

Why and How to Create a cDNA Library

Using Fluorescence and Avidin-Biotin to Detect a Protein of Interest in a Tissue Sample

Genetic Contributions to Sporadic Disease: A Recessive Locus at 12q24 Commonly Contributes to Patient Ductus Arteriosus

Using Plasmids to Grow Protein

- Take mRNA of interest and make a cDNA copy of it.
- Derivatize the cDNA with Biotin so it can bind to an avidin-enzyme complex, where the enzyme product is easily detectable (it fluoresces, say).
- Wash the complexed cDNA over a tissue sample, add the avidin-enzyme complex, and see where you get fluorescence. These are areas where the mRNA of interest was expressed (and bound to the cDNA we added, which in turn bound to the avidin-enzyme complex).

- cDNA libraries allow us to analyze mRNA structure and abundance with the more stable and manipulatable DNA copy.
- cDNA allow us to determine the intron/extron positions on DNA by allowing hybridization.
- cDNA is created by using a 5' DNA primer to make a single cDNA strand complementary to the mRNA, then using a second primer (from the other end) to create a DNA version of the original mRNA strand.
- Place each cDNA into a library of vectors, and you have a library.

- Plasmids are generally several kilobases long, contain an origin of DNA replication, a resistance gene, and a useless section into which we can stick foreign DNA.
- Plasmids are usually capable of propagating on their own.
- Human DNA, DNA ligase, and a restriction enzyme are inserted, bacteria are plated on a selective plate, and only those bacteria that incorporated the plasmids with our DNA of interest (as well as the resistance gene) will survive.
- The surviving bacteria can then be used as a factory for growing our protein of interest.

- PDA: In fetus, a duct (the DA) btw pulmonary artery and aorta allows fetal blood to bypass the lungs. In PDA, this doesn't close at birth as it should.
- * One method of treating is to inhibit prostaglandins (which keep arteries open).
- The authors found much higher rates of PDA in Iran than the States, and specifically in families with consanguinity.
- By using areas of short-term tandem repeat markers we can compare the likelihood that two alleles inherited were inherited from a common ancestor. If many people affected with a certain disease have identical haplotype blocks, that section is likely linked to the disease.
- The sequences with the most conservation btw people are most likely to be linked to the gene of interest. After identifying sequences, use PCR and look for common mutations in the shared LD blocks.

Tissue culture and production of monoclonal antibodies

Synthesizing DNA and Site-Directed Mutagenesis

Southern Blotting

Yeast two-hybrid screening

- We can synthesize DNA of up to ~100 bp in length.
- A growing nucleotide chain is anchored to resin. We wash a bunch of nucleotides over, each which has a blocking group so only one NT is added.
- Wash off the excess nucleotides, remove the blocking groups, and repeat with the next nucleotide in the chain. Repeat until you have your full strand.
- Site-directed mutagenesis: If you want an A instead of a T at a certain point in a plasmid, synthesize a complementary chain with the replacement you're looking for. Allow the synthesized chain to bind to the original one, then use "tricks" of some kind to make the base mismatch repair processes resolve the discrepancy in favor of the new mutation.

- Tissue culturing can be used to study viruses and their life cycles.
- You can transfect DNA into cultured cells, integrating added DNA into the host genome. Difficult and unreliable.
- Can use cultured cells as factories to produce useful proteins, including specific antibodies.
- Production of monoclonal antibodies involves injecting antigen into mouse, isolating b-cells, fusing them with tumor cells to increase production, allowing them to grow and then testing for the presence of antibodies against the antigen of interest.

- Genetics approach to identifying which proteins bind to a protein of interest.
 - Uses two hybrid proteins. The first is a bait protein (of interest) fused to a DNA-binding domain, and the second is a bunch of cDNA-encoded proteins bound to a Transcription-activating domain.
 - Allow bait to bind to promoter of his3 gene, then wash various proteins fused to TAD's over it. Whichever binds to the bait protein will allow TAD to activate transcription of his.

- A DNA sample is treated with a specific restriction enzyme.
- Gel electrophoresis is used to separate DNA fragments of different sizes (large fragments migrate slower b/c they bump more).
- Transfer the separated pieces to a membrane by changing direction of electric field.
- Expose the membrane to a labeled hybridization probe.
- See whether your fragment of interest (mutation, say) is present in your original DNA sample.

Sequencing DNA

Mitosis

Transgenic and knock-in mice

Nomenclature

- Consists of production of identical daughter cells.
- Chromatids are dragged into daughter cells by their centromeres.
- Daughter cells are diploid with normal chromosome number of 46.

- Take template DNA of interest and add a synthetic primer (dunno how you know what primer to add...)
- Add all 4 NTs and a little bit of labeled "ender" ddATP onto which more NTs can't be added. You will end up with strings of nucleotides of varying lengths, each of which ends in an A.
- Repeat with sets of all 4 NTs and a little bit of each of ddTTP, ddGTP, and ddCTP, adding up to a huge number of DNA fragments of varying lengths, fluorescently labeled according to their final NT.
- Denature your template from the synthesized bits by heating, run them on a gel to separate chains by size, and scan for radioactivity.
- Look for the color of the shortest strand, second shortest strand, etc. The NT corresponding to the shortest strand's color is the first base, etc.
- You could also do this without differentially labeling the 4 "ender" NTs if you ran each set in a separate column next to each other.

- Chromosome is only condensed enough to be visible during prophase and metaphase.
- Chromosome is composed of two identical chromatids.
- Metacentric, submetacentric, and acrocentric refer to middle-centromere, offset, and one arm only being a satellite.
- Always specify the total number of chromosome and number and kind of sex chromosomes.
- Always name extra or missing chromosomes.

- Allow mice to mate, harvest fertilized eggs, inject eggs w/ cloned DNA and implant in foster mother.
- You can select embryonic stem cells to see which picked up the gene of interest by including resistance against neomycin.
- After culturing, implant those ES cells that survive into blastocyst and implant in foster mother.
- F1 generation will be chimeric (mosaic), F2 will either have the gene or not.

Dosage Control, X-Inactivation, and Y Chromosome Function

DNA Microarrays and ChIP Technology

Polymerase Chain Reaction

Meiosis

- Isolate mRNA from one or several cells of interest.
- Use that mRNA to make labeled target/probe (terminology changes from one source to the next) cDNA.
- * If you wanted to compare cancerous and healthy cells, for example, you could label one type green and one type red.
- Obtain microarray with genes of interest or the entire genome, broken up into fragments and stuck into individual wells.
- Wash your array with both sets of target/probes, then compare fluorescence to see which mRNA are expressed in both cells and which are only in one or the other.
- All you can tell by this method is which genes are or are not being transcribed, since you're looking at mRNA.
- ChIP is used to identify where on DNA a particular protein binds. "Fix" chromatin (DNA plus the protein of interest) tightly together, then use antibodies for the protein of interest (stuck to a bead), shake your DNA to break it into chunks, wash out all DNA not bound to

- Aneuploidy for any autosome disrupts development. To deal with the XY/XX conundrum, then, we have X-inactivation.
- In females, at about the 1000 cell mark of development, each cell randomly inactivates one of the two X's.
- Partial explanation for mechanism: Inactivation starts at the Xic location on one chromosome and spreads along it, probably by Xist RNA coating the entire inactive chromosome.
- Relatively high frequencies of X-aneuploidy survival, b/c of inactivation mechanism.
- Functions of Y chromosome are primarily male sex determination and spermatogenesis.
- Clonal inheritance from father to son implies that sequence changes simply accumulate with time (no recombination).

- Results in change of DNA content from 2N to 4N to 2N to N in 4 haploid daughter cells.
- Recombination occurs between homologs in meiotic prophase, with homologous segments exchanged among non-sister chromatids.
- 40-50 total crossovers, or about two per chromosome, during meiosis.
- 1st division separates homologous chromosomes (after recombination), second division separates sister chromatids.
- Spermatogenesis takes place over about 64 days, begins in puberty, and continues forever.
- Oogenesis pauses at prophase of the 1st division. In ovulation, the 1st div. continues, producing an oocyte and polar body. The 2nd div. occurs post-fertilization, and produces a 2nd polar body and an ovum, which fuses with a sperm to create a diploid zygote.

- Requires a DNA template that you want to multiply, 2 primers complementary to each end, taq or another heat-resistant DNA polymerase, and a bunch of all 4 dNTPs.
- As you raise and then lower the temperature, your resistant polymerase synthesizes complement chains to each of the two strands of your original DNA template, they "melt" and denature, and the cycle repeats.
- You can use mRNA and reverse transcriptase for RNA PCR.

Bayes' Theorem

Linkage, the Exception to Mendel's Law of Independent Segregation, and LOD

Aneuploidy and Non-disjunction

Structural Rearrangements

- Linkage is when a parental combination of traits occurs more often than expected among F2 generation.
- Genetic distance between genes A and B = number of recombinant gametes / total gametes * 100%, with 1 map unit (cM) as a recombinant frequency of 1%.
- Linkage analysis is an attempt to link a phenotype (disease) to a DNA marker whose position is known.
- Log Of the Odds is used to compare the heredity of a phenotype and a marker in a particular family and predict whether they're linked.
- LOD = \log_{10} (likelihood of data if two loci are linked by a certain distance) / (likelihood of data if they're not linked). LOD of 3 or higher is considered definite evidence of linkage, and the distance producing the highest LOD is taken as the map distance.

- Used to define more accurate probability of a parent having a certain genotype, given what we learn from her offspring.
- $\text{Prob}(C|E) = \text{Prob}(C)(\text{Prob}(E|C) / \text{TotalProb}(E))$.
- The probability that parent is carrier, say (C) given the offspring she's produced (E) is equal to the original probability she was a carrier (C, just from parents) * the probability she would have had the offspring she has, if she was (E|C) divided by the total probability of her having those offspring (E, which is E|C + E|notC).

- Structural rearrangements occur when chromosomes break and rejoin abnormally.
- Most common structural rearrangements are balanced translocations, which result in a normal amount of genetic material and normal phenotype. This involves translocation btw NONhomologous chromosomes.
- During meiosis, the homologous parts of the 4 involved chromosomes (2 translocated and 2 normal) align, and various segregations can occur.
- A Robertsonian translocation involves the production of a 45 cell b/c of translocation btw acrocentric chromosomes. Balanced translocations are harmless to the individual, but lead to unbalanced translocations in the 2nd generation.

- Aneuploid cells have either +1 or -1 chromosomes. The vast majority of -1 autosomes do not survive.
- Most aneuploidies result from nondisjunction at meiosis I, which results in two daughter cells with both parental chromatids, and two with none.
- Most common human aneuploidies at birth are autosomal and sex chromosome trisomies, b/c meiosis is suspended before the first disjunction occurs.
- Occasional nondisjunction during mitosis leads to mosaicism in the zygote.

Why Study Mendelian Disorders?

Types of Genetic Heterogeneity

Hardy-Weinberg Equation and its Assumptions

Types of Genetic Variation

- Allelic: Most disease genes have large numbers of disease-producing mutations. This produces a spectrum of functional consequences in the protein product of the gene, from variants with variants with zero function to those functional conditionally, to silent mutations. This contributes to the individuality of each patient.
- Locus: One phenotype can be caused by mutations in many different genes. The study of locus heterogeneity for a disease can give us insight into the number of gene products required for a particular system to function normally.

- Numbers are large
- Relatively straightforward
- Have high predictive value for identifying systems involved in particular disease. Identifying the protein coded for by a particular mutant gene and the system(s) in which it participates suggests other candidates for disorders with similar phenotypes, or other genes involved in the same complex trait.

- normal allele variants (no effect)
- conditional variants (normal function most of the time, but errors when "stressed" by the environment)
- disease variants, which include causative (Mendelian) and risk (altered susceptibility in complex traits)
- chromosomal disorders, which are actually quite common (1/200 live births)

- Uses Mendel's laws to determine allele and phenotype frequencies in a population.
- $p^2 + 2pq + q^2 = 1$ AND $p + q = 1$
- Assumes random mating (trait is not selected for or against during mating), uniform population, and no migration that CHANGES allele frequency.

The Vagaries of X-Linked Genes

Genotype/Phenotype Correlations

What Makes an Allele Dominant?

What Makes an Allele Recessive?

- Modifier genes alter the phenotype associated with mutations in another gene.
- An allelic series is a set of mutant alleles at a given locus that give different phenotypes.
- Poor correlation between genotype at the locus and phenotypic severity indicates that other variables (environmental, modifier genes) play a major role in the disease.

- Males' single X chromosome produces all of the proteins coded for by X's, so even a "recessive" allele will have a phenotype.
- Females are mosaics with some cells expressing the genes on the paternal X and some those of the maternal X.
- If alleles of one or more X-linked genes confer an advantage at the cellular level (even if not for the organism as a whole), there may be a selective advantage leading to a skewed pattern of X-linked inactivation.
- The tips of the p and q arms of the X contain sequences homologous to those on the Y chromosome, and escape X-inactivation in females.
- No father to son transmission!
- For lethal X-linked disorders, 1/3 of such males will be the result of new mutations.

- Usually passed from asymptomatic carrier parents to one or more children.
- May result from loss of function, often of an enzyme which can serve its purpose even if only 50% is present.
- Each of us are estimated to be carriers of 5-10 extremely deleterious (lethal) recessive alleles.
- Can be multiple recessive alleles at a locus, allowing for homoallelic or heteroallelic abnormal individuals.

- Usually doesn't skip generations (with exceptions).
- Dominant loss of function: A 50% reduction in functional protein product leads to a certain phenotype.
- Dominant negative: The mutant protein is not only nonfunctional, but interferes with the products of the normal allele--often seen in multimeric proteins.
- Dominant gain of function: The mutant protein gains excessive amounts of normal function, or gains a new function entirely.
- Among dominant alleles there can be variable expression (phenotypic severity), and penetrance (probability that there is ANY phenotypic expression).

Types of Variation

Linkage Equilibrium and Disequilibrium

Gene Definitions and Facts

Repetitive Elements

- When a mutation first occurs, it is "linked" to other SNP near it, forming a haplotype block. As the haplotype is transmitted, recombination reduces the size of the original unique haplotype block.
- There is a relatively large haplotype block in N. Europeans b/c of the population bottleneck from which they arose.
- Linkage equilibrium is finding a particular combo of alleles together in the frequency you'd predict from their general frequency. Disequilibrium is finding them together more often, which suggests linkage.

- Insertions and deletions
- Short tandem repeat polymorphisms (short runs of simple DNA sequences 1-4 bp in length. Similar to CNV, often used as markers.
- SNPs. Young SNPs are those that have arisen recently in human history, and will have a low frequency of the least common allele.
- Copy Number Variation involves large (>1kb) segments that vary in number from one to several copies. These show up again in anticipation diseases.

- About 45% of the genome consists of repetitive elements, which are useful because they can be used as markers.
- Short Interspersed Nuclear Elements, LINEs, retrovirus-like elements and DNA transposons are "molecular fossils" of past insertion events of transposable elements.
- Transposable elements move around the genome and contribute to its dynamic nature.
- SNPs are a common form of sequence variation, and can be used as markers.
- Copy number variants are segments of the genome that can be duplicated or deleted, and can be used as markers, though also can be involved in disease.

- About 22k genes (protein-coding) in humans, plus those for tRNA and microRNA.
- Euchromatin is loosely packed, gene rich, and frequently transcribed.
- Heterochromatin is tightly packed (dark stains), gene poor, and infrequently transcribed.
- Gene density varies greatly within and between chromosomes. The 3 chromosomes involved in viable trisomies (13, 18, 21) are both small and not dense.
- Homologs are genes w/ common ancestor, orthologs are equivalent genes in 2 species, and paralogs are homologous genes in 1 species arising from duplication.

Transmission-Disequilibrium Test

What to Do After Identifying a Region?

All Disorders are Complex

Genome-Wide Association Studies (GWAS)

- Look for genes within the identified area that may be strong candidates based on the biology of the disease and the protein encoded in the gene.
- Look for sequence variants that might account for a difference in the gene product.
- Use protein interaction studies (ChIP!) to look for other components of the implicated system.

- Similar to linkage analysis, TDT looks at trios of parents and their affected child(ren). If a particular marker is transmitted more than the predicted 50% of the time to affected children, it is likely either a susceptibility allele, or in disequilibrium with one.

- A dense set of SNP markers (500k-1mil) spread across the entire genome are genotyped in a large number of cases (people with the phenotype of interest) and controls.
- This approach allows you to find either the disease gene, or an SNP linked to the disease gene.
- One limitation is that you may only identify a large genomic region, which contains many possible causative genes.
- A second limitation is the danger involved in trying to match your control and case populations to avoid population stratification.
- Helps to identify more complex traits, where many genes may have a small effect. You do need to validate your results to make sure.

- Though we classify genes as monogenic, complex, and chromosomal, this is a simplification.
- Monogenic disorders are complex because they can have a spectrum of disease alleles.
- Variation in other genes can modify the severity, progression, and response to treatment of the genes.

Mitochondrial (Matroclinal) Disorders

Epigenetics

Anticipation

Mosaicism (not X-linked)

- "Above Genetics", epigenetics refers to heritable changes in gene expression that aren't encoded in the DNA sequence itself.
- DNA methylation is passed through replication by maintenance methylase, and is involved in "silencing" genes.
- Histone methylation, phosphorylation, etc. can be passed down, and is involved in regulating transcription.

- Can arise when mutations occur in either somatic or germ-line cells during early development.
- If it arises in somatic cells, only that person will express the mosaic phenotype.
- If it arises in germ line cells then a somatically normal parent may have multiple afflicted offspring, even though it initially looks as though it's a new mutation (meaning no chance the next child will have the disorder).

- Mitochondria have their own genome, which is a 16k double-stranded circular molecule encoding 13 proteins (involved in cellular respiration), 2 rRNA and some tRNA.
- Because the sperm leaves its mitochondria outside of the egg upon fertilization, offspring inherit only their mother's mitochondrial genome.
- Mutations in the mitochondrial genome can be either homoplasmic (all genomes in the cell have the same mutation) or heteroplasmic (mix of normal and mutant).
- There is often a progressive course, especially if the mutation is a deletion, b/c the mutant mitochondrial genomes will replicate faster than normal ones.

- Anticipation refers to the tendency of certain dominant or X-linked phenotypes to become more severe in successive generations.
- Can result from unstable repeat expansion mutations, where a gene contains a series of a variable number of short nucleotide repeats in one (of several) portions of the gene. In some families, the number of repeats increases dramatically, which can first form an allele with normal function but increased risk for additional expansion, and then form a disease phenotype allele.
- Responsible for some neurodegenerative disorders.

Allele Frequency

Pharmacogenetics

Imprinting

Uniparental Disomy

- Genetic variability affects the way people react to drugs, which has to be taken into consideration in drug design.
- Some people may inactivate drugs faster or slower, requiring carefully titrated dosages.
- Some people may experience adverse reactions to drugs b/c of changes in drug receptor or... something else.
- Pharmacogenetics is of interest to the industry b/c otherwise efficacious drugs will be dropped from development if even a small fraction of patients experience serious adverse reactions.

- For autosomal recessives, count the number of affected people and then use $p^2 + 2pq + q^2 = 1$
- For autosomal dominant mutations, count the fraction of people affected and divide by two (since it would be rare to have two)
- For X-linked mutations, assuming they're not lethal, simply count the fraction of affected males.
- For a dominant X-linked non-lethal mutation, you have to include females (and account for their two X chromosomes).

- Both members of a chromosome pair are from one parent.
- Isodisomy is two copies of the same chromosome, and leads to homozygosity for all those genes, exposing rare recessive alleles.
- Heterodisomy is both homologs from one parent.
- Can be caused by chromosome loss followed by duplication of the remaining homolog, or fertilization with aneuploid zygote followed by loss of one chromosome.

- A specific form of epigenetic modification where only one of two parental genes is expressed.
- The parent whose gene is expressed is maintained, but the gene silenced is reset each generation. Thus, a mutant allele could be passed down the paternal line without ever being expressed if that particular gene was always silenced in the chromosome inherited from mom.
- This can lead to trouble especially in uniparental disomy.

What Causes Cancers?

Cancer-Causing Genes

Blood Groups and Types

Types of Tumors

- Oncogenes: Normally stimulate normal growth.
- Suppressor Genes: Normally inhibit growth or stimulate apoptosis.
- Repair Genes: Normally limit mutations.

- Viruses
- Other infectious agents
- A breakdown in immunity (as seen in AIDS)
- Changes in gene expression
- Mutations in specific genes

- A tumor is any growth, period. This includes hernias and cysts (non-neoplastic tumors).
- Neoplastic tumors can be benign (polyps, adenomas) which means they can't move, or malignant (leukemias, carcinomas, sarcomas, etc.)
- A tumor is fundamentally a group of cells that have lost their 1:1 ratio of cell death to birth to something higher.

- A blood group is the set of red cell antigens controlled by a genetic locus with (basically) three alleles: A and B (codominant) and O.
- Serum from individuals with various alleles contain antibodies against any antigens their red cells don't contain.
- Giving RBCs with antigens for which a patient has antibodies leads to agglutination and possibly death.
- Rh+ people have antibodies against the antigen rhesus. If a mother is Rh- AND the father is Rh+ AND she has already given birth to one Rh+ child AND the mother and the first child were ABO compatible (so her antibodies didn't kill any of his RBCs immediately), she will have antibodies against Rh+ cells, and the 2nd kid's in trouble.

Non-Structural Proteins

Stages of One-Step Virus Growth

Two-Hit Hypothesis

Types of Viruses

- Eclipse: Very few infectious particles found in or outside of the cell (genome is incorporated into host)
- Intracellular accumulation: Newly-synthesized virions found inside but not outside.
- Rise period: Increasing numbers found outside cell as well.

- The virions (packages) of some viruses contain small amounts of other proteins, including RNA-dependent RNA polymerase (to produce viral mRNA), reverse transcriptase, integrases to help stick retroviral DNA into host genome, transcription factors, and proteases that process viral proteins and inhibit translation of host proteins.

- Helical vs. Icosahedral: Helical NA surrounded by a protein-shell capsid, icosahedral is surrounded by a roughly spherical protein shell.
- The multiple copies of identical structural units that comprise the capsid are genetically economical, b/c they require the smallest possible protein subunit for a given internal volume.
- Naked vs. Enveloped: Lipoprotein envelope surrounding the capsid is a modified version of the cytoplasmic or nuclear membrane, and may be involved in infection.

- Suggested by the 10k increase in incidence of familial form of a particular cancer.
- According to the two-hit hypothesis of cancer, both alleles need to be knocked out for tumor initiation. In the familial form, offspring start off life with one hit. There is a 10k higher chance of getting just one hit in a certain gene than two.

Phases of Infection (Poliovirus, Model RNA Virus)

Vaccination

Basic Infection Processes

Phases of Infection (naked icosahedral virus)

- Whole Virus Vaccines can be either killed (inactivated) or attenuated live.

* Killed viruses have lost their ability to reproduce and express genes, but still express their antigens. They generally impart immunity of short duration.

* Attenuated viruses can be administered orally, result in an inapparent infection and impart long-lasting immunity. They're generally cheaper, but there is the possibility of reversion to a pathogenic form, especially in immunocompromised hosts.

- Like most RNA viruses, it multiplies exclusively in the cytoplasm.

- Viral Genome has messenger polarity and can immediately begin synthesizing viral protein.

- One of the proteins synthesized is an RNA-dependent RNA polymerase, which synthesizes new virus genomes.

- Polyovirus proteins are synthesized as a giant strand which is then cleaved. The protease that cleaves it also interferes with capped cellular mRNA, "kidnapping" the host machinery for only its own proteins.

- Early Phase: Prior to DNA replication, this phase consists of transcribing mRNA to direct the synthesis of the T antigen and other

* The T antigen induces the host cell to replicate, initiates viral DNA replication, activates late mRNA synthesis and represses early mRNA synthesis, and inhibits apoptosis.

- Late Phase: Proteins transcribed during this phase include capsid proteins are made in much greater amounts, b/c they are packaged in all of the progeny virions.

- This temporal regulation is beneficial because delaying the production of virion proteins to the late phase allows the exponential accumulation of many copies of the viral genome during the early phase, which could not be accomplished if those copies were being packaged into virions.

- Attachment: Interactions btw proteins in virion surface and receptors on host cell surface.

- Endocytosis: Enveloped cells are taken into the cell through endocytosis, and then the acidic pH of the endosome triggers fusion of the viral envelope with the endo. membrane. Non-enveloped viruses may form a channel in the endosome and inject their DNA into the cytoplasm.

- Direct fusion: Some cells directly fuse their envelope membrane with the cytoplasmic membrane, delivering a naked capsid to the cytoplasm.

- Viral genome then either migrates to the cell nucleus and is incorporated in the host genome or remains in the cytoplasm and is transcribed there, forming viral proteins and new DNA, which is then packaged up, grabs some host membrane (and sticks in its own proteins), and travels on its merry way.

HIV Strategies for Resistance

Influenza

Retroviruses

Measuring Mutation of Retrovirus Genomes

- Influenza viruses infect humans and animals, with those affecting both of the most clinical importance. Each strain of flu contains a different collection of envelope glycoproteins, which come in two flavors: Hemagglutinin (more important antigenically) and Neuroaminidase (less important antigenically, but involved in preventing agglutination).
- The genome of the flu virus is segmented into negative sense s.s. RNA molecules, with each segment encoding 1 or 2 proteins. This allows reassortment between different strains infecting a single cell.
- After infection with a particular strain, we build up antibodies to that strain's antigens (HA & NA). As the strain mutates into new strains, our antibodies slowly become less effective against it.
- Vaccination against flu consists of inactivated (killed) viruses or partially purified HA protein. Short-lived immunity and poor mucosal response.

- To measure the rate of nucleotide substitution in retroviruses, we use a helper cell line (makes all the proteins required for virus replication and packaging, but not the virus itself), and transfect it with a DNA vector that encodes a defective virus.
- The defective virus carries the sequences needed for replication and packaging, but instead of viral genes has a mutated gene for neomycin resistance.
- After one round of replication, neo-resistant frequency is measured and mutation rate calculated.

- HIV contains genes for internal protein requirements, its polymerase, and its envelope proteins. The greatest genetic variability is seen in the envelope protein genes.
- Mutations that disrupt antibody binding to envelope proteins while still allowing attachment of HIV to host cell receptors have a selective advantage.
- Also, because CD4 contacts are to the alpha carbon backbone rather than A.A. side chains (as opposed to the antibodies, which bind those side chains), the virus can "hide" its side chains until after binding to CD4, shielding the envelope from neutralizing antibodies until it's too late.

- Retroviruses are positive-strand diploid viruses that replicate through a chromosomally integrated DNA intermediate.
- Retrovirus virions contain a reverse transcriptase and tRNA which prime DNA synthesis, and produce a double-stranded DNA copy of the genome while in the cytoplasm. This migrates to the nucleus, is integrated into the chromosome (with integrase), and then is transcribed and replicated.
- Virus is assembled at the cell membrane and released by shedding (non-cytolytic!)
- Variability is introduced during conversion of viral RNA to DNA (infidelity of reverse transcriptase) and transcription of DNA into RNA (infidelity of RNA polymerase II).
- Retroviruses also undergo recombination during reverse transcription, adding to variability.

Virus Tissue and Cell Susceptibility

Virus Variability

Viral Envelopes

Antigenic Drift and Shift

- Viruses with the most genetic variability are most able to evade host immune responses, elude vaccination-protection, and acquire drug-resistance.
- Viral variation can lead to changes in host range (what species it infects), virulence (amount needed to kill), tissue tropism (what tissues it infects), ability to elicit an immune response, ability to interact with host antibodies, and susceptibility to antiviral drugs.
- RNA viruses tend to have mutation rates of 1 per 10-1,00k nucleotides per replication cycle, which is 1000 to 100,000 times less than chromosomal genes. HIV has even higher rates of mutation. This is because viral RNA polymerases lack editing functions.
- Because the mutation rates of RNA viruses are so high, for a given virus the genome cannot be precisely defined.
- An isolate is a sample of viruses from a single individual, a strain is a collection of highly related genomes, and a genome is the sequence that comprises a single particle's genetic content.

- Viruses have both species and cell/tissue specificities.
- Cellular susceptibility is determined at several stages of the virus multiplication cycle, including whether appropriate receptors are expressed on the cell surface, whether the virus has access to those cells, and whether the cell contains the appropriate machinery for viral replication.

- Antigenic drift refers to the slow emergence of new antigenic variants as a result of mutation and selection. Though this year's antibodies won't be perfect against next year's virus, they will probably provide some protection.
- Antigenic shift is the result of recombination, usually between a human strain (with RNA capable of multiplying in human cells) with a previously animal-only strain (with HA and NA that our bodies aren't used to). Antigenic shift produces recombinants with a huge selective advantage in human populations, and is the reason we see occasional epidemics or pandemics in the human population. Reintroducing a variant that has not been seen for variations will produce the same basic results.

- Many animal viruses have a lipid bilayer envelope.
- The envelope of the influenza virus, for example, contains two integral glycoproteins, HA and NA. These proteins have internal and external domains that help create the envelope, and the external domains are involved in attachment to host cells and penetration.
- The external glycoproteins are the most sensitive targets of antibodies, and are species-specific (as far as viruses have species).

Outcome of Cellular Infection by Viruses (cont'd)

Virus (in)Activation of Common Cellular Pathways

Outcome of Cellular Infection by Viruses

Neutralization of Viruses by Antibodies

- Intrinsic cellular defenses: Cytocidal cell destruction is the most common mechanism by which viruses produce cellular and tissue pathology, and as cells lyse virus is released.
- * Necrosis is an unprogrammed, rapid cellular degeneration resulting in lysis of populations of cells in a tissue.
- * Apoptosis is programmed cell death in individual cells.
- * Autophagy occurs as part of normal development in many tissues, and involves a directed breakdown of cellular components.
- Viruses may also encode genes that specifically block apoptosis, since infected cells will often try to self-destruct to limit infection. Inhibition of the tumor suppressor p53 is one mechanism of blocking apoptosis. This allows the establishment of persistent infections and "virus factories."
- Viruses can subvert normal cellular transcription/translation through proteases that cleave host cell

- Cell susceptibility to virus is initially determined by the expression of cellular receptors for the virus.
- * Viral mutations can block virus's ability to attach to cells (as in attenuated poliovirus vaccine).
- Receptor expression is required but not sufficient for cell susceptibility.
- Interaction btw virus and cellular receptors is sufficient to alter host cell function (cellular co-receptors may be required, for example, for the virus to progress from binding to infection).
- * A naturally occurring deletion in the human CCR5 gene provides resistance to HIV for homozygous individuals, as infection is prevented.

- Viruses usually elicit the production of antibodies by the host.
- Virus neutralization is the decrease in infectious titer of a viral preparation following exposure to antibodies. Antibodies can interfere with any of the steps leading to release of viral genome in host cells.
- Neutralization can be reversible (low-affinity binding, requires a high ratio of antibody:virus b/c it depends on saturation of virus surface).
- Stable neutralization doesn't require saturation b/c the antibody molecule binds two sites on the virion and produces a conformation change in the capsid that prevents membrane fusion or uncoating.

- A cytotoxic infection rapidly kills the cell and produces a burst of new infectious diseases.
- * Acute infections are usually cleared w/i weeks because of strong immune responses to the virus.
- A persistent infection infects the cell and produces infectious virus without affecting the cell's viability.
- * Persistent infections are not cleared b/c of weak immune response or limited access to the site of infection by the immune system (as in the brain). Virus and viral proteins are continually produced, as in HIV and Hepatitis.
- A latent infection has no detectable viral products and no effect on the cell, but can later become cytotoxic.
- * Latent virus infections result when the virus is not cleared initially, and the virus exists either integrated or as an extrachromosomal element with occasional reactivation (and symptoms). Herpes, HIV and Hepatitis.
- Slow virus infection includes an initial acute infection followed by a period of low-level viral persistence, with eventual reactivation.

Innate Immune Response to Viral Infection

More Persistent and Latent Infection Stuff

- Virus persistence can be caused by poor immune responses or rapid virus mutation, infection of inaccessible tissues or cells of the immune system, or even expression of cellular proteins that block apoptosis (and transform a cytolytic infection into a persistent one).
- In a latent infection, little or no viral protein and viral nucleic acid is present, resembling lysogeny in bacteria. The viral genome can also be present as extrachromosomal DNA that either only forms virions at certain stages of cell differentiation or in response to stress (as in Herpes).
- In HIV, virus DNA is integrated into the host genome of bone marrow stem cells and only activated for transcription when the cells leave and begin to differentiate in the blood.
- In herpes virus, an initial infection usually occurs in childhood, the virus is taken up by adjacent nerve endings and migrates to the sensory ganglia.
- Latently infected cells express non-coding microRNA that suppress viral expression unless stressful events

- Innate immune responses are fast-acting, genetically hard-wired, and do not require prior exposure to the virus.
- Type I interferons are one of the mediators of innate responses to viruses. Virus binding to cellular receptors or exposure to (the normally unusual) d.s. RNA activates interferon synthesis.
- Interferons propagate an anti-viral state to nearby cells, which both inhibits viral infection and replication and makes cells better activators of the immune system.
- Interferons induce the synthesis of proteins that target RNA replication, transcription, d.s. RNA, etc.