

DESIGNER GENES 2007
Molecular Genetics, Biotechnology, and Population Genetics
Dr. Bill Wellnitz (wwellnit@aug.edu)
Augusta State University, Augusta GA

What is the event: see rules, but basically students are answering a variety of questions pertaining to various aspects of molecular biology and population genetics. These questions should be mainly problem solving in nature, but in some cases, expect the unfortunate overuse of vocabulary terms

What to expect: Ideally, the event should be run as stations, but many places will probably use a mere paper pencil set up, especially at regional. Again, an ideal situation will have at least 2x more questions than there are teams involved. Such a structure will drastically reduce ties from a grading standpoint. In a good event, only a few teams should be expected to answer all questions.

In any station event, students 1) must be able to split up the tasks, and 2) move quickly, with all belongings to next station, since in most cases they can not come back to a station. Expect some graphs and tables at most state events and certainly at national. **Spend a few minutes with students discussing/recognizing basic exponential functions (may be useful in many events, and on many AP exams) Significant figures and exponents are two crucial skills for people in almost all areas of science.**

Students may (must?) bring a non-programmable calculator, but nothing else.

What kind of students make a good team: Teams members should be those who can think logically and fairly quickly, and be good at solving all sorts of problems. Any one who can do well in various kinds of word problems, with a reasonable background, will probably do well. At regional, and even state level, consider a team of one upper class person and one lower class person. The latter may have been recently exposed to some of the material, whereas the former might be better at tying things together. Students must be able to work multi step problems.

How to prepare: Start with the basic hs bio text, and do the problems at the end of molecular biology chapters. Now repeat the same kind of activity using the appropriate chapters in an intro college bio text or the AP text. Pay particular attention to the chapter on biotechnology and “ethics.” A general genetics book may also be helpful, but many parts of these books are probably way too advanced for hs students.

I have included on the CD(largely because I couldn't attach all of the pages in the form I wanted) three chapters from a book I wrote for college genetics classes (*Genetics Problem Solving Guide, 2nd ed.*—the book is now out of print, but there are many helpful hints in it). Your students may find both the concepts/explanations and the problems helpful. Each problem has a very detailed answer. Different chapters of this book will be referred to below in various places. Most of the comments/questions etc in ch. 11 about bacterial control are beyond the scope of this event (there will however be some who chose to include these things)

Enough preaching. Now the nuts and bolts—the lubricating aspects can be found on the CD.

I. DNA structure (chapter 12)

A. Base composition: A always pairs with T, G always with C

A-T
C-G
G-C
T-A

Therefore, the amount of A in DNA will = amount of T and G = C. Both of these parameters must be met for DNA to double stranded. If both are not met, DNA is single stranded.

Purines: A,G; have two nitrogen containing rings

Pyrimidines: C, T, U; contain one nitrogen ring

An A-T pair is held together by 2 hydrogen bonds; A G-C pair is held together by 3 hydrogen bonds

Students may be expected to determine length of DNA molecule :
10 bases/ 10^{-9} m= 10^9 bases/ ?

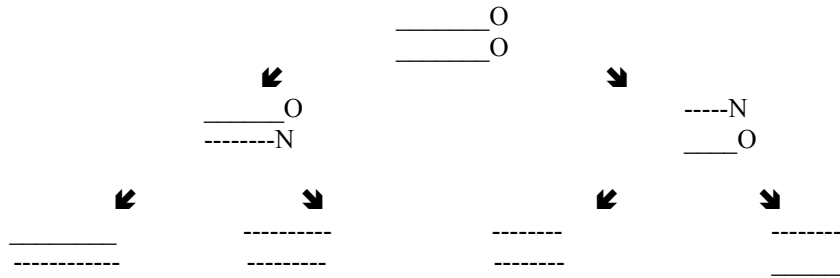
Other calculations may involve rate of transcription and/or translation.
Be sure students can make conversions between m and km etc.

- B. Each DNA strand has two different ends: P and OH (5' and 3'). The two strands run in opposite directions.
- C. **Melting temperatures:** as the temperature increases, the two strands of DNA begin to separate as H bonds are broken. The melting temperature is when half of bonds are broken. Since A-T pair has only two bonds, molecules with a high A-T ratio will have a melting temperature lower than a molecule with a high G-C content.

II. DNA replication (*text books*)

In most organisms process is **semi conservative**: strands separate and each strand is used as template for complementary strand. This result was initially discovered by Meselson and Stahl using Nitrogen that was heavy (N^{15}) and light (N^{14}). Having DNA made of either all heavy or all light produces molecules with slightly different densities.

- A. Result: each new molecule will have one old(O) and one new strand(N)



Questions:

1. If a DNA is labeled with a heavy isotope of nitrogen, and allowed to replicate in the presence of only light nitrogen, what fraction of the molecules will be of H/L composition after three rounds of duplication? (*1/4; 2 of the 8*)
2. If 28% of bases in DNA are A, what per cent is expected to be G?(*22*)
3. If A= 29, T=21, G=32, and C=18 %, what is nature of DNA (*single stranded*)
4. If it takes 15 minutes, using PCR, to make a copy of the molecule, how many copies will be present after 2 hours? (*see PCR below; ans 256*)

III. Polymerase chain reaction (*see <http://www.dnalc.org/ddnalc/resources/animations.html> for animations and descriptions of many molecular biology techniques; many are quite good*)

1. A procedure that allows one to make many copies of a DNA molecule in a test tube.

Note: A common misconception is that all the DNA an organism is replicated, but in fact it is usually about 200 nucleotides(Under extremely careful conditions it might be possible to get maybe 1000 bases). One typically uses certain primers as starting points so that a particular DNA, for example, a gene, or part of a gene, is replicated. This replicated DNA can be used to create a probe for certain DNA, say in forensics, or to construct some type of genetic hybrid.

2. Strands must be separated by an increase in temperature.
3. Sample is cooled slightly, primers are added along with a special DNA polymerase, and all four nucleotides.
4. Strands are replicated, temperature elevated to separate the newly made molecules, and the above procedure repeated. Changing temperatures is usually done by a machine called thermocycler or PCR machine. Many copies can be made in a few hours. (see animation listed above.)

IV. Transcription (*chapter 10*)- process of making RNA from DNA; usually only one strand of RNA is made from one of the two strands of DNA

A.

Differences between RNA and DNA

property	DNA	RNA
stranded	usually double stranded (some viruses use single stranded)	usually single stranded; some viruses use double stranded
Sugar	Deoxyribose	Ribose
Bases	A, G, C, thymine	A, G, C, uracil
location	usually nucleus (and mitochondria/chloroplast; bacterial chromosome (not in nucleus))	usually cytoplasm

B. Uracil replaces where Thymine would normally be.

Ex: If DNA is G G C T A A

C C G A T T* and bottom strand is used to make RNA, sequence of RNA will be:
G G C U A A

Practice: write out DNA sequences and then write an RNA from a given strand

V. **Translation** (*Texts, ch 10*)

- A. Process of making proteins from amino acids and messenger RNA; requires GTP instead of ATP for energy
- B. Bases read in groups of 3; 3 bases = 1 codon for 1 amino acid; see genetic code. It is always a good idea to set up the proportion :
- C. 3 bases/1 amino acid = # bases given/# amino acids given. # BASES WILL ALWAYS BE GREATER THAN # AMINO ACIDS
- D. Students should be able to use code to produce amino acid sequence
- E. If statement says the beginning of gene, must look for AUG, and then begin marking off groups of 3

Ex: ACGC AUG CCA UGC UUC ACGUAG

This message would make 5 amino acids; begin at AUG, and count successive groups of 3. UAG is stop codon and does not code for any amino acid.

F. If by chance the ends of DNA are indicated, remember that all nucleic acids run in opposite directions. So the 3' OH end of DNA will produce 5' P end of RNA. Same rules for DNA strands. Transcription begins at P end of RNA and proceeds to the OH end, just like we read left to right. Proteins also have two different ends: N (amino) and C (carboxyl). If ends are not indicated, assume that N end is on left and C on the right

Test: it is essential that students know how to use the genetic code that is any text. Most tests will have many questions on its use. The code should be provided. There is no way students should memorize the code. They must however know start and stop codons! Use of code will also be involved with mutations.

VI. **Mutations** (*text, ch 10*) will change one or more bases in DNA and hence in RNA. Some mutations will replace one amino acid, some will produce no change. Adding or removing a base will totally alter the message.

A. Kinds of mutations:

- 1. **point** (sometimes referred to as missense mutation): one base is replaced by another base. Depending upon location of this change, a) either one or no amino acids will be changed, b) a start or stop signal may be altered, resulting in no protein or a longer than normal protein, c) intron (see below) splice site may be altered, yielding larger, usually non-functional proteins
- 2. **frameshift**: one or more bases is added or deleted, making the wrong, and usually garbage codons after the mutation

B. Consequences of mutations in terms of dominance, etc. Many times, but not always, the change of one amino acid will usually produce a recessive trait. One possible exception is if the protein is part of a larger complex. In this case the amino acid change may produce a dominant trait Mutations in control regions are usually dominant

Practice: produce many different sequences of bases and predict amino acid sequence using the code.

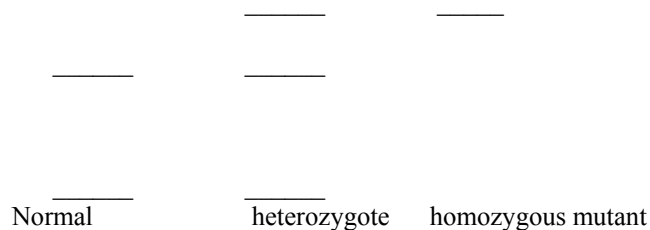
VII. DNA fingerprinting (ch 12)

1. DNA is isolated and then digested with **restriction enzyme**.
2. DNA fragments are separated in an electric field (**electrophoresis**)
3. Fragments stained or exposed to X-ray film to visualize.
4. Compare samples. Similarity/identity **suggests** the samples are related, as a perpetrator and victim. Differences **allows** for exclusion.

Short tandem repeats (STR)—a few bases repeated many times; number of repeats varies from individual to individual. Sizes of STR are used in various kinds of DNA fingerprinting

Restriction Fragment Length Polymorphisms (RFLP): different sized fragments produced, depending upon sequence with a gene. Originally a major player in fingerprinting. Now used often to detect individuals carrying certain mutations in specific genes

Example: A normal gene may have one site for an enzyme, and a mutation eliminates this site.



Test: Students may see a gel with many such patterns from many individuals in a family. They may then be asked with ones have the disease, which parent did the disease come from etc. Or if a scale is given on the STR (and it is often done), one may be asked who or how many have say 5 repeats.

DNA arrays: Filter paper or something similar has many single stranded pieces of DNA on it; each piece represents a single identified allele of the gene. When DNA is added, only an exact complement will produce color.

Now, DNA from an individual is cut with restriction enzyme, the strands are separated, and then mixed with the array. Only the section in the array that is *exactly complementary* to the test DNA will produce a color.

Test: A good question may show arrays that are mixed with DNA from different family members. Students are asked which children have the same DNA as say, their mother.

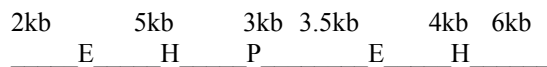
Example: The finger print below shows the DNA of a child (C), the child's mother (M) and 2 potential fathers (F1, F2.) Which individual is most likely the father of the child?



VIII. Restriction mapping and generating the map (ch. 12)

1. DNA is cut by a particular restriction enzyme, the fragments are separated as above, and the size of each fragment is determined (usually from a graph of known standards).
2. The above procedure is repeated with a different restriction enzyme. With the second enzyme, different size fragments will be produced.
3. A sample of the DNA is now digested with both enzymes simultaneously, and the above separation procedure performed.
4. To create a map, begin by drawing each digestion results separately.
5. Now look at combined digest. Determine which pieces from one digest remain after the double digest. These are pieces that do not contain a site for the second enzyme.
6. For those fragments that get smaller in the combined digest, we must now add up their sizes.

The map below shows the sites for restriction enzymes and the size of fragments produced. (kb =kilo base)



E= *EcoRI*
 H= *HindIII*
 P=*PstI*

List the fragment size in order from smallest to largest. Any other order will be counted wrong.
 What size fragments would be produced when the DNA is cut with *EcoRI*? (2kb, 10kb, 16.5)

What size fragments would be produced when the DNA is digested with *PstI* ? (10.5 kb, 13.5)

What size fragments would be produced when the DNA is cut with both *HindIII* and *PstI*.? (7kb,3 kb, 7.5kb, 6kb)

When a DNA molecule is digested with either enzyme A or enzyme B, two fragments are produced, but the fragments are different for each enzyme. If the molecule is digested with both enzymes, four fragments are produced. Is the DNA molecule circular or linear? (circular; number of fragments in mix = sum of individual digests; linear : number of fragments in mix is one less than sum of individual)

Match the following techniques with its use. Some answers may be used more than once; others not at all. Each use will have only one answer.

- A. Southern Blotting C. Western Blotting E. none of these
B. Northern Blotting D. Eastern Blotting

Transferring electrophoresed RNA to filters (B)

Transferring electrophoresed DNA to filters (A)

Determining if particular binding protein was inserted into the membrane (E; possibly C)

Transferring hybrid DNA to an enucleated egg. (E)

The gene for thinkase is isolated and found to be 6.3 kb long. Digestion of this DNA with *Eco RI* yields fragments of 4.1 and 2.2 kb; digestion with *BamHI* yields 3.0, 2.5, and 0.8 kb; the double digestion yields 0.8, 1.1, 1.4, and 3.0 kb.

A mutant of the gene is isolated and then digested with *BamHI* to yield fragments of 5.5 and 0.8 kb. What did the mutation do and where is the mutation located?

- A. Added restriction site at 3.0/2.5 junction B. Added restriction site at 2.5/0.8 junction
C. Eliminated restriction site at 3.0/2.5 junction * D. Eliminated restriction site at 2.5/0.8 junction

IX. DNA sequencing (ch. 12)

1. DNA plus enzymes plus all four bases plus one modified base (usually dideoxy). When modified base is added, that particular strand can not be replicated further. Incorporation of modified base is random. Hence, many different size fragments are produced.

2. Fragments are separated by electrophoresis

3. Read sequence up from bottom of gel. Bottom gel will begin with the 5' end.

Note: strand sequence read from gel will be complimentary to original.

What enzyme is used to covalently link two or more pieces of DNA together? (*DNA ligase*)

The following strand is used as the template strand for the dideoxy DNA sequencing procedure.

5' C A T C C A T C A G C A A G C A C A T T 3'

If the reaction tube contains all four radioactive nucleotides plus di-deoxy T, how many fragments will be produced? ((7—sizes 3, 5,8, 9,12,15,19)

A DNA sequence of about how many bases can be routinely synthesized by automated machines?

- A. 10 C. 10,000
B. 100 * D. 1,000,000

X. Gene control (ch. 11)

1. Much on the control is at the level of transcription. Over 3-D shape is important. Adding methyl groups or some other group to DNA or proteins bound to DNA may change their shape and prevent or allow more binding to DNA. Similar statements can be made about the addition of groups to DNA. Some may ask many more vocabulary-type questions about what is the role of this protein or that protein.

2. For eukaryotic genes, the initial, or primary transcript, must be processed by having introns removed, the ends are often modified, and then leave the nucleus. Each area can also result in another level of control.

3. Some RNA molecules will be sequestered in the cytoplasm by binding to one or more proteins. Translation depends upon adding more proteins (and hence changing shape) or removing some of the proteins.

4. Rate of translation will also depend upon proteins, overall shape, as well as sequence of control regions of the RNA. Number of times a message is translated may depend upon sequences at one or both ends.

5. **telomeres**—sequences that are added to ends of chromosomes seem to control how many times a given DNA molecule can replicate. Once the last telomere is lost, gene expression may also be lost. Such a phenomenon may be involved in aging, and suggests why it is better to use young (embryonic) cells for cloning.

Control of transcription:

1. Turning off transcription—one or more small molecules will interact with one or more proteins . This interaction cause shape changes in the proteins and they now can bind to the promoter region and prevent transcription
2. Turning on transcription—one or more proteins (*repressors*) bound to the promoter, preventing RNA polymerase binding, and preventing transcription, Small molecules will bind to repressors, changing their shape so they can no longer bind DNA. Polymerase can now bind and make RNA.

Summary of types of gene control

Type of system	Normal situation	Result of “environmental“ change	Example
Repressible	Usually on	Production of specific RNA is turned off; no protein made	genes for amino acid synthesis in bacteria
Inducible	Usually off	Production of specific RNA is turned on; specific protein is made	genes for sugars other than glucose metabolism muscle specific proteins
Constitutive	Always on	RNA and protein always made but quantity may change	Proteins involve din energy production

Enhancers—sequences of DNA bases that are very far (often 1000s of bases) from promoter. The DNA molecule seems to bend so that enhancer sequence is close to promoter and almost “pushes RNA polymerase along.

Eukaryotic RNA has modified bases (*cap*) on the 5’ end, usually poly A on 3’ end. Length of poly may determine number of times the message is translated.

By removing and splicing different introns, a cell can make more than one protein from a gene sequence.

Xa. Recombinant DNA (*see chapter 12*): (also known as genetic engineering, hybrid DNA) combining DNA from two or more sources to create one new molecule. This hybrid molecule is then exposed to certain kinds of cells in hopes of changing the genetic makeup of the cells. Cells that have acquired the new DNA are identified, often by resistance to an antibiotic. Hybrid molecule must be able to replicate in the cells that have acquired it.

Procedure:

1. Desired gene is isolated (often not an easy task); procedure requires at least one restriction enzyme and some other things.
2. Vector molecule needed; this molecule has many different poperties: restriction sites for many different enzymes, gene for selection, sequence for replication.
3. The two DNAs are mixed together in test tube, and joined together with *DNA ligase*.
4. Hybrid molecules are mixed with cell, possibly through injection.
5. Cells that acquired new DNA identified (resistance, color, etc). Not all cells will acquire the new DNA.

Uses:

1. Gene therapy
2. Production of mass quantities of useful protein, eg, human insulin
3. Creating different crops
4. Creating agents that may be used in biological warfare.

Problems

1. Not all cells acquire the new DNA, and often those that do are very rare.

2. Acquisition may be only temporary.
3. Acquisition may alter other parts of the physiology of the cell.
4. Who decides which genes should be transferred.
5. May be cost prohibitive
- 6.

Genetic Modified Organisms (GMO): organisms that have some of their DNA from a different species.

1. Various plants have been modified for resistance to herbicides, in theory, allowing for more crops, produced cheaply.
2. “ “ “ “ “ to certain insect pests.

Problems:

1. Some people may have allergy to added product, and mixed food may be produced.
2. Pollen from GMO can travel distances to other plants.

XI. Mitochondria DNA and inheritance: all mitochondrial DNA comes from mother as the mitochondria of sperm do not enter the egg at fertilization. **Thus, the affected mother will pass trait onto ALL progeny whereas affected father will pas on trait to NONE of his children.** The fingerprint pattern is sometimes used to determine family heritage.

DNA molecule it self is circular, has no introns, and no histones (all supporting evidence that mitochondria arose from bacteria—I could envision a question that asks for some kind of support of this idea.)

Human mitochondria:

37 genes, only 13 code for proteins; others code for tRNAs and rRNA

Mutations lead to degenerative disorders, often involving muscles (ex: Leber optical atrophy)

XII. Population Genetics: concerned with how frequent various alleles are in population

Equilibrium—implies that allele frequencies remain same from one generation to the next. If not at equilibrium, something is happening to change the frequency: in breeding, selected survival, migration, etc.

p = frequency of dominant allele; q = frequency of recessive allele

$p + q = 1$

Ex: 18 AA 52Aa 30 aa

$p = [\text{number of hom dom} + 1/2(\text{hets})]/\text{total number of individuals}$

$p = (18 + 1/2(52))/100 = 0.44$ $q = 1 - 0.44 = 0.56$

Or count alleles $2 \times 18 = 36$

$1 \times 52 = 52$

$88/200 = 0.44 = p$

If at equilibrium, $p^2(\text{AA})$, $2pq(\text{Aa})$, $q^2(\text{aa})$

Expect 0.1936 AA, 0.5028Aa, 0.3036 aa very close in this case

If 4/100 show the homozygous recessive trait, what are allele frequencies?

$q^2 = 0.04$; $q = \sqrt{0.04} = 0.2$; $p = 0.8$,

Other handouts available thru email by contacting me:

1. GA. Regional designer genes questions, 2004 or 2005 or 2006 (All tests also on Wright Center page)
2. GA state cell biology questions, 2004 or 2005 (good for AP biol classes; Cell biology will return as an olympiad event, probably 2008, but certainly 2009)
- 2a. GA 2005 Life Science Process (B) skills event; good for intro boil and provides a nice review for students
3. Cell biology handouts from previous workshops, both in Georgia and the national workshop.
4. DNA “concept map”—various roles played by DNA etc; good for introductory biology class, nice review for advanced students
5. DNA finger print simulation exercise—good for class activity; has about 20 different “sequences” and another “sequence obtained from a crime scene or pregnancy test. Helps students understand DNA finger printing. Good for any level bio class.