mt)
- ASO testing for sickle cell anemia
- **Diagram on slide**

- Amplify the region by PCR
- Hybridize with ASO
- White = none of it (AA or SS)
- Tell you their genotype

Genetic Analysis Using Microarrays

- Each gene chip can contain up to 500 000 different fields, each representing a different sequence
  - May design gene chips containing SNPs associated with many known diseases
  - Determine genotype with a simple full genome scan
    - Will all newborns be screened this way
    - Would you want to know your genotype?
      - Insurance companies might want to know….. ethical dilemma

Microarray Analysis Comparing Gene Expression

- May be important diagnostic tool
  - E.g. patients with activated B-like cells have a high rate of death (16 in 21) compared to GC-B like cells (6 in 21)
  - Each type may respond to treatments differently
- Want to know the type of cell they have so we know what treatment would be the best (especially for cancers)
- **Diagram on slide**

STRs

- Short tandem repeats (STRs) or microsatellite sequences, contain very short (2-6 bp) tandem repeats and are highly polymorphic
- Examples are the dinucleotides repeat (GT)^n and the trinucleotide repeat (CAG)^n
- Many are polymorphic and are used for genetic mapping and forensics
- STRs are usually typed by PCR with primers flanking the sequence
  - A population may have many different allele lengths for STRs
  - An individual may be either homozygous or heterozygous for a particular STR
- DNA between genes isn't coding for anything
- Polymorphic – no slection pressure against it
  - Not coding for anything
  - Could differ people by STRs using PCR
- **Diagram on slide**

DNA Profiling Based on DNA Microsatellites

- Have a repeating motif repeated from 7-40 times in different regions of the genome
    - Thousands of different STRs are found in the human genome, only a small subset are used for DNA fingerprinting (13)
- An STR profile is created by amplifying each STR region by PCR and then running the sample on a gel (capillary gel electrophoresis) to determine the size of the PCR fragment
- **Diagram on slide**
- Different ethnicities have a different number of repeats – this is how you can profile them
- Multiplex PCR – many primers in one test tube and all run at the same time

Capillary Electrophoresis Readout for STR Profiles in a Rape Case

- **Diagram on slide**

Human Identity Testing

- Forensic cases – matching suspect with evidence
- Paternity testing – identifying father
- Historical investigation – e.g. Thomas Jefferson
- Missing person investigation
- Mass disasters – putting pieces back together
- Military DNA dog tags
- Convicted felon DNA database

Following the Y Chromosome

- During Jefferson's time at the White house, there was speculation that he fathered children with one of his slaves
- Sally Hemmings lived in the Jefferson home while in Paris (as ambassador) and in Virginia (Moticello)
    - Jefferson was thought to have fathered two or three children with Hemmings
- None of Jeffersons male children survived to have children of their own
    - However a male descendant of Field Jefferson was alive today (uncle)
    - The Jefferson Y chromosome was compared to the Hemmings family Y chromosome
        - They were the same

Chapter 22 – Quantitative Genetics

Discontinuous Traits

- Most of the traits we have examined so far have been discontinuous traits
    - Either yellow or green
    - Either tall or short
- Each trait was markedly different from one another
    - Single identifiable phenotype was observed
- For discontinuous traits, the phenotype often gives insight to the genotype of the organism
- Complicating factors such as penetrance, expressivity, pleiotrophy, and epistasis make determining the genotype more difficult

Continuous Traits

- Many traits exhibit a wide range of possible phenotypes
    - E.g. human birth weight, adult height, and number of eggs laid by fruit flies
- Traits like these are called continuous traits
    - Since these traits are usually described by quantitative measures (e.g. how tall, how many), these traits are also known as quantitative traits
- **Diagram on slide**

Questions Asked in Quantitative Genetics

- What is the influence of genes and the environment on the variation in phenotypes?
- How many genes determine phenotype?
- Are the contributions of the genes equal?
- Are the effects of the alleles additive?
- When selection occurs for a particular trait, how rapidly can that trait change?
- What is the best method for selecting and mating individuals to produce desired phenotypes in the progeny?
    - Agriculture – how much milk is produced by cows

Polygene Hypothesis for Quantitative Inheritance

- A trait may have a range of phenotypes because of enviro influences
    - Same genotype may produce a range of phenotypes
    - Multiple genotypes may produce same phenotype
- Work in the early 20th century indicated that both enviro and genotype influence quantitative traits
    - These are multifactorial traits
    - Phenotype cannot be explained using a single locus

        - Explanation for these traits is the polygene or multi-gene hypothesis for quantitative inheritance

Not All Polygenic Traits show Continuous Variation

- Meristic traits – phenotypes described by whole numbers
    - E.g. seeds in a pod or eggs laid by a chicken
- Threshold traits – small number of discrete classes, found in a number of human diseases
    - E.g. type II diabetes
        - Increase # of predisposing alleles increase their chance of getting into the threshold (need to control enviro)
        - Past threshold you will have it (enviro may or may not contribute)
- **Diagram on slide**

Polygene Hypothesis for Kernel Color

- The work of Hermann Nilsson-Ehle in 1909
- Crossed true breeding red kernel to white kernel plants
    - F1 all were the same intermediate shade between red and white – pink
    - F2 kernels were 15 red (all shades) to 1 white
        - When categorized the red shades he found a phenotypic ratio of 1:4:6:4:1
- **Diagram on slide**
- Dominant = contributing alleles (produce pigment)
- Recessive = non-contributing
- Did not observe 9:3:3:1 ratio but a 1:4:6:4:1 ratio
- Alleles can either be contributing alleles or non contributing alleles
    - Each contributing allele will allow for the synthesis of some pigment. Kernel coloration is depend on the number of R or C alleles (the r and c alleles are noncontributing)
        - E.g. RRCC dark red, rrcc white

Polygenes

- Multiple gene hypothesis fits some observations of quantitative inheritance
- May be explained by the action and segregation of allelic pairs at a number of different loci called polygenes
    - Each polygene has a small effect on the overall phenotype
- **Diagram and Example on slide**

Calculating the Number of Polygenes

- 1 / (4^n) = ratio of F2 individuals expressing either extreme phenotype (one or the other)
- E.g. 1 / (4^n) = 1/16
  - 4^n = 16
  - nln4 = ln16
  - n = ln16/ln4
  - n = 2 genes
- **Diagram on slide**

Study of Polygenic Traits Requires Statistical Analysis

- Large data sets from a population may lead to data with a normal distribution
  - Characteristic bell shaped curve
  - Several statistical concepts can be used here
    - Mean
    - Variance
    - Standard deviation
    - Standard error
    - Covariance

Stats in Genetics

- We know that the phenotype of an organism is partly genetic and partly enviro
- As geneticists we need to know how much of the variation that exists in a population is genetically determined and how much is determined by the enviro
- Quantitative genetics addresses this question of nature vs. nurture, or the relative roles of genes vs. the enviro
  - Note that this is an oversimplification of the problem

Variation

- Variation in the phenotype of a population is Vp
  - The proportion of this variation that is genetic variation is Vg, while the variation from the enviro is Ve
  - Vp = Vg +Ve
- how do we partition the phenotypic and genetic components?
  - We need to understand some statistical tools first

Samples and Populations

- Lets say we want to know the weight of all babies born in Ontario in 2006
  - This would be a laborious analysis to collect the data from thousands of records
  - Scientists commonly collect data from a subset of the group. They might collect data from only 500 babies born in Ontario
    - Using this data, an estimate of the average weight of babies born in Ontario in 2006
- The group of ultimate interest (all infants born in Ontario in 2006) is called the population
  - The subset (or group of 500) used to give us an estimate of the total population is called the sample

The Mean

- The mean or average indicates the center of the distribution of phenotypes in a normally distributed phenotypic trait
  - Calculated by adding up all the individual measurements and dividing by the number of measurements that were added
- **Diagram on slide**

The Variance and Standard Deviation

- Variance is a measure of how much the individual measurements spread out around the mean
  - How variable the individuals and their measurements are
- Two distributions can have the same mean, but have dramatically different distributions or variances
- **Diagram on slide**
- The sample variance is symbolized by s^2, is defined as the average squared deviation from the mean
- **Diagram on slide**

Standard Deviation

- Often preferred over variance since it shares the same unit of measurements with the data (variance is in the units squared)
- A theoretical normal distribution can be described by the mean and standard deviation
  - 68.3% of individuals will have values within one SD above or below the mean
  - 95.5% of the values fall within two SD’s
  - 99.7% of the values fall within 3 SD’s

Quantitative Genetic Analysis – Inheritance of Ear Length in Corn

- two true breeding strains of corn, each displaying little variations in ear length
    - Mexican sweet corn (mean length 6.63 cm)
    - Tom Thumb popcorn (mean length 16.80 cm)
    - Assume that parentals were homo for the genes controlling their length
    - F1 plants would thus be hetero for most genes, and importantly, would all have the same genotype
- Vp = Vg +Ve
    - Genotype is 0 – no genetic variation
- The F1 plants all have the same genotype
    - The range in ear length phenotypes must result from factors other than genetic influence
- Mean of F2 is close to the variation of F1
    - Due to genetic factor in addition to enviro
- **Diagram on slide**
- If the enviro was responsible for variation in the parental and F1 generations, then it should have the same effect of the F2
    - Enviro should not have a greater effect on the F1 compared to the F2
    - Increased variability of F2 is likely due to the presence of greater genetic variation in the F2 compared to the F1 or parents

Some Generalizations to Quantitative Inheritance Studies

- Mean of F1 is usually intermediate btwn means of parents
- Mean of F2 is usually approx. equal to F1
- F2 almost always shows more variability around the mean then F1
- The extreme values in the F2 extend closer to the two parental values than do the extreme values of the F1
- **Diagram on slide**

Heritability

- Proportion of a population’s phenotypic variation that is attributable to genetic factors
- Important to know genetic contribution of traits
    - In agriculture such as weight gain in cattle, milk production in cows and # of eggs laid by chickens

    - In natural populations such as body size, fecundity and development rates can help us understand natural selection and evolution
- Often misused
    - Family of a certain stature or intelligence is assumed to be genetically determined
        - All members of the family also have the same enviro

<u>Phenotypic Variance</u>

- Phenotypic variance (Vp) is a measure of all variability for a trait
    - Genetic contribution to the phenotypic variance is called the genetic variance (Vg)
    - Environmental contribution to the phenotypic variance is called environmental variance (Ve)
        - Includes any nongenetic variation such as temp, nutrition, and parental care
- Vp = Vg +Ve

<u>Covariance (COVg,e)</u>

- 100% of the variation among individuals is accounted for by genetic and envrio factors
- However, the sum of genetically and environmentally caused variance may not add to the total phenotypic variance
    - Here the genetically and environmentally caused variance covary
        - Individuals above average genetic quality receive above the average resources leading to a covariance btwn genotype and the enviro
        - Another term is needed to account for this situation COVg,e

<u>Genotype-by-Environment Interaction (Vgxe)</u>

- Relative effects of genotypes differ among enviro
    - in cold enviro AA plants are 40 cm tall on average, while aa plants are 35 cm tall
    - in warm enviro AA plants are 50 cm tall, while aa plants are 60 cm tall
        - both plants grow taller in warm enviro so there is an enviro component
        - there is also a genetic effect with the relative performances of the genotypes differ in the two temperatures
            - while both genetic and enviro differences contribute to the phenotypic variance, the effects on the genotype and enviro cannot simple be added together
- **Diagram on slide**

<u>Phenotypic Variation</u>

- Vp = Vg +Ve +COVg,e + Vgxe
- Note: individual components may equal zero depending on the genetic composition of the population, the specific enviro and the manner in which the genes interact with the enviro

Additive Genetic Variance (Va)

- Some of the genetic variance can be accounted for by the action of the additive effects on different alleles
    - Genes that have additive effects and variation resulting from this sort of gene action is called additive genetic variance (Va)

Dominance Variance (Vd)

- Some genes exhibit dominance and this produces a source of dominance variance (Vd)
- Will only contribute to the Vp when both the homo recessive and either the homo dominant or hertero are present in the population
    - In an F1 that is backcrossed to a homo dominant parents, there will be no increase in variability

Interaction Variance (Vi)

- Epistatic interactions may occur among alleles at different loci to determine the phenotype
    - This presence adds another source of genetic variation called epistatic or interaction variance (Vi)
- May have several different genotypes at one locus, but their penetrance could be affected by variation at another locus
- Genetic variance can be broken down into
    - Vg = Va +Vd +Vi

Environmental Partitioning

- General enviro effects (Veg)
    - Individuals raised in nutritionally deprived conditions
- Special enviro effects (Ves)
    - Transient changes in phenotype such as pigmentation changes of skin upon exposure to sunlight
- Common family enviro effects (Vecf)
    - Egg deposition of insects on certain plants. Larvae feed on the plants. Some plants have toxins and other offer better nutrition value
- Maternal effects (Vem)

    - Growth within the uterus of different mothers and the nutrient value of breast milk are examples
- Vp = Va + Vd + Vi + Veg + Ves + Vdf +Vem + 2COVg,e + Vgxe

Broad Sense Heritability

- How much of the phenotypic variation (Vp) can be attributed to genetic variance (Vg)
    - This is broad sense hetitability
        - H^2 = Vg/Vp
- Heritability of a trait can range from 0 to 1
    - If 0, then none of the variation in phenotype among individuals results from genetic differences
    - If 1, then all of the phenotypic variance is genetically determined
        - Can be 0 if no variability in the population

Narrow Sense Heritability

- There is an important relationship btwn additive genetic variation and artificial and natural selection
    - Dominance and epistasis are complicating factors for both the breeder and mother nature
- Additive genetic variance allows one to make accurate predictions about the resemblance of parents and offspring
    - Geneticists often determine the narrow sense heritability which is the proportion of the phenotypic variance that results only from additive genetic variance
- h^2 = Va/Vp
- Can select for particular alleles

Heritability Issues

- Broad sense heritability does not define the complete genetic basis of a trait
    - Depends on genetic variation which may not be present. When a trait is not variable, we cannot determine the heritability
- does not indicate what proportion of an individuals phenotype is genetic
    - Is a characteristic of a population not an individual
- Is not fixed for a trait
    - Can only be applied to a specific population in a specific enviro
- High heritability for a trait does not imply that differences btwn populations for the same trait are genetically determined

    - Enviros might be different
- Trait shared by members of the same family do not necessarily have high heritability
    - These are familial traits and can be a result of a shared unique enviro

Calculating Heritability

- If the additive genetic component of variation is important in determining a trait, then we would expect the offspring to resemble their parents
- We can analyze this relationship using correlation and regression analysis
    - Important to use the midparent value which is the mean of the tow parent's phenotypic values
- Represent the relationship by plotting midparents values vs. the mean phenotype of the offspring
    - Each point represents one family
    - Slope provides info about the magnitude of the narrow sense heritability
        - When the slope is 1, genes with additive effects determine all of the phenotypic variance
        - When the slope is less then 1 but greater then 0, then other genetic influences may play a role
- **Diagram on slide**

Twin Studies and Heritability in Humans

- Traditional heritability studies not possible in humans
- Human twin studies are useful tools to examine how much phenotypic variance is due to the genotype vs. the enviro
    - Mono twins are derived from splitting of a single egg, genotypically identical
        - Genes expressed?
            - X inactivation in females, methylation
    - Di twins originate from two separate eggs and have 50% of genes in common
        - But they share the uterus so the enviro is same as mono twins
- For a given trait
    - Phenotypic differences btwn identical twins is equivalent to enviro variance (Ve)
    - Phenotypic differences btwn di twins represents both Ve and approximately half of Vg
    - We can estimate the broad sense heritability by comparing these two studies
- Another approach
    - Looks at twins reared together vs twins reared apart in different settings

- Concordant for a trait if both express it or both don’t
- Discordant if one expresses and one doesn’t
    - Compare MZ and DZ twins and determine heritability (problems?)

Population Genetics Outline

- Population genetics is the field of genetics that studies heredity in groups of individuals for traits that are determined by one, or only a few genes
    - Both population and quantitative genetics use Mendilian principles and both are amenable to mathematical treatment

Population Genetics

- A population is a group of individuals who share a set of common genes, live in the same area and actually or potentially interbreed
- All of the alleles shared by these individuals constitutes a gene pool

Goals

- Computing frequencies of various alleles in the gene pool
- Computing the frequencies of different genotypes in the population
- Computing the changes in frequency that occur from one generation to the next

Questions

- How much genetic variation is present in a population
- Are genotypes randomly distributed in time and space or is there a pattern in their distribution
- What processes affect the composition of a populations gene pool
- Do these processes produce genetic divergence among populations

Genotypic Frequencies

- Are a way to study the genes in a particular gene pool by quantifying the genotypes (pairs of alleles) at a given locus
- To calculate, count individuals with one genotype and divide by total individuals in population. Repeat for each genotype in the population
- Frequencies are proportions that can range from 0-1. For any given trait, the sum of the genotypic frequencies in a population should be 1

Allele Frequencies

- Alleles frequencies give more info about the structure of the population than genotypic frequencies
- May be calculated in two different ways
    - From observed number of different genotypes at a particular locus

  - From genotypic proportions
- Allele frequencies at an X-linked locus are more complex because one sex will have only one X-linked allele while the other has 2

The Hardy Weinberg Law

- This is a simple explanation showing how Mendelian segregation influences allelic and genotypic frequencies in a population
- There are three parts to the law: one set of assumptions and two major results

Assumptions of the H-W Law

- This law is a simplification of complex events. There are certain assumptions that must be present for the law to apply
  - The population is infinitely large, to avoid genetic drift. (impossible so large populations are studied when possible)
  - Mating is random with regard to traits
  - There is no natural selection of traits
  - No mutations occur
  - No migration occurs

Predictions of the H-W Law

- If the conditions are met, the population will be in genetic equilibrium, with 2 expected results
  - Allele frequencies do not change over generations, so the gene pool is not evolving at the locus under study
  - After one generation of random mating, genotypic frequencies will be p^2, 2pq, and q^2 and will stay constant in these proportions as long as the conditions are met. This is H-W equilibrium, which allows predictions to be made about genotypic frequencies

Derivation of the H-W Law

- Zygotes are formed by random combinations of alleles, in proportion to the abundance of that allele in the population
- When a population is in equilibrium, genotypic frequencies will in the proportions of p^2, 2pq, and q^2. This result is the expansion of the square of the allelic frequencies
  - (p+q)^2
- Mendilian principles acting on a population in equilibrium will work to maintain that equilibrium

- **Diagram on slide (Important)**

Using the H-W Law to Estimate Allelic Frequencies

- If one or more of the alleles is recessive, one can't distinguish btwn heterozygous or homozygous dominant individuals
- Can use H-W law to calculate the allele frequency based on info about the number of homo recessive individuals
- **Albinism Example**

Extensions of the H-W Law to Loci with More than 2 Alleles

- Often more than 2 alleles are possible at a given locus
- If 3 alleles are present with frequencies p, q, r, the frequencies of the genotypes at equilibrium will be
  - (p +q +r)^2 = p^2 +2pq +q^2 +2pr + 2qr +r^2
- **Blue Mussel Example**

Extensions of the H-W Law to Sex Linked Alleles

- In species where sex is chromosomally determined, humans for example, females have 2 X's and males are XY
- In females, H-W frequencies are the same as for any other locus. In males, frequencies of the genotypes are the same as frequencies of the alleles in population
- Males receive their X chromo from their mothers, so the frequency of X-linked allele will be the same as the frequency of that allele in their mothers. For females, the frequency will be the average of both parents
- With random mating, the difference in allelic frequency btwn sexes will be reduced by half in each generation. One generation after allelic frequencies become equal in males and females, the genotypes will be in H-W equilibrium

Genetic Variation in Space and Time

- The genetic structure of population can vary in space and time
- An allele frequency cline is a clear pattern of variation across a geographic transect, usually correlated with a physical feature such as temp and rainfall
- Statistical tools are used to quantify spatial patterns of genetic variation. These are important in conservation biology
- **Diagram on slide**

Genetic Variation in Natural Populations

- Important component in natural populations
    - Determines the potential for evolutionary change and adaption
    - Provides clues about roles of various evolutionary processes
    - Allows predictions about a populations chances for long term survival
- No variation = no evolution
- No variation due to no different alleles to be accounted for

Measuring Genetic Variation at the Protein Level

- Understanding genetic variation in a population was difficult before molecular biology because most phenotypes are the result of multiple genes and there was no technique available to determine allele frequencies
    - Couldn't tell it if was homo dominant or hetero until electrophoresis
- Protein electrophoresis separates proteins on the basis of size, charge and conformation and so often can separate the gene products of different alleles
    - Remember how this was used to show there were 2 forms of hemoglobin sickle cell anemia
- Amount of genetic variation within a population is usually measured by 2 parameters
    - Proportion of polymorphic loci (those with more than one allele within a population)
    - Heterozygosity (the proportion of an individuals loci that are hetero). Determine for individuals and then average to obtain estimate of hetero of the population
- Unfortunately, proteins of similar sizes and charges will conform in gel electrophoresis, so the allele differences are likely to be underestimated
- **Diagram on slide**

Genetic Variation at the DNA Level

- Restriction fragment length polymorphisms (RFLPs) are a quick way to map genes and gain an idea of how many DNA differences occurs within a population. Limitation of RFLP analysis is that it only assess variation in sites for particular restriction enzymes
- **Diagram on slide**
- Decrease when they need to be maintained
- DNA sequence analysis shows that there is more variation and therefore many more alleles of most genes than previously believed
- Different regions will have different levels of variation. Exons are less likely to vary than intros and flanking sequences and changes within exons are often synonymous

    - These are nucleotide changes that do not change the amino acid sequence (usually at the wobble codon)
    - Coding sequence is resistant to change
- DNA length polymorphisms result from deletions and insertions of short stretches of nucleotides, especially in non coding regions
- Microsatellites or short tandem repeat polymorphism (STRs) are very simple repetitive sequences that occur in different numbers of times in different individuals
    - More then 8 000 STRs have been mapped in the human genome
- Both are useful in determining genetic variability within a population

Mutation Creates New Alleles in a Gene Pool

- Usually a mutation converts one allelic form of a gene to another
- May be neutral, detrimental or advantageous depending on the enviro
    - Enviro changes may favor alleles different form those previously favored
- The frequency of alleles in a population is determined by interaction of mutation rates and natural selection
- Mutation is the ultimate source of genetic variation in populations

Sources of Variation: Mutation

- Mutations are a source of variation
    - Provides genetic changes in a population that can be acted upon by natural selection to produce changes
- Rate of change in gene frequency is low
    - Mutation rates are usually very low
    - Defined as: the probability that a copy of an allele changes to some other allelic form in one generation
- Increase in frequency of a mutant allele = mutation rate x freq. of nonmutant allele
- E.g. population were completely homo for A. mutations to a occur at a rate of 1/ 100 000 per newly formed gamete
    - In the next generation the frequency of the a alleles would be only
        - f(a) = 1.0 x (1/ 100 000) = 0.00001
        - f(A) = 0.99999
- The rate of increase of the new allele is very slow and it gets slower every generation
    - Because there are fewer copies of the old allele still left to mutate
- We can look at the change in the frequency of one allele over time (from Pn to Po) with a constant mutation rate (mew)

    - **Know equation on slide**
- **Example on next slide**
- **Diagram on slide**

Acondroplasia Example: Mutation

- Acondroplasia (A_) is a dominant form of dwarfism in humans
    - Individuals with this skeletal disorder have an enlarged skull but short arms and legs
- **Example on slide**

Cystic Fibrosis and Mutation

- An evolutionary perspective can often lead to medical discoveries
- Autosomal recessive cystic fibrosis caused by loss of function in a gene for a cell surface protein called cystic fibrosis transmembrane conductance regulator (CFTR)
- Frequency of mutant allele is about 2 percent in European populations
    - Yet, until recently, most individuals with two of these mutant alleles died before reproducing
- Chloride is usually pumped out of the cell
    - They lack this and water moves into the cell
    - This gums up everything
    - Can't move mucus out of their lungs and other organs
- The puzzle: with selection acting against the presence of the CFTR mutant allele, why is the mutant allele maintained within the population at a high frequency of 2 percent?
- Two possibilities
    - Mutation selection balance. Mutations are constantly creating new mutant alleles to replace those lost by selection
        - For this scenario to work the mutation rate would have to be extremely high (5x10^-4) you could almost feel the pain of your DNA changing at this rate
    - Heterozygous superiority hypothesis
        - Here, selection against the homo mutant is counterbalanced by selection in favor of the hetero
        - Usually selection for the hetero is for the resistance to a disease
        - Recent work suggest CF hetero may have enhanced resistance to typhoid fever

Cystic Fibrosis and Typhoid Fever

- Typhoid fever is caused by a bacterium which infiltrates the lining of the intestine

  - Lab strains of mice that were hetero for CFTR delta 508 (analog of human mutation) acquired 86% fewer bacteria then mice without the mutation
  - Do humans also show this increased resistance?
    - If so, then CFTR is like sickle cell anemia exhibiting hetero advantage

Selection in Natural Populations

- The effects of natural selection on allele frequencies has been studied in the lab and in natural populations
- Detailed field studies have been complied for insects exposed to pesticides
  - **Example on slide**

Fitness

- Fitness is a term used by geneticists to quantify an organism's ability to reproduce (often symbolized by w)
  - Each member of a population will have their own fitness value
    - 0 if it dies or fails to reproduce
    - 1 if it produces 1 offspring
    - 2 if it produces 2 offspring etc.
  - Average of all these values is symbolized with w (with line above it)
    - Populations of stable size will have it equal to 1
    - Population that is growing will have it greater than 1
    - Population that is shrinking will have it less than 1
- **Diagram on slide**

Relative Fitness

- E.g. a moth species with two pigmentations
  - A dark form (A_) which survives better in forest habitats
  - A light form (aa) which survives better in open fields
- In each enviro we arbitrarily define fitness of superior genotypes to be equal to 1
  - Inferior genotypes as deviation from 1
  - Fitness deviation symbolized by the letter s
  - S is called the selection coefficient – measures intensity of selection acting on the genotypes
- **Diagram on slide**

Biston Betularia

- Moth from Great Britain which inhabits wooded areas
- Commonly known as the peppered moth, it exists in two color forms, light (cc) and dark (C_)
- From 1850 on the dark form increased in frequency from 1 to 90% around heavily industrialized centers
    - Selection against the light form on the darkly polluted landscapes
- Recently the level of population has decreased and the light form of the moth has made a comeback

Directional Selection

- Genotypes of one or the other phenotypic extremes are selected
- Often used in plant and animal breeding
- In nature often occurs when one of the extreme phenotypes is selected against usually as a result of changes in the enviro
    - Medium ground finches in the Galapagos islands
    - Beak size varied enormously in the population
    - Periods of drought favored the survival of the large beaked birds since they had a greater variety of seeds to eat

H-W Law and Human Populations: CCR5 Gene

- Gene encodes a protein called C-C chemokine receptor-5
    - Chemokines are signaling molecules in immune system
    - Also a receptor for strains of HIV-1
        - Binds CCR5 and gains entry into cells
- Mutant form of CCR5 gene called CCR5 delta 32
    - 32 bp deletion in exon #4
    - HIV-1 cannot bind CCR5 delta 32 and these people are resistant to infection
- **Diagram on slide**
- Examined the percentage of Europeans with the CCR5 delta 32 mutation
    - CCR5 is a surface protein present on macrophages and T cells
    - HIV binds CD4 on macrophages and then CCR5 to gain entry
        - Individuals hetero for CCR5/CCR5 delta 32 are susceptible to HIV but sho prolonged progression into AIDS
        - Homo CCR5 delta 32's are resistant to infection by HIV (although 3 are known to be infected)
        - CCR5 delta 32 is only found in Europeans or their descendants

  - HIV over time invades the macrophages they then get into T-cells and wipe out immunity
    - AIDS
  - Hetero's have a harder time getting into the T cells

Origin of CCR5 Delta 32 in the European Population

- Not due to natural selection due to HIV
- The black death of 1346-1352 killed btwn a quarter to a third of all Europeans by bubonic plague
  - Produced a protein killing white blood cells – may act through CCR5
- Smallpox caused by the variola virus also use the CCR5 receptor for infecting cells
- Both the resistance to plague and small pox has led to the increased prevalence of CCR5 delta 32 in the European population

Stabilizing Selection

- Favors the intermediate phenotypes, with both extreme phenotypes being selected against
- Reduces population variance over time
- Great example is human birth weight
  - At the genetic level, individuals closer to the optimal birth weight will have a higher fitness
  - Infant mortality increases on either side of the optimal birth weight of 7.5 pounds (more drastic at low end)
- Why does normal birth weight drop on either side?
  - Complications during birth in the past (mother dies and then child due to lack of nutrients)
- **Diagram on slide**

Disruptive Selection

- Selection against the intermediates and for the extremes
  - Essentially opposite to stabilizing selections
- Not as common as the other two forms of selection
- An example might include a population of a hetero enviro
  - Darwin's finches with large and small beaks that can feed on different resources
  - Medium would be knocked out
- **Diagram on slide**

Migration and Gene Flow can Alter Gene Frequencies

- Migration is when individuals move btwn populations
  - If the 2 populations differ in their allelic frequencies, then this will alter the gene frequencies in the population
- Migration is essentially the flow of genes btwn populations – gene flow

Migration and Gene Flow

- At some point a population may become divided
  - Depending on the enviro, the allele frequencies may change in these two populations
- Migration is the movement of individuals btwn populations
  - As the migrants interbreed with the local population their alleles are now part of the gene pool
  - We now have to calculate the frequencies of the alleles in this situation
- Quickest way to change a population
- **Migration Example**

Gene Flow Example

- A DNA sequence polymorphism was identified called FY-NULL, that has two alleles, FY-NULL*1 and FY-NULL*2
  - Frequency of FY-NULL *1 is 0 in three African populations, and 1 in three European populations
- Find African-American populations derive 11.6-22.5% of their ancestry from Europeans and European-American populations derive 0.5-1.2% from Africans
- **Diagram on slide**

Migration Example: ABO Blood Group Locus

- The frequency of the I^B allele of the ABO blood group locus is present in a gradient from east to west
  - Highest frequency in central Asia and lowest in Spain
  - Forms a genetic footprint of human history
- Mongolians would invade places for years and have sex with the females
- North and south America = 2 different populations that migrated
- **Diagram on slide**

Random Genetic Drift

- Results from random events in small populations (sampling error)

- The number of reproducing individuals is too small to ensure that all alleles in the gene pool will be passed onto the next generation in their existing frequencies
    - Some alleles will be underrepresented and other alleles will be overrepresented in the next generation

__Genetic Drift__

- Coin flipping example
    - Chances of getting significantly more heads than tails in a small population, such as 4 out of 5, is much more likely than in a larger population such as 800 out of 1000
- Effective population size is the number of contributing gametes to the next generation
    - Includes # of breeding males and females
    - If one male contributes most of the gametes, his alleles will be present at a higher frequency in the next generation (nonrandom mating)

__Bottlenecks__

- Arise when populations expand from a small number of ancestors
- A bottleneck occurs when a population is either broken up into smaller groups or the whole population is decreased to a small population
    - Vulnerable to genetic drift
- Unequal allele frequencies that will get passed onto next generations
- **Diagram on slide**

__Bottleneck Example__

- Prairie chicken numbered in the millions in 1860
- Loss of habitat and other human activity caused a drastic decrease in their numbers
- Genetic variation in the prairie chicken was found to have decreased in the surviving population – genetic drift
- As the population decreases more and as the genetic variation decreases the species will become extinct

__Founder Effect__

- Genetic consequences when a human population originates from a small number of individuals
    - Population may increased in size later, but all of the genes in the population are derived from those of the original founders
- Commonly found in many human populations

- Examples – closed communities (Amish, Mennonites)

Causes of the Founder Effect

- Colonization of an area previously untouched by their species
- Passage through a bottleneck
- Physical barrier (i.e. islands)
- Cultural/religious isolation

Founder Effect Examples

- Albinism in the Navajo of the southwestern US (Arizona)
    - 1:1750 Navajo
    - 1:36000 Europeans
    - 1:10000 African-Americans
- Mutation found to be a deletion in the P gene which codes for a membrane protein
    - Mutation absent from the Apache
    - Navajo and Apache tribes separated btwn 600-1000 years ago
    - Mutation estimated molecularly to have arisen btwn 400-11000 years ago

Maple Syrup Urine Disease

- Autosomal recessive disorder
- Inability to metabolize branched chain AA (leucine, isoleucine and valine)
    - AA build up
    - Acidosis, seizures and urine that smalls like maple syrup
    - May lead to death
    - Complex therapy involved
- Found primarily in the Pennsylvania Dutch community

Ellis-Van Creveld Syndrome

- Autosomal recessive disorder in EVC gene
    - Normal function unknown
- Leads to short stature, extra fingers and toes, malformed wrist bones as well as abnormalities in teeth and fingernails
- Originated from small 200 German migrants – Amish