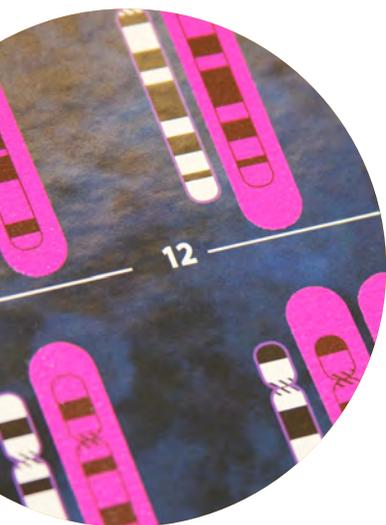


Biotechnology Discoveries and Applications

Extensions to high school science curriculum in Alabama



The 2010 guidebook

 **HUDSONALPHA**
INSTITUTE FOR BIOTECHNOLOGY



Biotechnology Discoveries and Applications

2010

How this guide is arranged

Recent research findings are grouped on the first eight pages and provide a quick update on the genetics/genomics/biotechnology field. **A new addition to the guidebook, this section represents discoveries, treatments or applications that have been announced during the past year.** Some are described in only a few sentences while others get a more thorough explanation.

Each new finding connects to one of twenty-three key technologies or concepts described in detail on subsequent pages. Language and concepts are intentionally geared to a high school or public audience.

Within each overview, linking course of study objectives are identified for Alabama High School Courses:

Look for the  symbol in blue.

Where relevant, the experiments and activities developed by HudsonAlpha are also described:

These are identified by the  symbol in orange.

SUMMER 2011 OPPORTUNITY FOR ALABAMA LIFE SCIENCE TEACHERS

Genetic Technologies for Alabama Classrooms (GTAC) *a two-week teacher academy*

July 10-22

Applications will be accepted
December 15-February 1

www.hudsonalpha.org/education



The HudsonAlpha Institute for Biotechnology invites high school educators to apply for GTAC, an intensive two-week genetics and biotechnology academy July 10-22, held in Huntsville. The academy is designed to help Alabama educators more effectively teach genetics by updating content knowledge, identifying common student misconceptions and gaining familiarity with hands-on genetic activities and classroom tools. Participants will interact with researchers studying genetic influences in cancer, neurological disorders and infectious diseases, and tour biotechnology companies dedicated to improving diagnosis and treatment of human disease.

Educators chosen to attend the academy will receive room, board and a stipend as well as a collection of materials and resources related to genetics and biotechnology. This is an excellent opportunity for a life science teacher that seeks to update an understanding of genetic concepts and their application to healthcare, agriculture and energy.



“GTAC has rejuvenated me and refreshed my excitement to be in the classroom. It has also awakened my love of genetics and opened my eyes to the wonderful things going on in our state.” - 2010 GTAC Participant

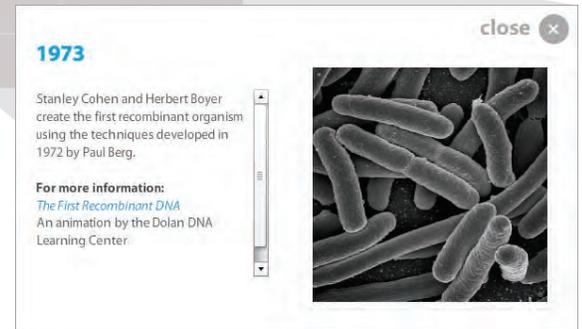
THE PROGRESS OF SCIENCE

Interactive timeline highlights history of biotechnology



The Progress of Science is made possible through the generous support of Wesfam Restaurants, Inc./ Burger King

“The Progress of Science” is an online timeline that details over 200 major accomplishments and milestones in genetics and biotechnology during the past 10,000 years. The digital timeline is an interactive navigation tool that offers details on each major event and links out to other online resources where available. The timeline is frequently updated, keeping the content current for classroom discovery. The Progress of Science can be accessed at <http://timeline.hudsonalpha.org>.



HUDSONALPHA iCELL™

Online program lets students explore cells inside and out

The advent of interactive whiteboards, classroom computer access, online courses and an increasingly technology savvy student and teacher population has increased the use of digital education resources across the nation. Alabama serves as a model with regards to the ACCESS program that links all high schools in a digital and distance format. HudsonAlpha is actively involved in creating three digital education products to be used in classrooms across the state.

HudsonAlpha iCell™ is a three-dimensional, interactive representation of an animal, plant or bacterial cell complete with organelles. Users can rotate the cell in 3-D space, mouse-over any component to display a text description of that part and double-click to zoom and re-center. Users can also chose from among multiple levels of text complexity to explain each part of the cell. Exploring HudsonAlpha iCell™ reinforces key concepts on cell structure and function.

The activity is available at www.hudsonalpha.org/education/digitaleducation/icell.

The program runs completely online, meaning it is compatible with any interactive whiteboard technology or projection system and nearly all personal computers made within the last eight years. Additionally, a free downloadable version is available for the Apple iPhone, iPad and iPod Touch.



CONTENTS

I. Executive Summary	7
II. New Findings	8
III. Course of Study Connections	16
IV. Foundational Concepts and their Applications	21
A. Key Technologies	22
1. DNA Sequencing	22
2. RNA and Protein Analyses	23
3. Bioinformatics	24
B. Applications	25
1. Agriculture	25
2. Cancer	26
3. Comparative Genomics	27
4. Copy Number Variation	28
5. Criminal Justice and Forensics	29
6. Diagnosing Chromosome Disorders	30
7. Epigenetics	31
8. Genetic Information Nondiscrimination Act	32
9. Genetics of Eye Color	33
10. Identifying Genetic Influence on Disease	34
11. Infectious Disease	35
12. Noninvasive Prenatal Diagnosis	36
13. Personal Genome Analysis	37
14. Personalized Medicine	38
15. Pharmacogenomics	39
16. Recombinant DNA and Genetic Engineering	42
17. Stem cells	43
18. Studying the Genome to Understand the Sequence	44
19. Synthetic Biology	45
20. Therapeutic Approaches	46
V. Statutes and Session Law	47
VI. Recommended Reading For More Details	48

EXECUTIVE SUMMARY

Genetics and biotechnology have dramatically expanded during the last twenty-five years, powered by our understanding of DNA and how fragments of genetic code may be analyzed, combined and exploited. These fields have overlapped and spread across diverse disciplines such as human health, agriculture, defense and space exploration.

The completion of the Human Genome Project marked a major benchmark in genetic research. Armed with this data, other large-scale projects strive to understand the function of the DNA sequence and its contribution to human disease. These findings have identified genetic factors that increase disease risk, unlocking pathways for drug development and future treatment.

From an agricultural standpoint, biotechnology has incorporated nonplant genes which confer resistance to pests, herbicides or drought, creating over 50 different genetically modified plants. At the same time, the genomes of soybean, corn, loblolly pine, poplar and switchgrass have been or are being sequenced for their potential contribution to bioenergy.

Career opportunities in genetics and biotechnology, often grouped under the umbrella heading of **bioscience**, are growing rapidly. The U.S. Department of Labor notes that the number of jobs in scientific research and development services is projected to grow at more than twice the rate of other jobs, with strongest demand in biotechnology and similar life sciences research. Four of the ten fastest growing U.S. jobs from 2008-2018 are in bioscience or information technology utilized by bioscience.

Students with a strong STEM foundation compose the competitive pool of workers needed for today's knowledge-based economy. A scientifically-literate workforce is particularly critical in a field such as biotechnology, which combines concepts from multiple branches of science. Using our life science expertise and educational collaborations, the HudsonAlpha Institute for Biotechnology works to excite students about science and inspire them to seek careers in biotechnology and related fields. Through collaboration with the Alabama State Department of Education, HudsonAlpha has



A not-for-profit research organization, HudsonAlpha is located in a 270,000 square foot building in Huntsville, Alabama, the cornerstone of a planned 150-acre biotechnology campus. Founded in 2005, the institute is a joint venture between private philanthropy and support from the state of Alabama. HudsonAlpha aims to harness the power of biotechnology to improve human health, stimulate economic growth and inspire youth to seek careers in the field of science through educational outreach.

developed several hands-on, inquiry-based activities for middle and high school students. Over the last three years, we have also hosted in-class and online discussions, provided summer camps and internship opportunities, and offered professional development sessions for educators (see page 4 for information about the 2011 GTAC summer educator academy).

This guidebook, the third annual update, is an easy to use resource that helps life science educators keep students on the cutting edge of the

field. Introduced in this edition is an eight-page “New Findings” section that summarizes many of the major discoveries from the past twelve months. Each finding is connected to one of twenty-three topics described in detail beginning on page 22. All topics are linked to course of study objectives for science, health and career and technical classes.

For quick reference, the relevant course objectives and the linking content topic are listed in table format on pages 16-20. A list of suggested readings for further exploration can be found on page 48.



In Brief

Infectious Disease

Antibiotic treatment increases bacterial mutation rate

Recent research has shown that treating bacteria with low concentrations of antibiotics increases the mutation rate in the bacterial cells. Many bactericidal antibiotics work by disrupting the bacterial cell's DNA, proteins and lipids. Researchers treated *Escherichia coli* bacteria with low doses of three antibiotics: ampicillin, kanamycin and norfloxacin. With all three treatments the mutation rate of the bacterial cells increased up to eight-fold. Similar results were observed with a different *E. coli* strain as well as *Staphylococcus aureus*. These findings provide important insight into the increasing problem of drug-resistant bacteria and suggest that antibiotics themselves could be a key player in this resistance.

REFERENCE: Kohanski, M.A., DePristo, M.A. & Collins, J.J. Sublethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis. *Molecular Cell* 37, 311-320 (2010)

NDM-1 gene creates resistant "superbug"

A new class of antibiotic resistant bacteria has begun spreading across the globe, thanks to a gene called NDM-1. The gene codes for an enzyme that neutralizes the activity of most antibiotics. This is particularly worrisome as the enzyme makes bacteria resistant to the carbapenem antibiotics, which are the so-called "antibiotics of last resort" currently used when other antibiotics have failed. The NDM-1 gene can move from one bacterial strain to another through a process known as horizontal gene transfer, potentially leading to a host of resistant infections. Currently, there are no drugs in the development pipelines to combat NDM-1.

REFERENCE: Roberts, M. New 'superbug' found in UK hospitals. *BBC News* 11 August, 2010 <http://www.bbc.co.uk/news>

Synthetic Biology

The first "man-made" genome

This spring, researchers announced the creation of the first bacterial cell with a genome designed by a computer and made entirely in the lab. The scientists synthesized the 1.08 million basepair genome from *Mycoplasma mycoides*, a bacteria that often infects goats, but is not dangerous to humans. The genome was created by synthesizing DNA fragments from bottles of chemicals in the laboratory, assembling fragments into an entire genome and inserting it into a host *Mycoplasma* cell whose own DNA was destroyed. When the host cells grew and divided, they did so under the direction of the synthesized genome.

For most Americans, this was their first introduction to synthetic biology, a field that seeks to apply engineering principles to biology (see page 45 for more details). The announcement of the first synthetic organism quickly prompted a Congressional hearing examining the implications of synthetic biology. Supporters of the field believe it has the potential to synthetically create new drugs, environmental biosensors and biologically-based fuel sources.

REFERENCES: Gibson, D. G., Glass, J. I., Lartigue, C., et al. Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science* 329(5987):38-9. 2010
Khali, A. Collins J.J. Synthetic biology: applications come of age. *Nature Reviews Genetics* 11, 367-379 (2010)

DNA Sequencing Leaps Ahead

Advances in 'next generation' sequencing technology



Over the past five years, there has been a shift from utilizing traditional Sanger sequencing methods for genome analysis. A number of new approaches (referred to as 'next-' or 'second-generation sequencing') have revolutionized this technology for research and clinical settings.

In addition to standard DNA sequencing, scientists can isolate RNA from a population of cells, reverse-transcribe the RNA to DNA and sequence the fragments, creating a catalog of the RNA fragments present in the original cells. Second-generation sequencing can also identify the epigenetic tags that modify gene expression and chromatin structure.

The second generation technologies have resulted in dramatic reductions in

Population Genomics

Gaining insight into population histories worldwide

Many of the first published human genome sequences were from individuals of European descent. Several recent studies have aimed to learn more about human diversity by sequencing and genotyping individuals from varying ethnic backgrounds.

Genome sequencing was performed on two individuals from South Africa - a Bushman and a Bantu speaker. The

Bushman was found to be as genetically different from other African groups as from Europeans. Additionally, a large number of new single nucleotide polymorphisms (SNPs) not previously described in other populations were found, reinforcing the need to expand beyond European genome sequences to better understand the whole of human genetic variation.

Another genomic study focused on the genomic history of the Jewish population. Genotyping studies found surprising genetic similarity between Jewish groups across the globe, but could still distinguish the specific groups created by the Diaspora (dispersion of the Jewish population). In addition, the studies found greater than



sequencing costs (see the figure on pages 14-15). As a result, sequencing the entire genome of an individual for clinical decision making is on the verge of entering medical practice. During the last year, several researchers have successfully sequenced the genomes of entire families affected with Mendelian disorders and atypical forms of cancer to find the causative mutations. In some cases, this knowledge has dramatically influenced treatment decisions at the patient's bedside. Similar studies have begun to identify the genetic modifiers that influence disease severity for disorders such as Huntington Disease.

It is estimated that up to 1,000 individuals may have their entire genomes sequenced by the end of 2010. Other approaches focus on sequencing only

expected genetic similarity with other non-Jewish groups such as Italians. The information gained from these studies provides unique insight into the timing of major milestones in Jewish history.

The Tibetan highlands lie 13,000 feet above sea level. Low levels of oxygen at that altitude lead to symptoms of mountain sickness (an increased risk of stroke, hypertension and pregnancy complications) caused by an overproduction of red blood cells. Genomic studies are beginning to suggest why Tibetans only rarely experience these symptoms. The genomes of Tibetans and the nearby dwelling Han Chinese, who live at a lower altitude, were compared. A number of genes - several associated with lower levels of red blood cells -

the "exome" - the exon and known regulatory regions of the genome. This approach provides results faster and more cheaply than sequencing the entire genome, but may not capture all the sequences important in gene expression. Amazingly, "third generation" platforms, based on sequencing a single-molecule on a semiconductor chip or a quantum-dot nanocrystal, are currently in development. These machines have the potential to yet again dramatically reshape the sequencing field and its applications.

For more information, refer to the background article "DNA Sequencing" on page 22.

REFERENCE

Metzker, M. Sequencing Technologies - the next generation. *Nature Reviews Genetics* 11, 31-46 (2010)

were identified which have undergone DNA changes since these populations diverged several thousand years ago. These changes are possibly the fastest examples of human genetic adaptation to date.

For more information, refer to "Studying the Genome to Understand the Sequence" on page 44.

REFERENCES

Simonson, T.S., et al. Genetic Evidence for High-Altitude Adaptation in Tibet. *Science* 329, 71-75 (2010)
 Katnelson A. Jews worldwide share genetic ties. *Nature News* 3 June 2010 www.nature.com/news
 Oster, S. Tibetans Had 'Fastest Genetic Change Ever' in Humans. *The Wall Street Journal* 2 July 2010 www.wsj.com
 Skipper, M. Into Africa. *Nature Reviews Genetics* 11, 170-171 (2010)

Cancer

New research findings; new options for treatment

Recent developments in genetic sequencing technology have made it possible to sequence the genome of tumor cells. Tumor cells accumulate many genetic changes, some of which play a key role in the cells' transformation into cancer. Genome sequencing has been carried out on both melanoma and small-cell carcinoma. Many mutations were identified in both cancers, some of which provided clues as to their cause. Not surprisingly, mutations from the melanoma cells revealed UV-related DNA damage while small-cell lung carcinoma mutations suggest DNA damage caused by tobacco smoke.

In a similar study, researchers sequenced the genome of three cell types from a single cancer patient. Samples were taken from the patient's blood, her primary breast cancer tumor and a later brain metastasis. Comparing the genomes allowed the researchers to identify genetic changes that occurred during the process of metastasis. While many genetic changes were common to both cancer samples, several differences were seen between the primary tumor and subsequent metastasis, providing insight into the progression from normal to tumor to metastasis.

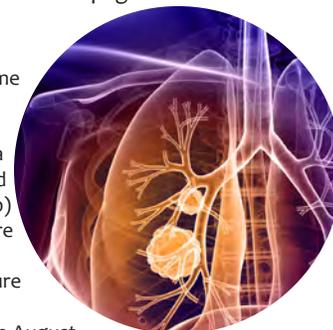
In addition to providing new understanding into the cause and progression of cancer, genetics also provides clues into the development and selection of cancer treatments. Therapies are now being developed that target the specific genetic fingerprints found in some cancers. A new drug has been developed for lung cancer that works very well in patients with a specific genetic variation in the ALK gene. While only about 4% of all lung cancers in the US have this genetic variant, the drug helps identify a potential new therapy pathway that can be modified for other cancer treatments.

The American Cancer Society and Livestrong determined that cancer costs the world's economy nearly \$900 billion a year in lost productivity. This is a greater figure than the toll from any other cause of death, including infectious disease and cardiovascular disorder. The study found that there were 83 million years of "healthy life" lost due to death and disability from cancer. Lung, bronchial and tracheal cancers have the highest overall cost (nearly \$180 million together), followed by colorectal and breast cancers. Note that these costs do not include those associated with direct medical care to treat the cancer, which the authors estimate would further increase the overall total.

For more information, refer to "Cancer" on page 26.

REFERENCES

Winslow, R. & Loftus, P. Advances Come in War on Cancer. *The Wall Street Journal* 7 June 2010 www.wsj.com
 Ding, L., et al. Genome remodeling in a basal-like breast cancer metastasis and xenograft. *Nature* 464, 999-1005 (2010)
 Gray, J. Genomics of metastasis. *Nature* 464, 989-990 (2010)
 Skipper, M. Genomic archeology. *Nature Reviews Genetics* 11, 94-95 (2010)
 American Cancer Society Press Release August 16, 2010: The American Cancer Society and LIVESTRONG® Launch First Global Economic Cost of Cancer Report <http://pressroom.cancer.org/>



NEW FINDINGS

Gene Therapy

A prescription for genes

A recent gene therapy trial was completed with twelve patients with Leber congenital amaurosis (LCA), an inherited retinal degeneration disease caused by mutations in the RPE65 gene. After the gene therapy treatment, all twelve patients experienced improvement in disease symptoms.

Another gene therapy trial is currently planned for SCID-X1 (x-linked severe combined immunodeficiency). This is an important step, as SCID-X1 trials were put on hold by the FDA several years ago when previously treated patients experienced increased rates of leukemia. Many changes have been made to SCID-X1 gene therapy protocols over the last few years to reduce the risk for leukemia and make the therapy safer and more effective.

While gene therapy research has historically focused on single-gene diseases, researchers at the University of Alabama in Birmingham are currently studying the use of gene therapy for treating multifactorial conditions such as cancer. One of the foci of the UAB Gene Therapy Center is determining methods for modifying the viral vector to be more safe and efficient. UAB has collaborated with other institutions for recent and continuing clinical trials of cancer gene therapy, with promising results.

For more information, refer to the background article Therapeutic Approaches on page 46.”

REFERENCES

Maguire, A.M., et al. Age-dependent effects of RPE65 gene therapy for Leber's congenital amaurosis: A phase 1 dose-escalation trial. *The Lancet* 374, 1597-1605 (2009)
Marcus, A.D. Study Shows Hope for Gene Therapy. *The Wall Street Journal* 7 July 2010 www.wsj.com

Interpreting an entire genome

Sifting for clinical clues among a mountain of genomic sequence

“Over 3,000 hours was required for analysis of this patient's sequence.”

In the future, an individual's genetic information will be a valuable clinical tool to guide medical management and treatment. A recent article in *The Lancet* addresses the challenge of how to begin incorporating a patient's genomic information into medical practice.

The genome of a 40-year-old male was sequenced and interpreted by a team of more than 35 scientists, physicians, and genetic counselors. The patient was healthy, but had a family history of coronary artery disease and sudden death. The patient's 148

gigabytes of genetic sequence were analyzed to identify areas of variation. Analysis revealed 2.6 million single nucleotide polymorphisms and 752 copy number variations.

The team undertook the enormous task of determining which of these variants had an impact on the patient's health by searching online genomic databases and the medical literature. Over 3,000 hours was required for analysis of this patient's sequence. The researchers identified more than 90 variants of potential clinical interest. This includes 69 drug-related variants that may be helpful as the patient's physicians select the type and dose of drugs to prescribe.

This study emphasizes the interrelatedness of diseases and genetic variants. The patient's sequence

suggested a decreased risk for hypertension, but an increased risk for obesity, which itself is a predisposing factor for hypertension.

There are many challenges to overcome before whole genome analysis is ready for clinical use. Obviously 3,000 hours of analysis is impractical under any healthcare setting and technologies must be developed to automate the interpretation process.

For more information, refer to the background articles “Personal Genome Analysis” and “Personalized Medicine”, on pages 37 & 38.

REFERENCE

Ashley, E. A. et al. Clinical assessment incorporating a personal genome. *Lancet* 375, 1525-1535 (2010)

Agriculture

From field and stream to the kitchen table

The genomes of several plants have been published this year including apple, corn, and wheat. Knowing the genome sequences of these organisms aids in understanding plant characteristics and developing genetically modified organisms with enhanced traits. For example, it may be possible to produce apples with higher levels of antioxidants or soybean crops with increased oil production for use in alternative fuels. In addition to soybean, other organisms are currently being genetically engineered as potential sources for biofuel, including algae, *Escherichia coli*, and eucalyptus trees.

The FDA is currently discussing the approval of the first genetically modified (GM) animal created for human consumption. GM Atlantic salmon contain a growth

hormone gene from a Chinook salmon that is linked to a genetic on-switch (a promoter sequence) from a distant relative of the salmon known as an ocean pout.

Salmon do not typically make growth hormone in cold weather, limiting their growth to only warmer months. The on-switch from the pout maintains hormone production year round. As a result, the salmon grows to market size in 16 to 18 months instead of 36. These large fish are unable to reproduce, preventing the GM version from interbreeding with the wild salmon population. The outcome of this FDA meeting will be one to watch as it sets the precedent for the future approval of other GM animals.

In other news, GM sugar beets were recently banned

on the grounds that their environmental impact was not adequately assessed prior to FDA approval. GM sugar beets comprise about 95% of all sugar beets planted in the U.S. and contribute to half of the sugar supply. The sugar beets are modified to include a gene that makes them resistant to Roundup™ weed killer, in the same way that many corn, soybeans, and cotton crops have been modified.

Although the FDA approves many genetically modified organisms (GMO), there is concern among some groups regarding the long-term impact of widespread GMO use. An industry group, called the Non-GMO Project, recently initiated a campaign for the labeling of foods as GMO or non-GMO.

Dog Genomes Surprisingly Simple

Many physical traits are controlled by very few genes

Most physical traits (like hair and eye color, body shape and height) are thought to spring from complex genetic interactions - the combined effects of several genetic and environmental factors. While research in both humans and plants confirms this view, recent studies in dogs suggests that many canine physical features result from only a small number of genes. Using DNA from more than 80 registered dog breeds, plus wolves, jackals, coyotes and feral African "village dogs" the researchers examined over 60,000 single nucleotide polymorphisms (SNPs) in order to associate specific genes with physical traits like body size and ear type. Body size seems to be controlled by six genetic regions while ear shape (erect or floppy) is linked to a single gene. Similar results were found for other traits.

The presence of a few genes that each exert large effects is most likely the result of population bottlenecks, both when dogs were first domesticated from wolves and later during the Victorian Era when most modern dog breeds were developed. A single mutation that produced a novel result (say short limbs or curly hair) was a prized finding and often maintained in purebred form. This form of selection simplified the genetic impact of many of the traits we today observe across individual breeds of dog.

REFERENCES

Bhanoo S. Wide Variety of Breeds Born of Few Genes. *New York Times* 16 August 2010
 Boyko A. et al. A Simple Genetic Architecture Underlies Morphological Variation in Dogs. *PLoS Biology* 8(8):e10000451 (2010)

One concern of the groups opposed to GM crops is that weeds may become resistance to Roundup™, much like bacteria develop antibiotic resistance. This issue is already a reality in some areas of the U.S. To date 10 resistant species have been identified in at least 22 states, causing farmers to find alternative and potentially more expensive, laborious, and toxic methods for dealing with the weeds.

For more information, refer to the background article "Agriculture." on page 25.

REFERENCES

Neuman, W. & Pollack, A. Farmers Cope With Roundup-Resistant Weeds. *The New York Times* 3 May 2010 www.nytimes.com
 Neuman, W. 'Non-GMO' Seal Identifies Foods Mostly Biotech-Free. *The New York Times* 29 August 2009 www.nytimes.com
 Schreck, S. GM Salmon Up For Debate. *Food Safety News* 30 August 2010 www.foodsafetynews.com
 Soybean Genome Sequenced: Analysis Reveals Pathways for Improving Biodiesel, Disease Resistance, and Reducing Waste Runoff *ScienceDaily* 14 January 2010 www.sciencedaily.com

In Brief

Studying the Genome to Understand the Sequence

1000 Genome Pilot Project

The 1000 genomes project has released data from its first three pilot projects. The ultimate goal of the 1000 genomes project is to build a detailed database of human genetic variation that helps researchers understand human diversity and genetic contribution to human disease. To achieve this goal, the project will sequence the genomes of 2,500 people from 27 populations worldwide. The pilot projects determined the most effective method for collecting and distributing the sequence data. The pilot data, including the whole genome sequences of 185 individuals and whole exome sequences (protein coding regions) of 700 individuals, is freely and publicly available on the web for use by researchers across the globe. The 1000 genome project, when complete, will produce an extraordinary amount of data. The pilot project data alone encompasses nearly eight trillion basepairs of DNA - more than 50,000 gigabytes of space.

REFERENCE: Wood, L, et al. 1000 Genomes Project Releases Data from Pilot Projects on Path to Providing Database for 2,500 Human Genomes. 1000 Genomes Project 21 June 2010 www.1000genomes.org

Coding and Noncoding Sequence Across Species

Most of the human genome consists of non-protein-coding DNA. Thanks to projects like ENCODE (see page 44) progress has been made in identifying the functional relevance of these regions, including transcriptional regulators and small regulatory RNAs. Similar studies are underway in other organisms. Below are the current estimates of genome classification for a number of model organisms.

Organism	Genome Size (Mb)	Percentage of DNA by category			
		Genic	Exonic	introns	intergenic
Yeast	12.2	73.5	72.9	0.6	26.6
Nemotode	100.3	59.2	28.1	31.2	40.8
Fruitfly	168.7	48.2	18.3	30.0	51.8
Human	3,107	45.1	2.8	42.3	54.9

REFERENCE: Alexander R.P. et al. Annotating non-coding regions of the genome. *Nature Reviews Genetics* 11:559-71 (2010)

Genetic Information Nondiscrimination Act

Claim of genetic discrimination tests new law

In 2009, the Genetic Information Nondiscrimination Act (GINA), went into effect, prohibiting employment discrimination based on genetic information. The first test of GINA has been raised, pitting a natural gas and electricity provider against an employee who claims she was fired after revealing that she carried a high-risk breast cancer mutation. The employee filed a complaint with the Equal Employment Opportunity Center. Based on a family history, the employee had tested positive for mutations in the BRCA2 gene, placing her at a high risk of developing cancer. She opted to undergo a double mastectomy and voluntarily revealed the reason behind the surgery to her employer. Although she had received positive performance reviews prior to her surgery, she was demoted upon her return to work. After a follow-up surgery several months later, she was fired.

REFERENCE: Vorhaus, D. GINA in Action: Woman alleges genetic test led to firing. *Genomics Law Report* www.genomicslawreport.com



NEW FINDINGS

Personal Genomics

Challenges surface for direct-to-consumer genetic testing and related projects



Personal genome analysis can be a powerful educational tool to help people understand complex genetics and interact with their genetic information in a real and “personal” way. The University of California – Berkeley announced in summer 2010 that they would be offering genetic testing

for all incoming freshmen and transfer students. The program originally planned to test students for three common genetic variants, supply students with their personal test results, and discuss the results as part of a campus wide discussion on personalized medicine. However, with input from the California Department of Public Health, the program will instead provide students with aggregate class data instead of personal test results.

Many in the personal genome industry hope that providing people with an individualized genetic risk for common disease will motivate these individuals to take steps to decrease the environmental risks. The Coriell Personalized Medicine Collaborative is studying this possibility, trying to determine whether

knowing genetic information improves health outcomes. Program participants are given their personalized disease risk information and asked to provide annual updates about how knowledge of their disease risks has impacted health behavior.

“consumers may make well-intentioned but actually harmful lifestyle or health-related choices”

The direct to consumer (DTC) testing company 23andMe recently published a research study using information and test results from their customer base. Genome wide association studies (GWAS) were performed to identify

genes linked to twenty-two human traits such as freckling, hair curl, and the ability to smell urinary metabolites of asparagus (asparagus anosmia). While these traits have little if any clinical significance, they illustrate the potential power of customer data contributing to genomic research.

There are significant concerns however, regarding findings from DTC companies. The genetic disease risks reported to consumers can vary based on which company performs the analysis. In a recent study, researchers sent samples from five individuals to two DTC companies and compared the disease risks reported. The companies agreed on risks for some conditions, but differed remarkably for others, largely because each company includes different genetic markers in their risk

King Tut's DNA

Using genetic testing to unravel an ancient family pedigree

Genetic testing on several mummies buried in Egypt's Valley of the Kings has allowed Egyptian scientists to unravel the tangled family tree of Tutankhamun, the boy-king who ruled Egypt from 1333-1323 BC.

Tutankhamun became pharaoh during a turbulent period – Akhenaten, the previous King, had overturned the entire religious foundation of Egypt, abandoning the ancestral gods to worship a single deity known as Aten. This practice was rejected early in Tutankhamun's reign in favor of the previous beliefs. After ruling only ten years, Tutankhamun died without any surviving heirs. He was buried in a small tomb in the Valley of the Kings, the traditional resting place for Egyptian kings and nobility. His successors,

perhaps in an attempt to remove reminders of the heresy, destroyed nearly all evidence of Akhenaten and Tutankhamun. The tomb of Tutankhamun remained hidden until its discovery in 1922. While thousands of artifacts were found that shed light on the life and times of this ruler, little was known about his family relationships including parentage, spouse and potential children.

DNA was carefully extracted from Tutankhamun and ten other mummies suspected to be members of his immediate family. To minimize the risk of obtaining DNA contaminated from previous archeologists or the Egyptian priests who had performed the mummification, cells were extracted from deep within the bones. The extraction process was challenging

due to the embalming substances used during the mummification process. Once DNA was obtained, eight genetic markers (short sequences of DNA that repeat in varying numbers) were analyzed to create a genetic fingerprint for each mummy, in much the same way genetic information is used today to solve a crime or for identification purposes.

The genetic information was used to identify family relationships among the mummies. A previously unidentified mummy was confirmed to be King Akhenaten, the father of Tutankhamun. Additional mummies were identified as Tutankhamun's mother and his probable wife. The DNA fingerprints also revealed that Tutankhamun's parents were siblings. Marriage among

close relatives was a common occurrence among Egyptian royalty, but such relationships increased the likelihood that recessive mutations would be passed to a child. Medical imaging suggest Tutankhamun was born with a clubbed foot and partial cleft palate, both potentially the result of genetic defects.

Tutankhamun's mummy was also tested for traces of infectious disease. DNA from the parasite *Plasmodium falciparum* was found, indicating malarial infection at the time of death. An unhealed broken leg was discovered in an earlier analysis. It has been speculated the malaria may have weakened his immune system, leaving him susceptible to complications from the leg injury and resulting in his early death.

In Brief

calculation. The authors of this study argue that DTC companies need to inform consumers of the actual fraction of overall disease risk that their test provides. Often this is a very small amount rather than the full risk for disease. If not fully aware of this gap in risk information, consumers may make well-intentioned but actually harmful lifestyle or health-related choices.

Since the beginning of the DTC testing industry, the FDA has not been involved in approving or regulating personal genome analyses. However, in July 2010 the FDA held hearings to explore overseeing these types of tests. This discussion is far from over. Many in favor of regulation believe that because the testing offered by DTC companies has medical implications it should be

regulated as a medical test. Others fear that increased regulation will hinder scientific innovation and an individual's access to personal genetic information.

For more details, refer to "Personal Genome Analysis" on page 37.

REFERENCES

Coriell Personalized Medicine Collaborative (CPMC) www.cpmc.coriell.org
Eriksson, N., et al. Web-based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits *PLOS Genetics* 6, 1-20 (2010)
Katsnelson, A. Consumer gene testing in the hotseat. *Nature News* 29 July 2010 www.nature.com/news
Ng, P.C., et al. An agenda for personalized medicine. *Nature* 461, 724-726 (2009)
Sanders, R. UC Berkeley alters DNA testing program. *UC Berkeley* 12 August 2010 www.berkeley.edu/news

DNA Sequencing*Exploring microbial populations*

Bacteria and archaea live in all environments capable of sustaining life (and in some extreme environments are the only forms of life). There are an estimated 5×10^{30} of these organisms present on earth. Metagenomics – the thorough study of the genomes of uncultured microorganisms – is helping researchers better understand their function and influence on animals, crops and humans. The first microbial genome was sequenced in 1995; thanks to advances in sequencing technology there are now more than 1,200 such genomes stored in the Genbank genetic database. The Human Microbiome Project seeks to characterize and understand the way microbes contribute to human health and disease. Microbes on the skin, gut and elsewhere outnumber human cells by about 10 to 1. Differences in the bacterial populations present in the digestive system have recently been linked to obesity, metabolic syndrome and inflammatory bowel disease.

REFERENCES: Wooley JC, Godzik A, Friedberg I *A Primer on Metagenomics. PLoS Comput Biol* 6(2): e1000667. (2010)
Flintoft L. *Human microbiome: A gut feeling for disease. Nature Reviews Genetics* 11, 237 (2010)

Comparative Genomics*Neandertal draft genome published*

Researchers published the draft genome sequence of the Neandertal, using DNA collected from the bones of three Neandertal females who lived in Croatia more than 38,000 years ago. Neandertals are estimated to have separated from the ancestors of modern humans between 270,000 and 440,000 years ago, but co-existed with modern Humans initially in the Middle East and later in Europe. When the genomes were analyzed and compared, several interesting findings emerged. Between 1-4% of the genomes from individuals of European and Asian ancestry are Neandertal in origin, yet Africans have no Neandertal segments. These findings suggest that a low level of interbreeding occurred between Neandertal and modern day humans, but only with populations that had migrated out of Africa. In addition, the protein coding sections of Neandertal and Human genes contain only 78 nucleotide changes that impact the structure of protein. Some of these changes are in genes implicated in energy metabolism, head shape and cognitive development.

REFERENCE: Green, R. et al. *A Draft Sequence of the Neandertal Genome. Science* 328, 710 (2010)

Human FOXP2 gene placed in mice

The FOXP2 gene is believed to play an important role in speech. Subtle changes in the gene sequence between humans and other animals is thought to have contributed to the ability of humans to coordinate the motions of tongue, lips and voice boxes to produce the spoken word. Researchers in Germany have created a strain of mice containing the human version of FOXP2. In the basal ganglia (a portion of the brain involved in human language) the nerve cells of the mice grew in a more complex structure. In addition, while all baby mice produce ultrasonic whistles if separated from their mothers, the mice containing the human FOXP2 gene uttered whistles with subtle, but significant differences. While FOXP2 is not the only gene required for speech, this comparative study lends support to its role as an important contributor.

REFERENCE: Enard W. et al. *A Humanized Version of Fxop2 Affects CorticooBasal Ganglia Circuits in Mice. Cell* 137(5), 961-971 (2009)



For additional details about this type of genetic testing, please see "Criminal Justice and Forensics" on page 29.

REFERENCES

Hawass Z. King Tut's Family Secrets. *National Geographic* 2010. <http://ngm.nationalgeographic.com/2010/09/tut-dna/hawass-text>
Hawass Z, et al. Ancestry and Pathology in King Tutankhamun's Family. *JAMA* 303 (7): 638-647 (2010)
JAMA and Archives Journals (2010, February 17). Study examines family lineage of King Tut, his possible cause of death. *ScienceDaily* www.sciencedaily.com.



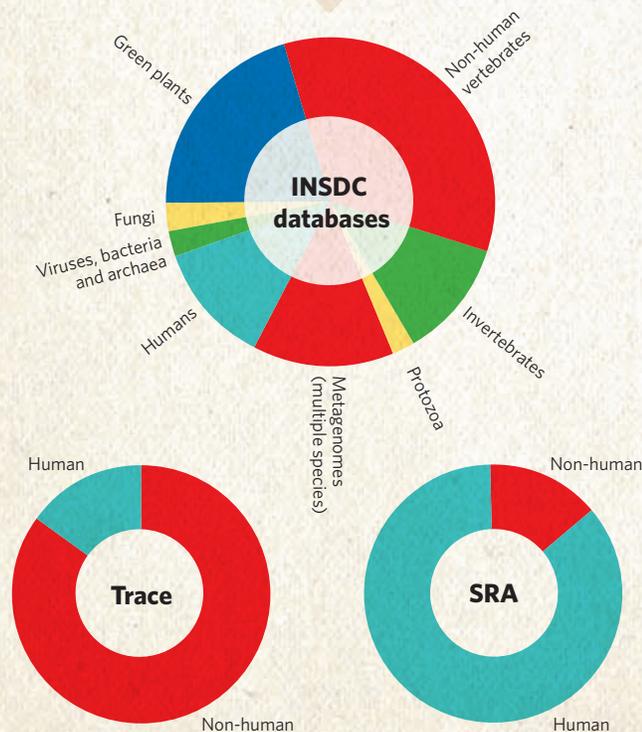
THE SEQUENCE EXPLOSION

At the time of the announcement of the first drafts of the human genome in 2000, there were 8 billion base pairs of sequence in the three main databases for 'finished' sequence: GenBank, run by the US National Center for Biotechnology Information; the DNA Databank of Japan; and the European Molecular Biology Laboratory (EMBL) Nucleotide Sequence Database. The databases share their data regularly as part of the International Nucleotide Sequence Database Collaboration (INSDC). In the subsequent first post-genome decade, they have added another 270 billion bases to the collection of finished sequence, doubling the size of the database roughly every 18 months. But this number is dwarfed by the amount of raw sequence that has been created and stored by researchers around the world in the Trace archive and Sequence Read Archive (SRA).

See Editorial, page 649, and human genome special at www.nature.com/humangenome

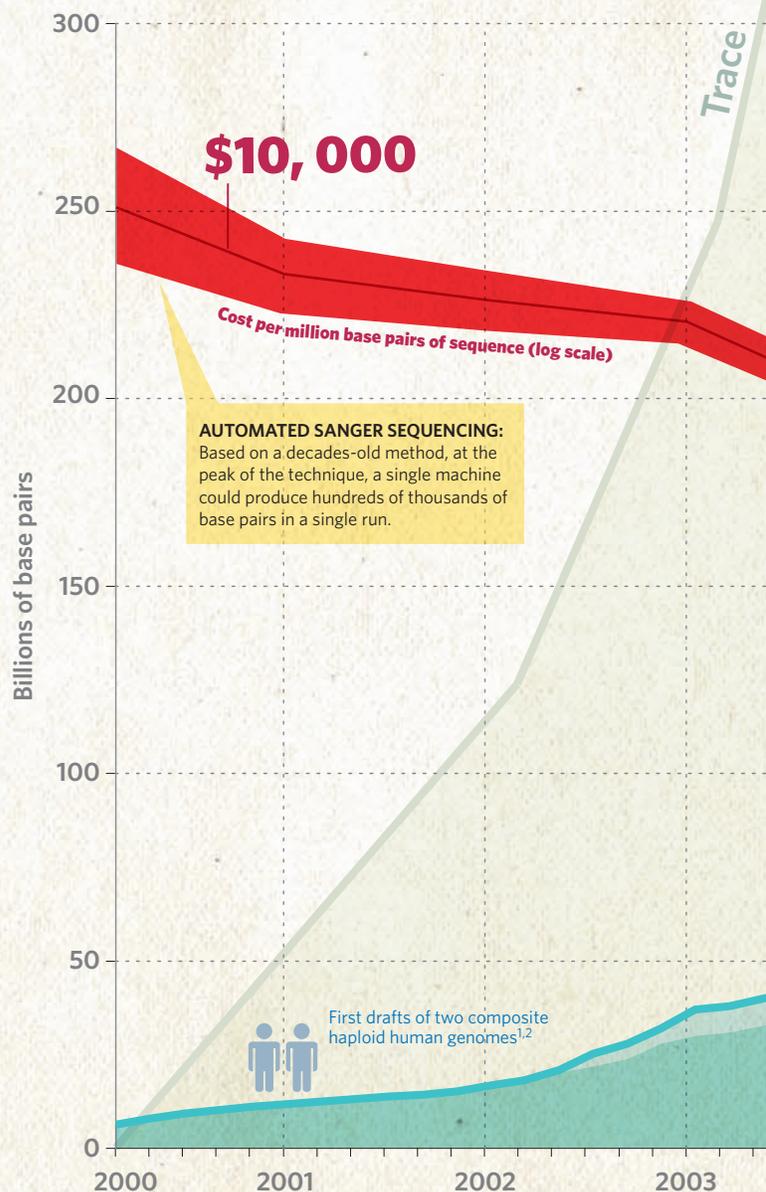
DNA SEQUENCES BY TAXONOMY

International Nucleotide Sequence Database Collaboration: The main repositories of 'finished' sequence span a wide range of organisms, representing the many priorities of scientists worldwide.



Trace Archive: Developed to house the raw output of high-throughput sequencers built in the late 1990s, the trace archive spans a wide range of taxa.

Sequence Read Archive: Houses raw data from next-generation sequencers. Dominated by human sequence, including multiple coverage for more than 170 people.



HOW MANY HUMAN GENOMES?

The graphic shows all published, fully sequenced human genomes since 2000, including nine from the first quarter of 2010. Some are resequencing efforts on the same person and the list does not include unpublished completed genomes.

COURSE OF STUDY CONNECTED TO GUIDEBOOK TOPICS

Objective and Applicable Subheading

Linking Scientific Concept

Course	Objective and Applicable Subheading	Linking Scientific Concept
<p>Biology</p> <p>2 Describe cell processes necessary for achieving homeostasis, including active and passive transport, osmosis, diffusion, exocytosis, and endocytosis.</p> <p>Identifying functions of carbohydrates, lipids, proteins, and nucleic acids in cellular activities</p> <p>4 Describe similarities and differences of cell organelles, using diagrams and tables.</p> <p>Identifying scientists who contributed to cell theory</p> <p>5 Identifying cells, tissues, organs, organ systems, organisms, populations, communities, and ecosystem as levels of organization in the biosphere.</p> <p>Recognizing that cells differentiate to perform specific functions</p> <p>6 Describe the roles of mitotic and meiotic divisions during reproduction, growth, and repair cells.</p> <p>Comparing sperm and egg formation in terms of ploidy</p> <p>7 Apply Mendel's law to determine phenotypic and genotypic probabilities of offspring.</p> <p>Defining important genetic terms, including dihybrid cross, monohybrid cross, phenotype, genotype, homozygous, heterozygous, dominant trait, recessive trait, incomplete dominance, codominance, and allele</p> <p>Interpreting inheritance patterns shown in graphs and charts</p> <p>8 Identify the structure and function of DNA, RNA and Protein.</p> <p>Explaining relationships among DNA, genes and chromosomes</p> <p>Listing significant contributions of biotechnology to society, including agricultural and medical practices</p> <p>Relating normal patterns of genetic inheritance to genetic variation</p> <p>Relating ways change, mutagens and genetic engineering increase diversity</p>	<p>RNA and Protein Analysis</p> <p>See HudsonAlpha Cell (pg 5)</p> <p>Stem Cells, See also Biotechnology Timeline (pg 5)</p> <p>Comparative Genomics, RNA and Protein Analysis, Stem Cells</p> <p>Cancer, Stem Cells</p> <p>Diagnosing Chromosomal Disorders, Noninvasive Prenatal Diagnosis</p> <p>Genetics of Eye Color</p> <p>Epigenetics</p> <p>Cancer</p> <p>RNA and Protein Analysis, Recombinant DNA and Genetic Engineering, Therapeutic Approaches</p> <p>Diagnosing Chromosome Disorders, Noninvasive Prenatal Diagnosis, Recombinant DNA and Genetic Engineering, Studying the Genome to Understand the Sequence</p> <p>Agricultural Applications, Cancer, DNA sequencing, Genetic Information Nondiscrimination Act, Noninvasive Prenatal Diagnosis, Personal Genomic Analysis, Personalized Medicine, Pharmacogenomics, Recombinant DNA and Genetic Engineering, RNA and Protein Analysis, Stem Cells, Synthetic Biology, See also Biotechnology Timeline (pg 5)</p> <p>Cancer, Comparative Genomics, Copy Number Variation, Identifying the Genetic Influences on Disease, Personalized Medicine</p> <p>Agricultural Applications, Cancer, Diagnosing Chromosomal Disorders, Epigenetics, Personal Genomic Analysis, Studying the Genome to Understand the Sequence</p>	

Linking Scientific Concept

Objective and Applicable Subheading

Course

Course	Objective and Applicable Subheading	Linking Scientific Concept		
Biology	8	Relating genetic disorders and disease to patterns of genetic inheritance.	Identifying Genetic Influence on Disease	
	9	Differentiate between the previous five kingdom and current six kingdom classification system.	Infectious Disease	
		Identifying ways in which organisms from the Monera, Protista, and Fungi Kingdoms are beneficial and harmful	Infectious Disease	
		Justifying the grouping of viruses in a category separate from living things	Infectious Disease	
	12	Describe protective adaptations of animals, including mimicry, camouflage, beak type, migration, and hibernation.	Comparative Genomics	
		Identifying ways in which the theory of evolution explains the nature and diversity of organisms	Comparative Genomics	
		Describing natural selection, survival of the fittest, geographic isolation, and fossil record	Comparative Genomics	
	Environmental Science	9	Describe land-use practices that promote sustainability and economic growth.	Agricultural Applications
	Forensic Science	4	Describe presumptive and confirmatory tests.	Criminal Justice and Forensics, DNA Sequencing
		5	Describe the importance of genetic information to forensics.	Criminal Justice and Forensics, DNA Sequencing
	Genetics	2	Describe factors such as radiation, chemicals, and chance that cause mutations in populations.	Cancer, Comparative Genetics, Identifying Genetic Influence on Disease, Infectious Disease, Studying the Genome to Understand the Sequence
			Describing effects of genetic variability on adaptations	Agricultural Applications, Comparative Genomics, Copy Number Variation, Criminal Justice and Forensics, RNA and Protein Analysis
4		Describe the process of meiosis and the cell cycle, including the hereditary significance of each.	Cancer, Diagnosing Chromosomal Disorders, Noninvasive Prenatal Diagnosis, Stem Cells	
5		Describe inheritance patterns based on gene interactions.	Diagnosing Chromosomal Disorders, Epigenetic, Genetics of Eye Color, Identifying Genetic Influence on Disease	
		Identifying incomplete dominance, codominance, and multiple allelism	Copy Number Variation, Epigenetics	
6		Describe occurrences and effects of sex linkage, autosomal linkage, crossover, multiple alleles, and polygenes.	Epigenetics, Identifying Genetic Influence on Disease, RNA and protein analysis	
7		Describe the structure and function of DNA, including replication, translation, and transcription.	DNA Sequencing, Recombinant DNA and Genetic Engineering, RNA and Protein Analysis	
	Describing methods cells use to regulate gene expression	Comparative Genomics, Epigenetics, Recombinant DNA and Genetic Engineering, RNA and Protein analyses, Therapeutic Approaches		
	Defining the role of RNA in protein synthesis	Recombinant DNA and Genetic Engineering, RNA and Protein analyses, Therapeutic Approaches		

Linking Scientific Concept

Objective and Applicable Subheading

Course

Genetics	8	Explain the structure of eukaryotic chromosomes, including transposons, introns, and exons.	Bioinformatics, Diagnosing Chromosomal Disorders, Studying the Genome to Understand the Sequence
	9	Differentiate among major areas in modern biotechnology, including plant, animal, microbial, forensic, and marine. Describing techniques used with recombinant DNA	Agricultural Applications, Bioinformatics, Cancer, Criminal Justice and Forensics, DNA Sequencing, Personalized Medicine, Pharmacogenomics, Recombinant DNA and Genetic Engineering, RNA and Protein Analysis Agricultural Applications, Recombinant DNA and Genetic Engineering, RNA and Protein Analyses
	10	Explain the development and purpose of the Human Genome Project. Analyzing results of the Human Genome Project to predict ethical, social, and legal implications. Describing medical uses of gene therapy, including vaccines and tissue and antibody engineering.	Bioinformatics, Criminal Justice and Forensics, DNA Sequencing, Identifying Genetic Influence on Disease, Studying the Genome to Understand the Sequence; See also Biotechnology Timeline (pg 5) Cancer, Copy Number Variation, Criminal Justice and Forensics, Genetic Information Nondiscrimination Act, Personal Genomic Analysis, Personalized Medicine, Pharmacogenomics, Therapeutic Approaches DNA Sequencing, Infectious Disease, RNA and Protein Analysis, Therapeutic Approaches
	II	Evolution	Agricultural Applications, Comparative Genomics
	IV	Continuity and Change	Agricultural Applications, Bioinformatics, Cancer, Comparative Genomics, Copy Number Variation, Criminal Justice and Forensics, DNA Sequencing, Genetics of Eye Color, Identifying Genetic Influence on Disease, Stem Cells, Studying the Genome to Understand the Sequence
	V	Relationship of Structure to Function	Epigenetics, RNA and Protein Analysis, Recombinant DNA and Genetic Engineering, Studying the Genome to Understand the Sequence
	VI	Regulation	Cancer, Copy Number Variation, Epigenetics, RNA and Protein Analyses
	VIII	Science, Technology and Society	Agricultural Applications, Cancer, Comparative Genomics, DNA Sequencing, Genetic Information Nondiscrimination Act, Identifying Genetic Influence on Disease, Noninvasive Prenatal Diagnosis, Personalized Medicine, Personal Genomic Analysis, Pharmacogenomics, Recombinant DNA and Genetic Engineering, Therapeutic Approaches, Synthetic Biology
	5	Evaluate negative and positive impacts of technology on health.	Agricultural Applications, Cancer, Identifying Genetic Influence on Disease, Noninvasive Prenatal Diagnosis, Personalized medicine, Pharmacogenomics, Recombinant DNA and Genetic Engineering, Stem Cells, Synthetic Biology
	Health	6	Discuss valid and essential information for the safe use of consumer goods and health products.
10		Determine the causes of disability and premature loss of life across life stages.	Cancer, Identifying Genetic Influence on Disease

Linking Scientific Concept

Objective and Applicable Subheading

Course	Objective and Applicable Subheading	Linking Scientific Concept
Technology Education	26 Explain uses and advantages of databases.	Bioinformatics
	27 Apply appropriate techniques for producing databases.	Bioinformatics
Agriscience	10 Determine characteristics and functions of plants. Explain how agricultural crops can be utilized as alternative fuel sources	Agricultural applications
Forensic and Criminal Investigations	7 Describe presumptive and confirmatory forensic tests. Examples: blood type comparison, DNA testing	Criminal Justice and Forensics
	8 Describe the importance of genetic information to forensics Using the process of gel electrophoresis for deoxyribonucleic acid (DNA) fingerprinting.	Bioinformatics, Criminal Justice and Forensics
Foundations of Health Sciences	10 Recognize legal responsibilities, limitations, and implications within the health care delivery setting. Examples: Patients' Bill of Rights, legal documentation requirements, Health Insurance Portability and Accountability Act (HIPPA)	Genetic Information Nondiscrimination Act, Personal Genome Analysis
Health Informatics	5 Describe legal and ethical regulations as they relate to health informatics. Examples: Patients' Bill of Rights, legal documentation requirements, Health Insurance Portability and Accountability Act (HIPPA)	Genetic Information Nondiscrimination Act, Personal Genome Analysis
Introduction to Agriscience	16 Analyze biotechnology to determine benefits to the agriculture industry. Example: Improved productivity, medical advancements, environmental benefits	Agricultural Applications, Bioinformatics, Recombinant DNA and Genetic Engineering
Introduction to Pharmacy	9 Identify classifications of selected drugs. Examples: analgesic, antibiotic, antiemetic	Personalized Medicine, Pharmacogenomics
	11 Differentiate among drug interactions, drug reactions, and side effects.	Personalized Medicine, Pharmacogenomics
Introduction to Biotechnology	1 Trace the history of biotechnology. Describing both scientific and non-scientific careers, roles, and responsibilities of individuals working in biotechnology.	See also Biotechnology Timeline (pg 5) Agricultural Applications, Bioinformatics, Criminal Justice and Forensics, Diagnosing Chromosome Disorders, DNA Sequencing, Pharmacogenomics, See also Biotechnology Timeline (pg 5)
	4 Correlate key cellular components to function.	See HudsonAlpha Cell (pg 5)
	5 Describe the process of meiosis and the cell cycle, including the hereditary significance of each.	Cancer, Diagnosing Chromosome Disorders, Noninvasive Prenatal Diagnosis, Stem Cells,
	8 Describe occurrences and effects of sex linkage, autosomal linkage, crossover, multiple alleles, and polygenes.	Cancer, Copy Number Variation, Genetics of Eye Color, Identifying Genetic Influence on Disease
	9 Describe the structure and function of deoxyribonucleic acid (DNA), including replication, translation, and transcription. Applying the genetic code to predict amino acid sequence	Recombinant DNA and Genetic Engineering, Studying the Genome to Understand the Sequence Bioinformatics

Linking Scientific Concept

Objective and Applicable Subheading

Course	Objective and Applicable Subheading	Linking Scientific Concept	
Introduction to Biotechnology	9	Describe methods cells use to regulate gene expression. Defining the role of ribonucleic acid (RNA) in protein synthesis	Cancer, Comparative Genomics, Epigenetics, RNA and Protein Analysis, Therapeutic Approaches Recombinant DNA and Genetic Engineering, RNA and Protein Analysis, Therapeutic Approaches
	11	Describe factors such as radiation, chemicals and chance that cause mutations.	Cancer, Infectious Disease
	13	Differentiate among major areas in modern biotechnology, including plant, animal, microbial, forensic, and marine. Describing techniques used with recombinant DNA	Agricultural Applications, Bioinformatics, Criminal Justice and Forensics, DNA Sequencing, Infectious Disease Agricultural Applications, DNA Sequencing, Synthetic Biology
	14	Explain the development, purpose, findings, and applications of the Human Genome Project. Analyzing results of the Human Genome project to predict ethical, social and legal implications Describing medical uses of gene therapy, including vaccines and tissue and antibody engineering. Using computer bioinformatics resources to provide information regarding DNA, protein, and human genetic diseases	Comparative Genomics, Copy Number Variation, DNA Sequencing, Identifying Genetic Influence in Disease, Personalized Medicine, Pharmacogenomics, Studying the Genome to Understand the Sequence Criminal Justice and Forensics, Genetic Information Nondiscrimination Act, Personalized Genomic Analysis Cancer, DNA Sequencing, Infectious Disease, Recombinant DNA and Genetic Engineering, RNA and Protein Analysis Bioinformatics, Cancer, Comparative Genomics, Copy Number Variation
	15	Describe the replication of DNA and RNA viruses, including lytic and lysogenic cycle.	Infectious Disease
Plant Biotechnology	1	Identify career opportunities associated with plant biotechnology.	Agricultural Applications
	14	Describe the ecological and economic importance of plants.	Agricultural Applications
	16	Explain the historical development of plant biotechnology. Identify medical advancements in plant biotechnology Describing environmental advancements in plant biotechnology	Agricultural Applications Agricultural Applications, Comparative Genomics Agricultural Applications; See also Biotechnology Timeline (pg 5)
	17	Describe methods of genetic engineering.	Agricultural Applications

FOUNDATIONAL CONCEPTS AND THEIR APPLICATIONS

Key Technologies

DNA Sequencing

In 1977 Fred Sanger and Alan Coulson published a method to rapidly determine the specific order of the adenine, thymine, cytosine and guanine nucleotides in any DNA sequence. This technology ultimately transformed biology by providing a tool for deciphering complete genes and later entire genomes. Improvements in process parallelization (running hundreds or thousands of samples simultaneously), automation and analysis led to the establishment of factory-like enterprises, called sequencing centers. These facilities spearheaded the effort to sequence the genomes of many organisms, including humans.

Today, the need for even greater sequencing capability at a more economical price has led to the development of new technologies based on different chemistries and refined for accuracy and speed. These “second generation” approaches reduce the necessary volume of reagents while dramatically increasing the number of simultaneous sequencing reactions in a single experiment. They are capable of producing nearly 150 times more sequence than the first generation systems, at 1/150th the cost. For example, the cost of sequencing all 3 billion letters in the human genome has dropped from \$15,000,000 to less than \$40,000.

The ability to quickly and economically decipher large swaths of DNA has opened doors to research previously deemed out of reach. Many of the discoveries outlined in this guide are in part due to this new technology.

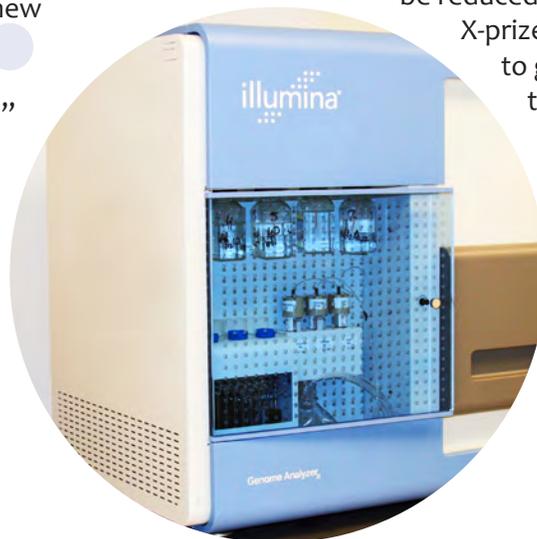
The first so-called “third generation” sequencing system debuted in 2009, producing an entire human sequence. Based on the analysis

Second- and third-generation sequencing technologies should be briefly discussed in Biology courses as part of course of study (COS) objective 8, particularly as it relates to significant contributions of biotechnology to society. These topics should be more thoroughly explored in Genetics classes, relating to COS objectives 7, 9 and 10, especially with respect to the impact such technologies have on identifying genetic risks, personalized medicine and pharmacogenomics. They may also be incorporated in the Forensic Science class in preparation for a discussion about DNA phenotyping (see page 29) as part of COS objective 4 and 5 or in an AP Biology course as part of the “Science, Technology and Society” and “Continuity and Change” general themes. This topic would also be appropriate for discussion in the Career/Tech Intro to Biotechnology course as part of objectives 1, 13 and 14.

HudsonAlpha educators have developed a high school lab activity, “Genes & ConSEQUENCES”, that connects the information produced by a DNA sequencing system to genes, mutations and disease. The activity incorporates biological databases used by genetic researchers on a daily basis and links changes in DNA sequence to common genetic disorders (see “Bioinformatics” on page 24 for more details). The lab was incorporated into the AMSTI high school program (Science in Motion) statewide during the 2010-2011 academic year.

of a single molecule of DNA, a major technological improvement, it is believed that these systems will become widespread within the next 2-3 years, further decreasing sequencing costs.

Looking towards the third (and fourth) generation sequencing systems, there remains a long list of necessary improvements. Chief among them is cost reduction: in order to deliver on the goal of sequencing a human genome for \$1,000, sequencing costs must be reduced by an order of 1-2 magnitudes. The X-prize in genetics serves as an incentive to groups working on sequencing technologies: a \$10M prize to the first group who sequences 100 human genomes in 10 days or less at a per genome cost of no more than \$10,000.



RNA and Protein Analyses

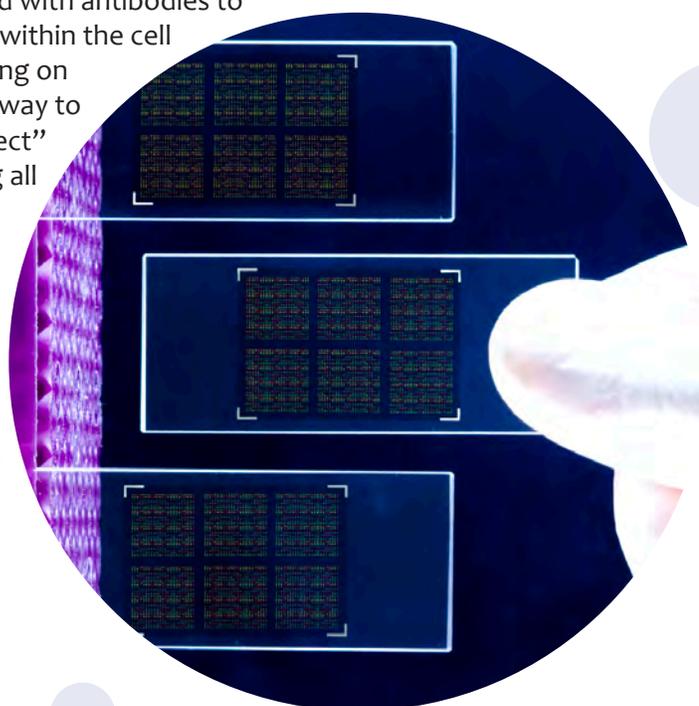
As sequencing techniques identify the genetic recipes of an organism, understanding the function of those genes becomes increasingly important. Many of the fundamental measurements used by molecular biologists seek to determine the presence, absence or relative amounts of RNA produced by a gene. Initially, these approaches examined one or only a handful of RNA sequences at a time. During the last decade, researchers developed techniques to study tens of thousands of RNA fragments simultaneously arrayed on a glass slide. Called “microarrays”, these could be used to identify which genes are active or silent in a given cell type, classifying, for example, the genes that distinguish a liver cell from a neuron or the set of genes activated or silenced across different types of cancer.

Second-generation sequencing technology has recently been extended to also identify RNA expression across cells. Scientists have shown that this approach, known as RNA-seq, yields more precise results than microarray analysis. It is expected that RNA-seq will become the standard tool for measuring genome-wide gene expression.

Large-scale, high-throughput technologies have also been developed to identify protein activity and interactions. This represents part of the emerging field of proteomics, which seeks to understand the entire protein complement (amounts, locations, interactions, and even activities) of an organism’s cells. For example, “tissue microarrays”, tiny slices of tissue from a single or multiple samples, can be tested with antibodies to identify the locations of proteins within the cell and their relative amounts. Building on these methods, efforts are underway to initiate a “Human Proteome Project” that would systematically catalog all the proteins manufactured in the body. The scale and complexity of this project is much greater than the Human Genome Project as a single gene can direct the production of multiple different versions of a protein and each protein can in turn be modified in a number of different ways.

RNA- and Protein-based technologies should be noted in a Biology course, as it relates to both COS objectives 2, 5 and 8 as they strive to identify the function of proteins and nucleic acids in cellular activities. These technologies can be examined in greater detail for either an AP biology course (under the “Relationship of Structure to Function” and “Regulation” themes) or a Genetics course, where they can be incorporated into activities that describe the occurrence and effects of genetic variability on populations (COS 2 and 6), methods used to regulate gene expression (COS objective 7), techniques using recombinant DNA and antibody engineering (COS objectives 9 and 10). These are also useful technologies to cover in the Career/Tech Intro to Biotechnology course, linking to COS objectives 9 and 14.

Many of the fundamental measurements used by molecular biologists seek to determine the presence, absence or relative amounts of RNA produced by a gene.

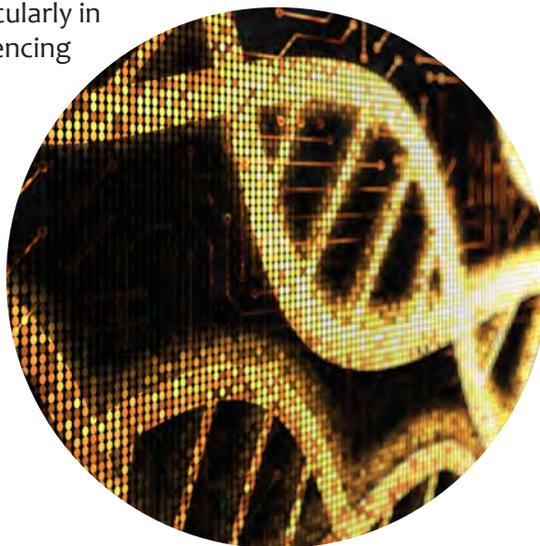


Bioinformatics

Acquiring DNA sequence has now become routine and new technologies can sequence a bacterial genome in a single day. Similarly, microarray experiments shed light on the RNA levels produced by tens of thousands of genes. Current analysis platforms are capable of generating terabytes of data in a single run. For reference, 1 terabyte is equal to 1,000 gigabytes - enough storage space to hold 500 copies of your favorite box office movie or the music libraries from nearly 125 iPod nanos.

Understanding the meaning of all that information is a daunting challenge. Deciphering the data requires a biological knowledge of what to look for, algorithms (computer programs) capable of detecting interesting features, and computers powerful enough to perform complex analyses efficiently and rapidly. Fortunately, advances in all three areas have kept pace and the resulting field of bioinformatics seeks to characterize functional sequences in genes and genomes through computational models. In addition, the data must be managed – stored in a form that is useful to the researcher and readily accessible. This has led to the development of many databases that store and provide data and analytical tools for researchers. The primary mission of all these databases is to provide unlimited free access to anyone, including Alabama students, interested in studying genomic sequences. It is no exaggeration to say that these databases and the immediate access to them through the Internet have changed the way that nearly all biological research is done.

Many bioinformatics experts, particularly in the early days of the genome sequencing efforts, were computer scientists who formed partnerships with biologists. With the growth of the field of genomics, it is not unusual today for a student to be trained in a truly interdisciplinary way by developing deep expertise in both biology and computational science.



The concept of bioinformatics is a critical component to understanding modern genomic discoveries. It provides tools capable of exploring the structure of chromosomes and predicting the likelihood of a genetic match in a forensics case. Bioinformatics databases also manage, search and store the data produced by the human genome project and more recent large-scale studies (Genetics COS objectives 8, 9 and 10). This topic should be incorporated in an AP Biology class under the general theme “Continuity and Change”, as well as Career/Tech courses in Forensic and Criminal Investigations (COS objective 8), Introduction to Agriscience (COS objective 16) and Intro to Biotechnology (COS objectives 1, 9, 13 and 14). Lastly, the creation, management and utilization of bioinformatics databases can be incorporated into the Technology Education course (COS objectives 26 and 27).

HudsonAlpha educators have developed a high school lab activity, “Genes & ConSEQUENCES”, that connects the information produced by a DNA sequencing system to genes, mutations and disease. The activity incorporates several biological databases used by genetic researchers on a daily basis. Students access a portion of the NCBI (National Center for Biotechnology Information) database known as BLAST. This program compares sequence data entered by the student to known sequences from a number of organisms, including human, and identifies genetic matches. Students then explore their matches on another NCBI database called Genes & Diseases. This dataset allows students to determine the chromosomal location of the gene and its role in disease. The lab has been incorporated into the AMSTI Science in Motion program statewide during the 2009-2010 academic year.

Application

Agriculture

Sequencing Plant Genomes for Food and Bioenergy Needs

Over the last decade, genome sequencing projects have begun for a number of plants, including rice, corn, soybean, canola, and orange. The goal of these sequencing efforts is a better understanding of the underlying genes that contribute to growth rate, seed and fruit characteristics and susceptibility to climate change or infectious agents. In addition, a number of plants have been or are being sequenced for their potential contribution to bioenergy. These include corn, soybean, loblolly pine, poplar and switchgrass. For example, soybean not only accounts for 70% of the world's edible protein, but soybean oil is the principle source of biodiesel. Detailed knowledge of the soybean genome, published in December 2008, allows for crop improvements and better applications of this plant to the generation of clean energy. Knowing which genes control specific traits, researchers could potentially change the type and quantity of oil produced by the crop as well as develop soybean plants that are more resistant to drought or disease.

Genetically Modified (GM) Crops

More than 13 million farmers across 25 countries currently plant biotech crops (also known as genetically modified organisms or GMOs). To date, over two billion acres of biotech crops have been harvested globally. At least 57 different plants have been the focus of biotech research over the last two decades. Of this number, eight different plants are in commercial production, and 15 different plants have received regulatory approval in the United States. Currently, biotech soybean is the principal genetically modified crop worldwide, followed by corn, cotton and canola. Herbicide tolerance has consistently been the primary trait introduced into the crops, followed by insect resistance and the combination of both traits. Biotechnology has enabled producers worldwide to produce higher yields on existing land. Biotechnology crops reduce the need for plowing to control weeds, leading to better conservation of soil and water and a



The application of genetic information and Genetically Modified Organisms to increase agricultural yields, improve nutritional content, craft insect resistance or increase bioenergy yields has a direct connection to COS objective 8 for Biology and COS objective 9 for the Environmental Science class. It can also be discussed in a Genetics course (COS objectives 2 and 9) and AP Biology as part of general themes “Evolution”, “Continuity and Change” and “Science, Technology and Society”. It also has a direct connection to Career/Tech courses in Agriscience (COS objective 10), Intro to Agriscience (COS objective 16), Intro to Biotechnology (COS objectives 1 and 13) and Plant Biotechnology (COS objectives 1, 14, 16 and 17).

HudsonAlpha has created a lab that allows students to test foods available at their local grocery store (such as chips and cookies) for the presence of genetically modified crops. The test primarily identifies various forms of herbicide resistant corn and soybeans and exposes students to DNA extraction, DNA amplification by the polymerase chain reaction (PCR) and DNA electrophoresis to separate fragments of varying length. Students test foods for both the genetic modification and a control gene from plant cells. This activity is an excellent link between key food safety techniques, foundational biotechnology methods and a subject of interest to all high school students – food. The ethical challenges associated with biotech crops and the varying global viewpoints are also presented for discussion, and the lab links to careers in food safety and inspection. The G-Mod lab is available to all Alabama high school students through the AMSTI Science in Motion program.

decrease in soil erosion and soil compaction. A reduction in plowing also allows farmers to significantly reduce the consumption of fuel and decrease greenhouse gas emissions. Plants modified to contain genetic variants that confer a degree of drought tolerance, particularly important for developing countries, will become available within the next five years.

Researchers are also developing biofortified food plants to boost the levels of nutrient, vitamins and minerals in foods such as rice, cassava, carrots and tomatoes. It is hoped that these fortified foods will reduce the incidence of global hunger and micronutrient malnutrition (taking in adequate calories, but lacking appropriate vitamins and minerals) which, according to a 2004 United Nations report, impacts up to half of the world's population.

Cancer

Cancer is a collection of diseases that are characterized by uncontrolled growth of cells and their spread to surrounding tissues. All cancers are genetic diseases, because changes in the genes that control cell growth and division are involved. However, only about 5% of cancers are strongly hereditary – primarily caused by mutations that are inherited from parent to child. Therefore, most cancers do not result from inherited mutations, but instead develop from an accumulation of DNA damage acquired during our lifetime. These cancers begin with a single normal cell that becomes genetically damaged. The transformation from that initial cell into a tumor is a stepwise progression. The number of genetic mutations that are required to convert a genetically normal cell into an invasive tumor is not known but most likely varies among cancer types. These genetic changes may involve single “letter” or base substitutions, large deletions or duplications, or chromosomal rearrangements impacting vast sections of the genome. Most cancer cells have a number of both large-scale chromosome abnormalities as well as single letter mutations.

Historically, the diagnosis and staging of cancers has been based on the appearance of the cancer cells under a microscope, and the spread to surrounding or distant tissues. Treatment decisions and options are often based upon this information. However, in many cases, individuals with similar-appearing tumors will show markedly different responses to treatment. We now know that differences at the molecular level, not visible under a microscope, are responsible for the varying outcomes.

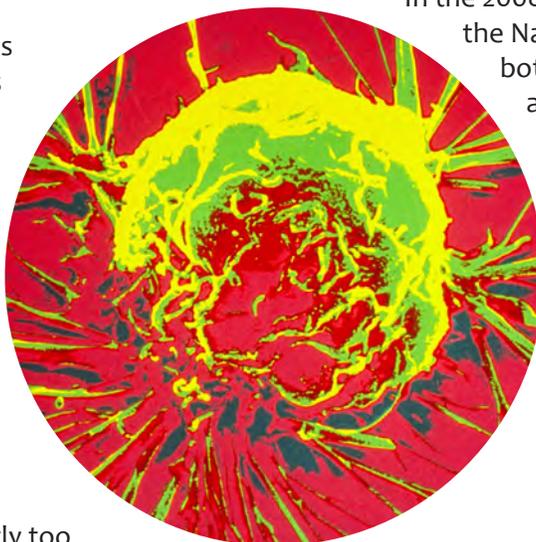
Microarray-based expression studies can be used to identify which genes are activated or silenced in the formation of cancer. Expression patterns can classify patients into groups that correlate with cancer subtypes and responses to a specific drug or clinical outcome. If validated, these differences can be used to predict outcomes for new patients, helping physicians identify the most optimal treatment or course of action.

Microarray experiments are currently too cumbersome to perform in a clinic, so it is not likely they will be used routinely to diagnosis patients. However, once a small subset of the genes

The idea that all cancers are genetic in nature and occur as a stepwise addition of mutations, many of which are initiated by environmental factors, is a useful addition to a discussion on common causes of disability and premature loss of life in a Health class (COS objective 10). These concepts should also be incorporated into Biology (COS objectives 6, 7 and 8), Genetics (COS objectives 2, 4, 9 and 10), and AP Biology (general themes “Continuity and Change”, “Regulation” and “Science, Technology and Society”). There are also several points of linkage with the Career/Tech Intro to Biotechnology course (COS objectives 5, 11 and 14). In all cases, the distinction should be made between a relatively small number of cancer types with strong inherited risks and most forms of cancer that are primarily due to mutations acquired throughout the life of the individual.

HudsonAlpha has developed a high school lab that focuses on various forms of cancer and methods for their detection. This lab gives students experience in drawing a family pedigree (a genetic family tree) and interpreting the pedigree with respect to a specific form of inherited colon cancer. The students will then complete and analyze a DNA-based diagnostic test to identify which family members have inherited the cancer-causing mutation. The lab activity also introduces students to a genetic counselor and laboratory technician for career exploration. The HNPCC lab has been incorporated into the AMSTI Science in Motion program and is currently available to high school life science teachers across Alabama.

most relevant to predicting disease or treatment outcome is discovered, it becomes possible to detect the corresponding protein levels in the cancer cells using specially labeled antibodies. For example, some of these proteins have been identified for breast cancer. Detecting whether each protein is present and at what level is useful in determining which therapy will be most effective for treatment. See the table on page 40-41 for specific genetic tests used in this manner.



In the 2008 “Annual Report to the Nation”, the National Cancer Institute noted that both the incidence and death rate for all cancers combined is decreasing. While cancer death rates have been declining for several years, this marks the first decline in cancer incidence, the rate at which new cancers are diagnosed.

Comparative Genomics

Although the human genome is perhaps the most famous sequencing project, scientists have assembled a genomic library of over 200 different organisms. Knowing the genome of each species provides insight into the function of its DNA; however, there is additional information gained by comparing genomes across organisms. This field of comparative genomics helps discover previously undetected genes, identify the regulatory regions that control gene activity and determine gene function as it relates to health and disease

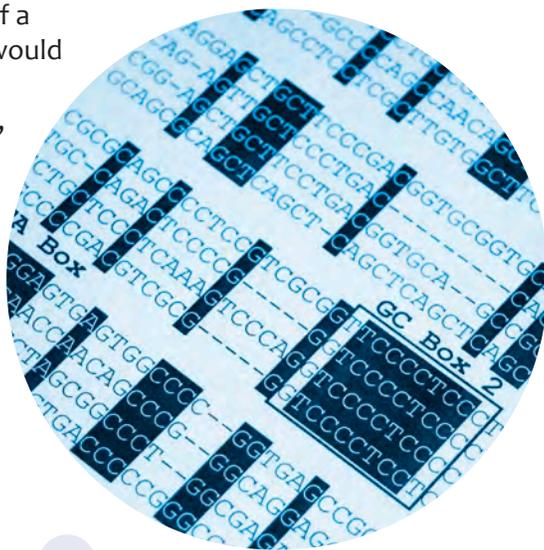
While humans may seem to have little in common with organisms such as fruit flies, roundworms or mice, they are all composed of cells that must take in nutrients and remove waste, interact with neighboring cells and the outside environment, and grow and divide in response to specific signals. To varying degrees, each of these organisms contains a digestive, circulatory, nervous and reproductive system and is impacted by disorders that impair these systems. During the evolutionary process, as organisms diverged and gave rise to new species, many key proteins such as enzymes, underwent little change. In general, the nucleotide and amino acid sequences of these key proteins have similarly been conserved across the species.

Scientists directly compare the DNA sequence of these organisms, using sophisticated computer programs that line up multiple genome sequences and look for regions of similarity. These similar segments or conserved sequences suggest the DNA sequence has an important functional role – for example, a gene or a regulatory element that controls the activity of a gene. Less critical DNA segments would accept sequence changes without clinical consequence: subsequently, these segments would vary among species. Genes that have relatively high sequence similarity are referred to as homologous genes or homologues. Comparative genomics provides

Comparative genomics provides evidence for the molecular process that underlies evolutionary theory and explains the nature and diversity of organisms, as outlined in the Biology COS objectives 5, 8 and 12 as well as in the Genetics COS objectives 2 and 7. Comparative genomics and its relationship to evolution intersects AP Biology, particularly with respect to general themes “Evolution”, “Continuity and Change” and “Science Technology and Society”. Career/Tech courses will also benefit from a discussion of comparative genomics, including Veterinary Science (COS objective 3) and Intro to Biotechnology (COS objective 9, 11 and 14).

a powerful tool for studying evolutionary changes among organisms, identifying genes that are conserved among species as well as gene and genetic changes that give each organism its unique characteristics.

Genomic comparison also extends to genes involved in disease. If we examine the current list of human disease genes, approximately 20% have a homolog in yeast and nearly two-thirds have one in flies and worms. Initial studies suggest these counterparts may function in nearly identical ways, meaning these organisms can serve as models for understanding human disease and potential treatment. For example, studying genes involved in DNA repair in yeast or bacteria has offered valuable insight into this process in humans and the role that mutations of these genes play in the development of some cancers.



Copy Number Variation

For years single nucleotide polymorphisms (SNPs) were thought to be responsible for the majority of human variation. Until recently, larger scale changes (1000+ nucleotides in length), known as Copy Number Variants (CNV), were thought to be relatively rare. However, scientists have discovered that CNVs occur much more frequently than was suspected. These structural changes alter the number of copies of a specific DNA segment.

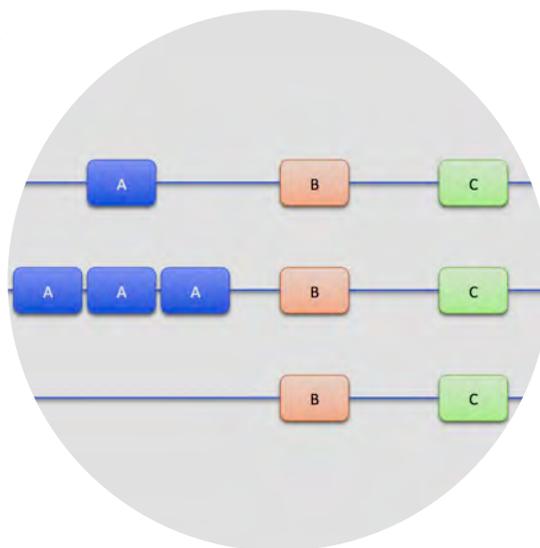
It came as a surprise to many scientists just how much DNA variation is due to copy number changes. Previous studies based primarily on SNPs suggested that any two randomly selected human genomes would differ by 0.1%. CNVs revise that estimate: the two genomes differ by at least 1.0%. While this may not seem like a major increase, remember that the human genome is composed of approximately 3 billion nucleotides, so the estimated number of nucleotides that vary between two random individuals has increased from 3 million to 30 million. Humans are still nearly 99% identical at the DNA sequence level, but the CNV research has broadened our understanding of how and where we differ.

It has been suggested that CNV regions influence gene activity by directly increasing or decreasing the number of copies of that gene, leading to a concurrent change in the amount of protein. Alternately, CNVs may alter the performance of nearby regulatory signals that activate or silence genes without directly impacting the copy number of the gene itself.

Preliminary studies have linked CNVs to lupus, Crohn's disease, autism spectrum disorders, Alzheimer disease, HIV-1/AIDS susceptibility, rheumatoid arthritis and Parkinson's disease. In some cases the associated CNV is rare, but in other diseases, the identified risk variant is quite common. It is also likely that CNVs may influence individual drug response and susceptibility to infection or cancer.

Relating genetic variation to human disease and inheritance is identified in the Biology COS under objective 8 and is described in detail in the Genetics COS objectives 2 and 5, particularly as it connects with genetic patterns of inheritance and multiple alleles. Genetic variation as it relates to human disease also is highlighted under objective 10, which explores the ongoing impacts from the Human Genome Project. AP Biology themes "Continuity and Change" and "Regulation" also intersect the topic of copy number variation, as does Career/Tech course Intro to Biotechnology (COS objective 8).

Preliminary studies have linked copy number variation to lupus, Crohn's disease, autism spectrum disorders, Alzheimer disease, HIV-1/AIDS susceptibility, rheumatoid arthritis and Parkinson's disease.



Criminal Justice and Forensics

DNA profiling, popularly known as DNA fingerprinting, has transformed personal identification, whether in forensic cases, missing persons, mass disasters or paternity disputes. It has become ubiquitous in law enforcement. It is used to exclude individuals suspected of crimes, help convince a jury of an individual's guilt and in some cases, set free individuals wrongly convicted of crimes.

DNA analysis is also used to suggest ancestral origins; there are several companies offering Y-chromosome and mitochondrial DNA studies to determine, for example, to which of the ancient tribes of Britain a man belongs or whether a man or woman has African, Native American or Celtic DNA markers. It is possible to use forensic DNA profiling in the same way to determine the ethnic or geographical origin of the individual from whom the DNA sample came, providing additional information that could be used to narrow the number of potential suspects. For example, in 2007, a DNA test based on genetic biomarkers indicated that one of the suspects associated with a bombing in Madrid was of North African origin. Using other evidence, police confirmed the suspect was an Algerian, confirming the test result.

It has been suggested that this testing could be extended to identify external and behavioral features as well. Scientists have recently identified the genetic variants related to hair, skin and eye color and are exploring other genes that influence traits such as facial height and width as well as nose and lip shape. This “forensic molecular photo fitting” may one day serve as a genetically-based police sketch. Today this approach is still primarily theoretical and currently has little concrete value. As noted throughout this guide, it will take years before the genetic markers associated with all physical and behavioral traits are known.



DNA profiling is a critical component of the Forensics science elective, as part of COS objectives 4 and 5, as well as the Career/Tech course Forensic and Criminal Investigation (COS objectives 7 and 8). It can also be explored in AP Biology as part of the general theme “Continuity and Change”, in Genetics as part of COS objectives 9 and 10 and in the Career/Tech course Intro to Biotechnology linked to COS objectives 1, 13 and 14. DNA phenotyping should be an extension of the discussion in all three of these classes, highlighting the concepts and technological challenges still facing the field. The ethical complications of phenotyping should also be incorporated into the discussion.

Legislatively, forensic phenotyping is allowed on a limited basis in some countries (such as the UK) and forbidden in others (Germany). However, for most of the world, legislation that addresses DNA forensic methods is silent about the ability to infer ethnicity or physical traits.



Diagnosing Chromosome Disorders

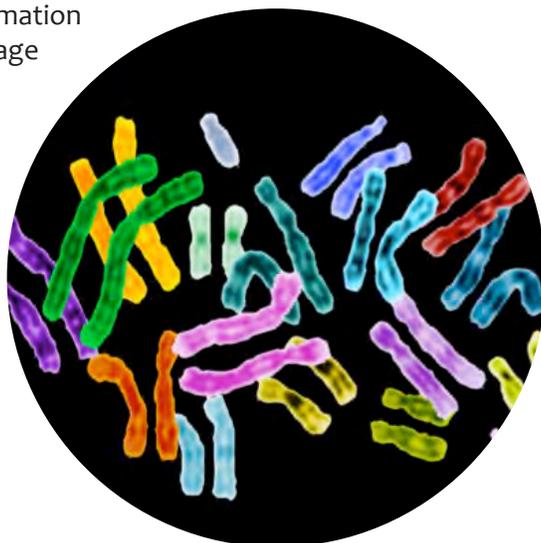
Although scientists have been able to microscopically observe chromosomes since the mid-1800's, a century passed before staining techniques were developed to examine them on a specific and individual basis. The chromosomes could then be arranged according to size and banding pattern for detailed examination - a display called a karyotype. Once it became possible to accurately identify individual chromosomes, abnormalities in chromosome number (such as trisomy 21, also known as Down syndrome) were discovered. Karyotypes can also identify deletions, duplications, and inversions of chromosomal segments.

Although abnormalities on the order of millions of base pairs can be detected using the basic chromosomal banding techniques, smaller alterations cannot be discerned. More recent technologies, such as fluorescence in situ hybridization (FISH) and array comparative genome hybridization (array CGH), allow a finer level of resolution, with the ability to identify submicroscopic chromosome changes.

Although array CHG is still relatively new, it appears to hold great promise for detecting chromosome disorders both large and small. Over the next 3-5 years, this technology will likely become the standard chromosome diagnostic tool to detect abnormalities in chromosome number, microdeletions and other chromosome imbalances. In 2009, clinicians in the UK developed a screening method based on array CGH to identify the most viable eggs obtained from older women undergoing in vitro fertilization (IVF). Array CGH was used to examine the chromosomes from the polar body, a by-product of egg formation that generally serves as a mirror image of the chromosomes found in the egg itself.

Chromosome studies, their behavior in cell division, the formation of egg and sperm and the concept of karyotyping are regularly discussed in Biology classes under the requirements of COS objectives 6 and 8. Karyotypes and their ability to diagnose chromosomal disorders are examined in Genetics classes as part of COS objectives 4,5 and 8, as well as in the Career/Tech course Intro to Biotechnology (COS objectives 1 and 5). The techniques of FISH and aCGH should also be discussed with students in these classes, although many of the technical details need not be described. It is important for students to realize that there are a number of genetic disorders that cannot be identified at the karyotype level, but the newer technologies bridge the gap between studies of stained chromosomes and DNA sequencing.

The HudsonAlpha education team has crafted a karyotype lab as a modification to an existing AMSTI Science in Motion chromosome lab for high school biology and genetics classes. In 'Disorder Detectives', students take on the role of a cytogeneticist working in a hospital or clinic and are given a case study and a set of human chromosomes. They arrange the chromosomes on a prepared board into a completed karyotype, analyze the karyotype and diagnose their patient. Many types of normal and abnormal chromosomal cases are presented. Students also explore the more recent techniques of FISH and aCGH to learn how these technologies provide the ability to diagnose increasingly small genetic imbalances. Geneticists, genetic counselors, and laboratory technicians are highlighted as careers that utilize these types of technologies. The module has been incorporated into AMSTI training at all 11 sites across Alabama and is currently in use by students.



Epigenetics

While identical twins (twins who share the same genetic information) generally look alike when young, obvious differences often emerge as they age. The differences may be due to the varied environment of each twin – for example, one may lift weights and become very muscular while the other never exercises and gains weight. Recent advances in the relatively new field of epigenetics suggest an additional role for the environment in health and disease by altering the activity of particular genes. Activating genes to begin the protein-making process is a key area of study. By identifying the signals that turn genes “on” and “off”, investigators hope to understand not only gene function under normal conditions, but also how improper on/off signaling may lead to disorders such as cancer, diabetes, heart disease and obesity.

Epigenetics encompasses modification to DNA, including the addition of small chemical tags called methyl groups. These modifications alter the patterns of gene activity, but do not change the actual DNA sequence. The modifications are not permanent, but can be remembered across thousands of cell divisions and at times from parent to child. This field includes some of the most fascinating biological phenomena, including X-chromosome inactivation, imprinting (when the DNA copy inherited from a particular parent is silenced, while the other copy remains active) and cellular differentiation (see the article on stem cells, page 43).

Studies of identical twins suggest that at birth, twins share similar patterns of epigenetic modification. As they age and are exposed to different diets and environments, the twin’s patterns become markedly different, leading to altered activation and silencing patterns.

Current research suggests environment alterations to these epigenetic patterns can change an individual’s risk for disease. For many mammals (humans included), differences in diet and level of stress during fetal

Epigenetic changes in DNA often lead to unusual patterns of inheritance for specific disorders. This could be discussed as part of a lesson on exceptions to standard Mendelian inheritance for Biology COS objectives 7 and 8, Genetics COS objectives 5-7, and Intro to Biotechnology COS objective 9. The relationship between the methyl modifications on the DNA and the gene silencing links epigenetics to AP Biology through general themes “Relationship of Structure to Function” and “Regulation”.

development and shortly after birth alter the pattern of on/off gene activity, leading to higher risk of obesity, type 2 diabetes and cardiovascular problems. These observations have a number of clinical and public health implications.

Epigenetics involves DNA modifications that alter the patterns of gene activity, but do not change the actual DNA sequence. This field includes some of the most fascinating biological phenomena, including X-chromosome inactivation, imprinting and cellular differentiation.



Genetic Information Nondiscrimination Act

While most Americans are optimistic about the use of genetic information to improve health, many have been concerned that genetic information may be used by insurers to deny, limit or cancel health insurance and by employers to discriminate in the workplace. There has also been concern that some insurers may choose to not insure healthy individuals who are genetically pre-disposed to future disease onset: such people incur more health-related costs for the insurance company than individuals who are not predisposed. A similar fear is that some employers might only employ or retain individuals who are not pre-disposed to future disease onset, since healthy individuals are more productive. Consequently, for many years lawmakers, scientists and health advocacy groups have argued for federal legislation to prevent genetic discrimination.

In 2009, the Genetic Information Nondiscrimination Act (GINA) took effect across America, paving the way for people to take full advantage of the promise of personalized medicine without fear of discrimination. The act had been debated in Congress for 13 years and was signed into law in 2008. GINA protects Americans against discrimination based on their genetic information when it comes to health insurance and employment. The law, together with existing nondiscrimination provisions from other laws, prohibits health insurers or health plan administrators from requesting or requiring genetic information of an individual or the individual's family members, or using it for decisions regarding coverage, rates, or preexisting conditions. The law also prohibits most employers from using genetic information for hiring, firing or promotion decisions.

GINA's protection does not extend to life, disability, or long-term care insurance. In addition, GINA does not prohibit a health insurer from determining eligibility or premium rates for an individual who is already exhibiting clinical symptoms of a disease or disorder.

Genetic discrimination should be briefly discussed in Biology courses as part of COS objective 8, particularly as it relates to significant contributions of biotechnology to society. It could be explored in AP Biology courses under "Science, Technology and Society" general theme and in Genetics classes in light of the ethical, social and legal implications of the Human Genome Project (COS objective 10). There are additional linkages to the Career/Tech courses Foundations of Health Science (COS objective 10), Health Informatics (COS Objective 5) and Intro to Biotechnology (COS objective 14).

In 2009, the Genetic Information Nondiscrimination Act (GINA) took effect across America, paving the way for people to take full advantage of the promise of personalized medicine without fear of discrimination.



Genetics of Eye Color

In 1907, Charles and Gertrude Davenport developed a model for the genetics of eye color. They suggested that brown eye color is dominant over blue eye color. This would mean that two blue-eyed parents would always produce blue-eyed children but never ones with brown eyes. For most of the past 100 years, this version of eye color genetics has been taught in classrooms around the world. It is one of the few genetic concepts that adults often recall from their high school or college biology classes. Unfortunately, this model is overly simplistic and incorrect – eye color is actually controlled by several genes.

In humans, eye color depends on the level of a pigment called melanin present in the iris. Melanin is produced and stored inside specialized cells known as melanocytes. Blue eyes contain minimal amounts of melanin. Irises from green–hazel eyes show moderate pigment levels, while brown eyes are the result of high melanin concentrations.

To date, eight genes that impact eye color have been identified. The *OCA2* gene, located on chromosome 15, appears to play the major role in controlling the brown/blue color spectrum. *OCA2* produces a protein called P-protein that is involved in the formation and processing of melanin. *OCA2* alleles (versions of the gene) related to eye color alter P-protein levels by controlling the amount of *OCA2* RNA that is generated. The allele that results in high levels of P-protein is linked to brown eyes. Another allele, associated with blue eye color, dramatically reduces the P-protein concentration.

While studies suggest that about $\frac{3}{4}$ of the eye color variation can be explained by genetic changes in and around *OCA2*, it is not the only genetic influence on color. A recent study that compared eye color to *OCA2* status showed that only 62% of individuals with two copies of the “blue eyed” *OCA2* allele actually had blue eyes. Blue eye color was also found among 7.5% of the individuals with the brown-eyed *OCA2* alleles. A number of other genes (such as



The multifactorial genetics of eye color should be discussed in Biology courses as part of COS objective 7, and in Genetics courses under COS objective 5, especially since most textbooks still explain this trait in terms of a single gene effect. It could also be explored in AP Biology courses under “Continuity and Change” general theme. In the Career/Tech Intro to Biotechnology courses, eye color genetics could be explored under COS objectives 8 and 11.

TYRP1, *ASIP*, and *ALC42A5*) also function in the melanin pathway and shift the total amount of melanin present in the iris. The combined efforts of these genes may boost melanin levels to produce hazel or brown eyes or reduce total melanin resulting in blue eyes. This explains how two parents with blue eyes can have green or brown eyed children (an impossible situation under the Davenport single gene model) – the combination of color alleles received by the child resulted in a greater amount of melanin than either parent individually possessed.

Identifying Genetic Influence on Disease

Much progress has been made in identifying the genetic causes of single gene diseases such as cystic fibrosis, phenylketonuria and Huntington disease. This has led to more accurate risk analysis, better testing approaches and, in some instances, more effective methods of treatment. Even though there are thousands of single gene disorders, they are rare, affecting less than 3% of the population.

In contrast, other diseases, including cleft lip, cardiovascular disease, psychiatric disorders, and cancer, affect much of the world's population. While these diseases have a strong genetic component, they arise from a combination of genetic risk factors that are also influenced by the environment. Few of the contributing genes are believed to make more than a modest contribution to overall risk, perhaps increasing it by 5 or 10%. It is the specific combination of multiple predisposing alleles (DNA changes) and environments that leads to physical symptoms. For this reason, they are often called complex or multifactorial disorders. Identifying the factors that influence disease is a major goal for biomedical research.

Traditional methods of determining the genes responsible for single-gene disorders do not work well for complex diseases. Fortunately, thanks to the advent of second-generation technology to cheaply analyze DNA changes, scientists are using a process known as genome-wide association (GWA) to identify the genetic factors involved in complex disease.

The basic premise behind GWA studies is straightforward: if a specific genetic variation increases the risk of developing a disease, that variation will occur more frequently - and hold up under rigid tests for statistical significance - in individuals who have the disease compared to those not affected. In other words, there is an association between the specific allele and the incidence of disease.

Successful genome-wide association studies test large numbers of variable DNA sites, using DNA microarrays (also called "gene chips") that contain up to one million microscopic spots of DNA. Each spot corresponds to a genetic change. While many of these changes occur with genes, others are in DNA sequences that may be important in regulation or expression of genes.

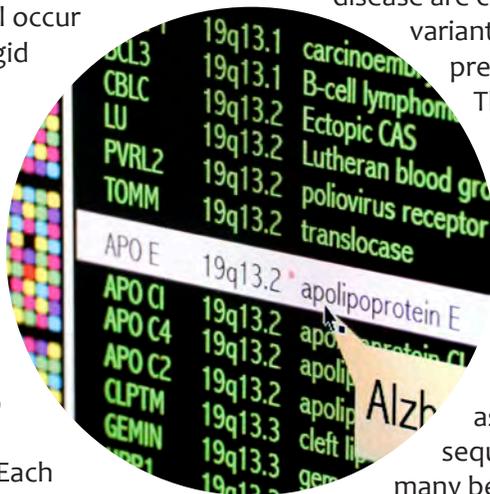
Relating genetic variation to human disease and inheritance is identified in the Biology COS under objective 8 and is described in detail in the Genetics COS objectives 2 and 5, particularly as it connects with genetic patterns of inheritance and multiple alleles. Genetic variation as it relates to human disease also is highlighted under Genetics objectives 6 and 10, which explore influence of multiple alleles as well as the ongoing impacts from the Human Genome Project. This would also be an appropriate discussion for an AP Biology course ("Continuity and Change and "Science, Technology and Society"), Health (COS objectives 5 and 10) and the Career/Tech Intro to Biotechnology course (COS objective 14).

This technology allows a researcher to simultaneously examine hundreds of thousands of genetic variants that span the human genome – a previously unfathomable accomplishment.

Until recently, researchers knew of almost no genetic variants involved in complex diseases. As of 2010, over 800 genetic single nucleotide polymorphisms have been associated with more than 150 complex diseases or traits. Most of the newly associated genes have not previously been linked to the disease of interest. Intriguingly, some genetic regions have been associated with multiple disorders, suggesting common chemical pathways that influence a number of different processes.

Even with these successes, the majority of the genetic risk for common disease remains undiscovered and the contribution by a single genetic variant to the overall clinical picture is often small. As a result, scientists are beginning to think that many of the genetic risks for disease are caused by a number of so-called "rare variants", genetic changes that are each present in less than 1% of the population.

This view represents a shift from previous beliefs that complex diseases were caused by variants that were much more common. Projects aimed at sequencing the genomes of a larger number of individuals will hopefully identify many of these "rare variants", allowing this hypothesis to be tested. In addition, as emerging technologies in DNA sequencing continue to drive down costs, many believe GWA studies will shift from examining specific sites of known genetic variation towards full sequencing of the entire genome. At that point, identifying even the rarest of variation becomes feasible.



Infectious Disease

The impact of infectious disease is a major healthcare challenge. Antibiotic resistant strains of pneumonia and staph infections are surfacing in hospitals, nursing homes and locker rooms. The 2009 H1N1 virus confirms long-held concerns about a pandemic influenza virus spreading unchecked across the globe. In both cases, the infectious agents seem to evolve with speed, evading treatment methods. What are we facing and how do these organisms change so quickly?

Infectious disease can be classified into two broad categories based on the infectious agent: bacterial or viral. Bacteria are single-celled organisms that live in nearly every environment on the planet including in and on the human body. Most bacteria associated with humans are beneficial and help with daily functions like digestion and protection. Other versions (strains) of bacteria are pathogenic, meaning they can cause illness or harm. If pathogenic bacteria enter the body, they may temporarily escape the body's immune system. Once recognized, the body's immune response attacks invading bacterial cells. Most healthy individuals will be able to fight off a bacterial infection, often with the help of an antibiotic. Antibiotics weaken the bacteria by interfering with its ability to carry out functions like protein synthesis and cell division.

In recent years there has been an increase in bacteria that are resistant to the effects of antibiotics, such as the antibiotic-resistant form of *Staphylococcus aureus*, better known as MRSA. Bacteria reproduce quickly, copying their DNA before each cell division. In some cases, the copying process introduces small DNA changes. By chance, these changes may make the bacteria more resistant to a particular antibiotic. If these bacteria spread to other individuals, then a strain with antibiotic resistance has formed. As additional changes occur, the bacteria may become resistant to a wide range of antibiotics (a "super-bug"), becoming difficult to effectively treat.

In contrast to bacteria, viruses are small packages of genetic material that infect and take-over a cell, converting it to a virus-producing factory. The take-over may occur

Similarities and differences between bacteria and viruses connects with the Biology course as part of COS objective 9. Discussions about mutation in both organisms and how it leads to diversity useful for both detection and treatment could be explored in a Genetics course under COS objectives 2 and 10. In the Career/Tech Intro to Biotechnology courses, infectious disease could be explored under COS objectives 11, 13, 14 and 15.

immediately after the individual is exposed, as happens with the flu, leading quickly to symptoms. Other viruses (e.g. the herpes simplex virus 1 that leads to cold sores) cause a delayed infection with symptoms appearing weeks, months or even years after exposure. Delayed infection viruses hide their genetic material in the cell until conditions are optimal for the virus to reproduce itself. Unlike bacteria, viral infections cannot be treated with antibiotics, although antiviral medications, such as Tamiflu, may be helpful in certain instances.

Viruses reproduce very quickly once activated and like bacteria randomly change their genetic material, often leading to new strains. In addition, if two viruses simultaneously infect the same organism, their genetic information may mix, leading to a completely new strain. This is what occurred with the 2009 novel H1N1 influenza virus. Studies have shown that 2009 H1N1 contains genetic material from pig- bird- and human-based flu viruses.

Understanding the genetic and molecular basis of these organisms allows scientists to develop better diagnostic test, treatments and preventatives. Although the genomes of pathogens have the capability to change rapidly, the genomes are small and often change in semi-predictable ways. Scientists may never be able to cure the flu or common cold, but through genetics and biotechnology more accurate and faster diagnostics can be made.



Noninvasive Prenatal Diagnosis

Prenatal diagnosis involves the use of tests during pregnancy to determine whether a fetus is affected with a particular disorder. These tests have been a part of prenatal medicine for over 30 years. Testing methods vary both in level of invasiveness to the fetus as well as the degree of accuracy. Generally, a set of non-invasive screening methods - such as maternal serum analysis or ultrasound - are initially performed. Suspicious results are followed up with more invasive diagnostic testing e.g. amniocentesis or chorionic villus sampling (CVS). These invasive approaches obtain amniotic fluid and/or fetal cells that are then biochemically or genetically analyzed. Genetic tests may be genome wide - such as karyotyping or array comparative genome hybridization (see page 30) - or more narrow in scope, e.g. testing a single gene. Both amniocentesis and CVS carry a small but significant risk of miscarriage.

Scientists have recently developed a testing method that is both noninvasive and diagnostic. In the 1990s it was discovered that fetal DNA crosses the placenta into the maternal bloodstream. Relatively straightforward techniques have been developed to isolate and analyze this DNA, beginning as early as seven weeks gestation. This test can be performed several weeks earlier than conventional techniques and carries no risk to the health of the fetus. As a result, a larger number of pregnant women may choose to undergo prenatal diagnosis. In 2009, two companies announced plans to introduce this form of noninvasive prenatal diagnosis into the clinic. Initially only trisomy 21 will be diagnosed, although as the technology matures it will likely be applied to other genetic disorders.

Whether this method ultimately replaces CVS and amniocentesis will depend upon the sensitivity and specificity of the testing. However a number of significant ethical issues are associated with safer, earlier prenatal diagnosis. For example, by offering early noninvasive diagnosis, will there be increased social pressure to have the test and terminate an “abnormal” pregnancy? What or who decides the definition of “abnormal”? As the genetic

Prenatal diagnosis is a standard part of discussions around egg and sperm formation and the abnormalities that can occur during meiosis. The advent of noninvasive techniques is an exciting addition for Biology (COS objectives 6 and 8), Genetics (COS objective 4) and the Career/Tech Introduction to Biotechnology (COS objective 5). The application of this new technology to health and society links to classroom conversations in AP Biology (“Science, Technology and Society”) and Health (COS objectives 5 and 6). Clearly, there are a number of ethical concerns related to noninvasive prenatal testing. Depending on the context of the conversation and the maturity of the class, these questions may be appropriate for exploration and detailed discussion.

components of many disorders become better understood, would noninvasive diagnostic testing allow parents, with only a blood test to identify mild, adult-onset disorders, as well as nonmedical traits such as eye color?



Personal Genome Analysis

The past few years have seen the rise of genomics research aimed towards sequencing groups of individuals, such as the “PGP-10”, ten individuals who have volunteered to share their DNA sequences, medical records and other personal information as part of the personal genomes project (PGP). The public profiles of the PGP-10 are freely available online at <http://www.personalgenomes.org/>. An additional large-scale genome sequencing project is the 1000 Genomes Project, an international research collaboration that hopes to sequence the genome of approximately 1200 individuals from across the globe. Sequencing such a large number of individuals will create an index of genetic variation including previously unidentified “rare variants”, genetic changes which scientists increasingly believe are responsible for much of the genetic influence on disease.

As an initial step in the direction of personalized, commercially available genomic sequencing, several companies have begun offering consumer genomics testing. Four companies (Navigenics, deCODEme, Pathway Genomics and 23andme) offer a similar product, namely a read-out of between 500,000 and 1,000,000 variable regions from across the genome. A small but increasing proportion of these variable regions has identified connections to ancestry, physical traits or disease risk, although the predictive value for medical decisions of many of these traits remains marginal or unclear.

The cost of this personal analysis varies between \$400 and \$2,500. Two additional companies (Knome and Illumina) offer to sequence the entire 3 billion base pairs of an individual’s genome for between \$48,000 and \$100,000.

In addition to genome-wide analysis, consumer genomics testing is available for individual genes, such as the *ACTN3* genetic variant involved in muscle strength and sprint ability. A number of companies offer parents genetic testing on their children, in the hopes of identifying characteristics linked to future careers.

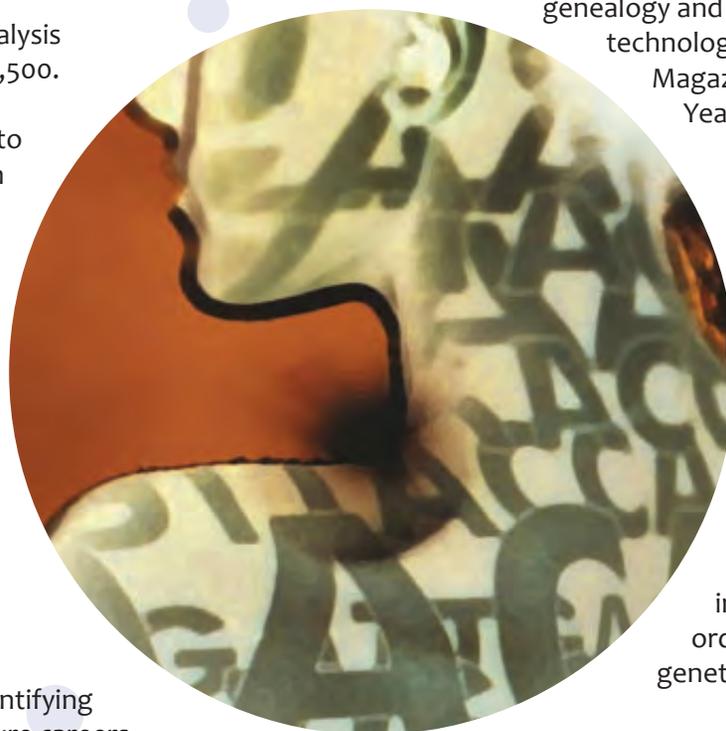
The first wave of personal genome studies offered direct-to-consumer should be a component of a Genetics course as part of COS objective 10 regarding ethical, social and legal implications from the Human Genome Project. The availability of personal information from the PGP-10 is also fertile ground for a discussion on the implications of genetic information. These topics can also be incorporated into a Biology course under COS objective 8 - significant contributions of biotechnology to society, the Career/Tech Intro to Biotechnology (COS objective 14) and an AP Biology course as part of the general theme “Science, Technology and Society”. Outside the traditional science classroom, this could form the basis of an excellent conversation with students in Health (COS objective 6), and the Career/Tech electives Foundations of Health Sciences (COS objective 10) and Health Informatics (COS objective 5) outlining valid and essential information for the safe use of consumer goods and health products.

Such programs are poor predictors of athletic aptitude, intelligence or musical or artistic talent. Much of the genetic and environmental influences on these traits are still unknown.

There is little data regarding the response of people who have received information about their genetic risk factors from one of these consumer genomic companies. At the same time, there is a growing recognition among personal genomic stake-holders that consumer genomics may provide a positive impact on an individual’s life and actions even if its direct health benefit is uncertain or marginal.

Regardless, there appears to be a strong consumer appetite for genetic information related to both genealogy and disease risk - the underlying technology was named Time Magazine’s 2008 “Invention of the Year.”

Even so, a number of scientists and health care providers have argued that these services are akin to “practicing medicine without a license”. The American College of Medical Genetics has issued a statement recommending “a knowledgeable health professional should be involved in the process of ordering and interpreting a genetic test.”



Personalized Medicine

At its core, personalized medicine uses information about a person's genetic background to tailor strategies for the detection, treatment or prevention of disease. This may include genetic screening tests to identify susceptibility to disease or more precisely pinpoint existing conditions. It may also be used to guide pharmaceutical choices, highlighting the brand and dose of medication best suited for a patient. The goal of personalized medicine is to help physicians and their patients identify the best course of action to prevent or manage a disease based upon the patient's genetic and environmental profile.

Drawing an analogy from the world of fashion, personalized medicine is the equivalent of a custom-made suit or outfit, designed with an individual's unique body measurements. This type of tailored approach provides a much better fit than purchasing something "off the rack."

As has already been noted in this guide, people vary from one another in many ways – what they eat, their lifestyle, the environmental factors to which they are exposed, and variations in their DNA. Some portion of this genetic variation influences our risk of getting or avoiding specific diseases. Certain changes in the DNA code influence the course of disease, impacting the age of onset for symptoms or the speed of progression. Genetic variation also contributes to differences in how drugs are absorbed and used by the body (see the section on pharmacogenomics on page 39).

This newfound knowledge is rapidly moving into the clinical setting. At the forefront are a series of drugs such as Gleevac™, Herceptin™ and Iressa™ known to be most effective in people with a specific genetic profile (set of genetic variants). Straightforward genetic tests are performed to identify who will benefit from these medications. At the same time, more precise diagnostic tests are in development that better classify disease subtypes or progression. The information identified in our genome will help develop a lifelong plan of health



The implications of personalized medicine impacts biology-based science courses, Health Education and pre-healthcare options at the high school level. Biology COS objective 8 and AP Biology theme "Science, Technology and Society" discuss significant contributions of biotechnology to society. Diagnosing genetic variants that increase the risk of human disease is a key focus of the Genetics COS objectives 9 and 10, particularly as it explores the ongoing impacts from the Human Genome Project and their application to disease. At the Health level, COS objective 5 asks students to evaluate negative and positive impacts of technology on health. Personalized medicine is an excellent candidate for this discussion, as well as showing application to the Career/Tech courses Introduction to Pharmacy (COS objectives 9 and 11) and Intro to Biotechnology (COS objectives 11 and 14).

maintenance tailored to our unique genetic profile. For an overview of current medical approaches based on genetic information, see the table "Selected Personalized Medicine Drugs, Treatments and Diagnostics as of March 2009" on pages 40-41.

One of the "holy grails" in personalized medicine is the so-called \$1,000 genome – the ability to sequence a human's genetic information at an economically feasible price. Recent advances in sequencing technology (highlighted in "New Findings" on page 8 and discussed in detail on page 22) are steadily moving the field closer to this figure. In addition to issues of cost, there are other challenges to personalized medicine, including concerns about patient privacy, confidentiality and insurability after taking a genetic test. Will the knowledge that specific genetic variation increases disease risk lead to greater or reduced prejudice or discrimination? How will access to genetic testing and personalized medicine be equitable? Does our current healthcare system need to change in light of this genetic approach and if so, which new model will be best?

Pharmacogenomics

Pharmacogenomics deals with how a patient's specific genetic variation affects the response to certain drugs. In part, the genetic variation among individuals helps explain why one drug may work spectacularly in one person, not at all for another and produce harmful side effects in a third. For example, variation in the *CYP2C9* and *VKORC1* genes impact whether someone is likely to develop a dangerous reaction to warfarin, a blood-thinning medication often prescribed for people at risk for blood clots or heart attacks.

A genetic test that identifies those susceptible to that reaction has now been developed to help doctors adjust warfarin doses based on each patient's genetic profile. For an overview of current pharmacological approaches based on genetic information, see the table "Selected Personalized Medicine Drugs, Treatments and Diagnostics as of March 2009" on pages 40-41. In addition, there are over 200 pharmaceutical products that either recommend genetic testing or point to the influence of genetic variability on the drug's response.

Pharmacogenomics has most rapidly developed in the field of cancer. For example, the *HER2* receptor, often found on the surface of a cell, helps regulate when the cell divides and grows. In many instances of breast cancer, the *HER2* receptor is present at very high levels, leading to increased cell growth and tumor formation. In these cases, the anti-cancer drug Herceptin™ is added to the patient's treatment plan where it increases the efficacy of chemotherapy.

Molecular testing is needed because only 25% of breast cancer patients will see any benefit from Herceptin™ -- the rest should be given another treatment. In a similar manner, Gleevec™ and Erbitux™ may be respectively prescribed for specific forms of chronic myeloid leukemia and colorectal cancer. Both medications prevent tumor cells from continuing growth but each operates in a very pathway-specific process that is unique to a subset of each cancer type. This type of therapy based on molecular targets is slowly but surely gaining in success as additional genetic pathways for disease are identified.



The implications of pharmacogenomics as a part of personalized medicine impact Health Education as well as Biology-based courses. Biology COS objective 8 and AP Biology general theme "Science, Technology and Society" discusses significant contributions of biotechnology to society. Diagnosing genetic variants that lead to specific drug recommendations is also a part of the Genetics COS objectives 9 and 10, particularly as it explores the ongoing impacts from the Human Genome Project and their application to disease. At the Health level, COS objectives 5 and 6 address negative and positive impacts of technology on health the safety of health products and like personalized medicine, pharmacogenomics is an ideal discussion topic. Classroom discussions concerning pharmacogenomics would clearly also be appropriate in the Career/Tech Intro to Pharmacy (COS objectives 9 and 11) and Intro to Biotechnology (COS objectives 1, 11 and 14) courses offered to Alabama students.

There are over 200 pharmaceutical products that either recommend genetic testing or point to the influence of genetic variability on the drug's response.

Table 1: Selected Personalized Medicine Drugs, Treatments, and Diagnostics as of March 2009*

Therapeutic product label contains pharmacogenomic information as:

- Information only
- Recommended
- Required

THERAPY	BIOMARKER/TEST	INDICATION
Herceptin® (trastuzumab) Tykerb® (lapatinib)	HER-2/neu receptor	Breast cancer: "...for the treatment of patients with metastatic breast cancer whose tumors over-express the HER2 protein and who have received one or more chemotherapy regimens for their metastatic disease."
Pharmaceutical and surgical prevention options and surveillance	<i>BRCA 1,2</i>	Breast cancer: Guides surveillance and preventive treatment based on susceptibility risk for breast and ovarian cancer.
Tamoxifen	Aviara Breast Cancer Index SM (<i>HOXB13, IL17BR</i>)	Breast cancer: Calculates a combined risk analysis for recurrence after tamoxifen treatment for ER-positive, node-negative breast cancer.
Chemotherapy	Mammostrat®	Breast cancer: Prognostic immunohistochemistry (IHC) test used for postmenopausal, node negative, estrogen receptor expressing breast cancer patients who will receive hormonal therapy and are considering adjuvant chemotherapy.
Chemotherapy	MammaPrint®	Breast cancer: Assesses risk of distant metastasis in a 70 gene expression profile.
Coumadin® (warfarin)	<i>CYP2C9</i>	Cardiovascular disease: "an increased bleeding risk for patients carrying either the <i>CYP2C9</i> *2 or <i>CYP2C9</i> *3 alleles."
Coumadin® (warfarin)	<i>VKORC1</i>	Cardiovascular disease: "Certain single nucleotide polymorphisms in the <i>VKORC1</i> gene (especially the -1639G>A allele) have been associated with lower dose requirements for warfarin."
Coumadin® (warfarin)	PGx Predict TM : Warfarin	Cardiovascular disease: Determines <i>CYP2C9</i> and <i>VKORC1</i> genotypes to predict likelihood of adverse events with warfarin therapy.
Coumadin® (warfarin)	Protein C deficiencies	Cardiovascular disease: Hereditary or acquired deficiencies of protein C or its cofactor, protein S, has been associated with tissue necrosis following warfarin administration.
Pharmaceutical and lifestyle prevention options	Familion® 5-gene profile	Cardiovascular disease: Guides prevention and drug selection for patients with inherited cardiac channelopathies such as Long QT Syndrome (LQTS), which can lead to cardiac rhythm abnormalities.
Statins	PhyioType SINM	Cardiovascular disease: Predicts risk of statin-induced neuro-myopathy, based on a patient's combinatorial genotype for 50 genes.
Atorvastatin	<i>LDLR</i>	Cardiovascular disease: "Doses should be individualized according to the recommended goal of therapy. Homozygous Familial Hypercholestremia (10-80mg/day) and heterozygous (10-20mg/day)."
Camptosar® (irinotecan)	<i>UGT1A1</i>	Colon cancer: "Variations in the <i>UGT1A1</i> gene can influence a patient's ability to break down irinotecan, which can lead to increased blood levels of the drug and a higher risk of side effects."
Erbix® (cetuximab) Gefitinib Vectibix® (panitumab)	<i>EGFR</i> expression	Colon cancer: "Patients enrolled in the clinical studies were required to have...evidence of positive <i>EGFR</i> expression using the DakoCytomation <i>EGFR</i> pharmDx TM test kit." <i>EGFR</i> positive individuals are more likely to respond to the drug than those with reduced <i>EGFR</i> expression.
Erbix® (cetuximab) Gefitinib Vectibix® (panitumab)	<i>KRAS</i>	Colon cancer: Certain <i>KRAS</i> mutations lead to unresponsiveness to the drug.
Erbix® (cetuximab) and Vectibix® (panitumab) Fluorouracil Camptosar® (irinotecan)	Target GI TM	Colon cancer: Provides information of the expression of key molecular targets— <i>KRAS</i> , <i>TS</i> , and <i>TOPO1</i> —to guide therapy.
Tagretol (carbamazepine)	<i>HLA-B*1502</i>	Epilepsy and bipolar disorder: Serious dermatologic reactions are associated with the <i>HLA-B*1502</i> allele in patients treated with carbamazepine. "Prior to initiating Tegretol therapy, testing for <i>HLA-B*1502</i> should be performed in patients with ancestry in populations in which <i>HLA-B*1502</i> may be present."
Immunosuppressive drugs	AlloMap® gene profile	Heart transplantation: Monitors patient's immune response to heart transplant to guide immunosuppressive therapy.
Ziagen® (abacavir)	<i>HLA-B*5701</i>	HIV: "Patients who carry the <i>HLA-B*5701</i> allele are at high risk for experiencing a hypersensitivity reaction to abacavir. Prior to initiating therapy with abacavir, screening for the <i>HLA-B*5701</i> allele is recommended."
Selzentry® (maraviroc)	CCR5 receptor (1)	HIV: "Selzentry, in combination with other antiretroviral agents, is indicated for treatment experienced adult patients infected with only CCR5-tropic HIV-1 detectable..."

Budesonide	IBD Serology 7	Inflammatory bowel disease: Identifies subset of patients who will benefit from budesonide.
Gleevec® (imatinib mesylate)	<i>BCR-ABL</i>	Leukemia: “Gleevec® (imatinib mesylate) is indicated for the treatment of newly diagnosed adult and pediatric patients with Philadelphia chromosome positive [indicated by presence of <i>BCR-ABL</i>] chronic myeloid leukemia (CML) in chronic phase.”
Dasatinib	Philadelphia Chromosome	Leukemia: “Dasatinib is indicated for the treatment of adults with Philadelphia chromosome-positive acute lymphoblastic leukemia (Ph+ ALL) with resistance or intolerance to prior therapy”
Busulfan	Philadelphia Chromosome	Leukemia: “Busulfan is clearly less effective in patients with chronic myelogenous leukemia who lack the Philadelphia (Ph1) chromosome.”
Purinethol® (mercaptopurine) Thiaguanine Azathioprine	TPMT	Leukemia: Guides adjustment of dose in treatment of acute lymphoblastic leukemia: “Patients with inherited little or no thiopurine S-methyltransferase (TPMT) activity are at increased risk for severe Purinethol toxicity from conventional doses...”
Tarceva® (erlotinib)	<i>EGFR</i> expression	Lung cancer: The test determines patients most likely to respond.
Capecitabine	DPD	Multiple cancers: “Rarely, unexpected severe toxicity (e.g., stomatitis, diarrhea, neutropenia and neurotoxicity) associated with 5-fluorouracil has been attributed to a deficiency of dihydropyrimidine dehydrogenase (DPD) activity.”
Pharmaceutical and surgical treatment options and surveillance	<i>MLH1, MSH2, MSH6</i>	Multiple cancers: Guides surveillance and preventive treatment based on susceptibility risk for colon and other cancers.
Chemotherapy	CupPrint™	Multiple cancers: Determines cancer classification for tumors of unknown primary origin.
Chemotherapy	Aviara CancerTYPE ID®	Multiple cancers: Classifies 39 tumor types from tumors of unknown primary origin, using a gene expression profile.
Elitek® (rasburicase)	G6PD deficiency	Multiple cancers: “Rasburicase administered to patients with glucose- phosphate dehydrogenase (G6PD) deficiency can cause severe hemolysis. ... It is recommended that patients at higher risk for G6PD deficiency ... be screened prior to starting ELITEK therapy.”
Drugs metabolized by CYP P450	Amplichip® <i>CYP2D6/CYP2C19</i>	Multiple diseases: FDA classification 21 CFR 862.3360: “This device is used as an aid in determining treatment choice and individualizing treatment dose for therapeutics that are metabolized primarily by the specific enzyme about which the system provides genotypic information.”
2C19: Celecoxib, Codeine, Diazepam, Esomeprazole, Nelfinavir, Omeprazole, Pantoprazole, Rabeprazole, Voriconazole 2D6: Acetaminophen, Aripiprazole, Atomoxetine, Carvedilol, Cevimeline hydrochloride, Clozapine, Fluoxetine HCl, Fluoxetine HCl and Olanzapine, Metoprolol, Propranolol, Propafenone, Protriptyline HCl, Risperidone, Tamoxifen, Terbinafine, Thioridazine, Timolol maleate, Tiotropium bromide inhalation, Tolerodine, Tramadol, Venlafaxine		
Rifampin Isoniazid Pyrazinamide	NAT	Multiple diseases: N-acetyltransferase slow and fast acetylators and toxicity- “slow acetylation may lead to higher blood levels of the drug, and thus, an increase in toxic reactions.”
Rituximab	PGx Predict™: Rituximab	Non-Hodgkin’s lymphoma: Detects CD-20 variant (polymorphism in the IgG Fc receptor gene <i>FcgRIIIa</i>) to predict response to cancer drug rituximab.
Celebrex® (celecoxib)	<i>CYP2C9</i>	Pain: “Patients who are known or suspected to be P450 2C9 poor metabolizers based on a previous history should be administered celecoxib with caution as they may have abnormally high plasma levels due to reduced metabolic clearance.”
Risperdal® (risperidone) Zyprexa® (olanzapine)	PhyzioType PIMS	Psychiatric disorders: Predicts risk of psychotropic-induced metabolic syndrome, based on a patient’s combinatorial genotype for 50 genes.
Gleevec® (imatinib mesylate)	<i>c-KIT</i>	Stomach cancer: “Gleevec® is also indicated for the treatment of patients with Kit (<i>CD117</i>) positive unresectable and/or metastatic malignant gastrointestinal stromal tumors (GIST).”

*This list is not intended to be comprehensive but reflects commonly used or available products as of March 2009. Some products, for which the FDA recommends or requires pharmacogenomic testing or which have pharmacogenomic information in their label, are listed at the FDA’s Web site (http://www.fda.gov/cder/genomics/genomic_biomarkers_table.htm). Other listed products that are novel, and/or that address large populations, have been identified via websites and public announcements.

Indications in quotes are taken from the therapeutic product label.

BCR-ABL = breakpoint cluster region – Abelson
BRCA 1,2 = breast cancer susceptibility gene 1 or 2
c-KIT = tyrosine kinase receptor
CYP = cytochrome P450 enzyme

DPD = dihydropyrimidine dehydrogenase
G6PD = glucose 6 phosphate dehydrogenase
HER2 = human epidermal growth factor receptor 2
NAT = N-acetyltransferase

TOP01 = topoisomerase 1
TPMT = thiopurine S-methyltransferase
TS = thymidylate synthase
UGT1A1 = UDP-glucuronosyltransferase 1A1

Recombinant DNA and Genetic Engineering

For centuries, humans have used selective breeding techniques to modify the characteristics of both plants and animals. Typically, organisms with desired traits like a high grain count, specific petal color or fragrance, consistent milk production or ability to herd livestock have been chosen to pass those traits to the next generation. These breeding practices, while very successful, require a large number of generations to yield the desired results. In addition, only traits that are naturally expressed in a species can be selected. For example, traditional breeding methods do not allow characteristics to be transferred from a plant to an animal.

Research during the last one hundred years has identified the relationship that exists between physically observed traits and the genetic information that codes for those traits. This understanding has been coupled with modern molecular laboratory techniques to transfer certain traits expressed in one species into a different (and maybe very distant) species. Scientists can modify the DNA of bacteria, plants and animals to add genetic information (and the associated characteristics) from a different organism. This process has historically been called genetic engineering but more recently is referred to as recombinant DNA technology or genetic modification.

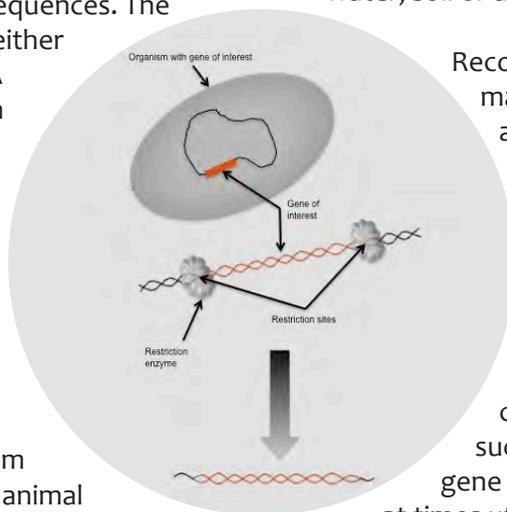
To make a recombinant organism, the gene of interest must first be isolated from the initial donor organism. To isolate the gene, scientists use restriction enzymes, proteins that can be thought of as molecular scissors that cut DNA at specific nucleotide sequences. The restriction enzymes cut the DNA on either side of the gene of interest. The DNA fragment containing the gene is then ligated (fused) into a different piece of DNA called a vector. The vector serves as a mechanism to carry the gene of interest into the host. It often includes additional genetic information such as selectable markers and genetic signals that control when and where it will be expressed. The vector is then introduced into a single host cell. From this cell, an entire organism, plant or animal is grown.

Recombinant DNA offers an excellent way to re-emphasize central dogma (the information in DNA is transcribed into RNA and then translated into protein) in the context of key molecular biology techniques, e.g. restriction enzyme digestion and DNA transformation. This approach of combining concept with application can be successfully incorporated into a number of life science as well as career/tech courses, many of which mention genetic engineering by name. This includes Biology (COS objective 8), Genetics (COS objectives 7 and 9), AP Biology (general themes “Relationship of Structure to Function” and “Science, Technology and Society”), Health (COS objective 5), Introduction to Agriscience (COS objective 16), and Introduction to Biotechnology (COS objectives 9, 13 and 14).

The organism must be tested to make sure the gene is functioning correctly and the organism is exhibiting the desired trait. Multiple generations are grown and tested before the crop, therapeutic drug or sensor is made commercially available.

Since the first recombinant DNA molecule was created in 1973, the technology has been used across a wide variety of fields:

- amending crops such as corn, soybean and rice, adding pest or herbicide resistance, or increasing nutrient content (see Agricultural Applications, page 25)
- modifying bacteria by adding genes that produce enzymes used in industry (Chymosin™ - used for making cheese)
- producing therapeutic products such as human insulin (Humulin™), blood clotting factors (rFVII™) and components of the immune system (Enbrel™)
- developing biosensors to identify toxins in the water, soil or air



Recombinant DNA forms the core of many key biotechnology applications and continues to result in new approaches that impact agriculture, healthcare and the environment. The technology is also at the core of gene therapy, a series of techniques aimed at introducing the correct version of a gene into the cells of a patient. Gene therapy is a complicated process, with only limited success to date. Silencing an overactive gene is a related form of therapy that

at times utilizes recombinant DNA. More information about this approach, known as RNAi, can be found on page 46.

Stem cells

Stem cells can be thought of as “master cells”, the raw materials from which a complete individual is crafted. The power of a stem cell lies in its “pluripotency” - the ability to divide and develop (differentiate) into any one of the 220 various types of cells found in the body. As cells differentiate, they lose this ability; a liver cell for example, can only renew itself to form more liver cells - it cannot become lung or brain.

Because of this pluripotency, stem cells have great medical potential. They could be used to recreate insulin-producing cells in the pancreas to treat type I Diabetes, to repopulate neurons destroyed due to Parkinson’s disease or to replace cells lost in spinal cord injuries. In the laboratory, stem cells have been used to successfully treat animals affected with paralysis, muscular dystrophy, Parkinson’s disease and sickle cell anemia.

Multiple types of stem cells have been identified or developed. Embryonic stem cells (ES cells) were the first category discovered. These cells are fully pluripotent, but only found in young embryos (the stage of human development from conception to eight weeks gestation). Because the process to collect ES cells destroys the embryo, some religious groups are opposed to their use.

In the tissues of many developed organs, scientists have identified so called “adult stem cells” that retain a portion of the ability to differentiate into other cell types. The primary role of adult stem cells is to maintain and repair the tissue in which it is found. For example, bone marrow contains adult stem cells, which can give rise to all the types of blood cells. This is why a bone marrow transplant can repopulate the blood and immune cells in a patient. It appears that adult stem cells may not have the full range of pluripotency found in ES cells, although researchers are exploring techniques to use adult stem cells for certain forms of therapy.

The concept of stem cells connects to several components of the standard Biology Course. It can be highlighted during explanation of the cell cycle (COS objective 6), although some biology curriculum models include discussions of stem cells during instruction on the Cell Theory instead (COS objective 4). In addition, exploring the similarities and differences between stem cells and differentiated cells would reinforce concepts about structure and function of cell and how specific functions are performed (COS objective 5) as well as the role of biotechnology in developing iPS cells (COS objective 8). Discussion of stem cells in relation to cell cycle is also connected to Genetics (COS objective 4) and Introduction to Biotechnology (COS objective 5). Highlighting the pros and cons of each stem cell type provides links to AP Biology (general theme “Continuity and Change”) and Health courses (COS objective 5).

Recent genetic discoveries have identified key genes that are active only in ES cells. Working in the laboratory, scientists have used this information to modify differentiated cells to “reactivate” these genes, in effect regressing the cells into pluripotent stem cells. These cells are known as induced pluripotent stem (iPS) cells and early research suggests they behave in much the same way as ES cells. Because iPS cells could be created by reprogramming a patient’s own tissues, they lack the ethical concerns posed by ES cells. In addition, because they are a genetic match, therapies using iPS cells would not be rejected by the patient’s immune system. While there are a number of technical hurdles that must be overcome before iPS cells are ready for clinical applications, several companies are beginning to explore treatment possibilities.



Studying the Genome to Understand the Sequence

Ten years ago the “completion” of the Human Genome Project (HGP) was announced with much fanfare. The published DNA sequence was akin to an operations manual or book of recipes, identifying the genetic instructions for how cells build, operate, maintain and reproduce themselves, all the while responding to varying conditions from the surrounding environment. While the completion of the HGP may have felt like the end of an era, in reality it was only the beginning. Scientists had very little knowledge of how cells utilized the information found in each genetic recipe to function and interact. Nor was there a clear understanding of how genes keep humans healthy or predispose them to disease. A representative genome had been sequenced, but how many differences would be found if peoples from around the world were compared? How did the human sequence compare to those of other organisms? Sequencing the human genome raised more questions than it answered.

Two large-scale projects aimed at expanding our understanding of the human genome have begun to answer many of these questions. The International HapMap Project was created to compare the genetic sequences of different individuals. The HapMap identifies DNA variants across the genome and examines how the variants are distributed within and across world populations. The project does not connect the variation to a specific illness, but rather provides the raw information that researchers can use to link genetic variation to disease risk.

ENCODE, the Encyclopedia Of DNA Elements, was launched to identify and classify the functional elements in the human genome that activate or silence regions of DNA. Based on preliminary data released in 2007, the majority of DNA in the human genome has some sort of functional role.



The history of and findings from the Human Genome Project are addressed in the Genetics COS objective 10. The subsequent HapMap and ENCODE studies shed light on the effects of genetic variability on adaptation (Genetics COS objective 2 and AP Biology general themes “Continuity and Change” and “Relationship of Structure to Function”) and the structure of eukaryotic chromosomes (Genetics COS objective 8). The influence of genetic change and mutation on increasing diversity is also a key concept in the HapMap study that is identified in the Biology COS under objective 8. These findings also have merit for discussion in the Career/Tech Veterinary Science (COS objective 3) and Intro to Biotechnology (COS objectives 9 and 14) courses.

HudsonAlpha has modified an existing AMSTI Science in Motion lab dealing with extracting DNA. This is a foundational activity that a Biology class would perform before exploring DNA or the findings of studies such as HapMap or ENCODE. The original lab followed a very simple protocol and left no room for inquiry or student input. The expanded lab provides students an opportunity to learn about the composition and structure of cells and their DNA. Students chose from a variety of plant and animal samples (fruits, fish, liver etc). Then, using a hands-on, inquiry based approach, the students design and make the necessary buffers to break open cell membranes and extract DNA, using everyday household materials.

This challenges the long-standing view that the human genome consists of a relatively small set of functional elements (the genes) along with a vast amount of so-called “junk” DNA that is not biologically active.

Just like the HGP, information generated from HapMap and ENCODE is freely accessible by scientists and the public around the world.

Synthetic Biology

Synthetic biology seeks to apply engineering principles to biology. It has an ultimate goal of designing and building biological systems for specified tasks (e.g. drug development, material fabrication and energy production). The field is a collaborative effort between not only engineers and biologists, but also chemists and physicists.

Synthetic biology aims to use engineering methods to build novel and artificial biological tools. This is done using an established engineering approach - defining the specification for a device or system and then using a set of standard parts to create a model that meets that specification. The basic building block is a biopart - a fragment of DNA with a specific function such as producing a protein or activating a “start/stop” switch. Bioparts are combined into devices that carry out a desired activity, like producing fluorescent protein under a given condition. Multiple devices can be connected into a system, which performs more complex, higher-level tasks.

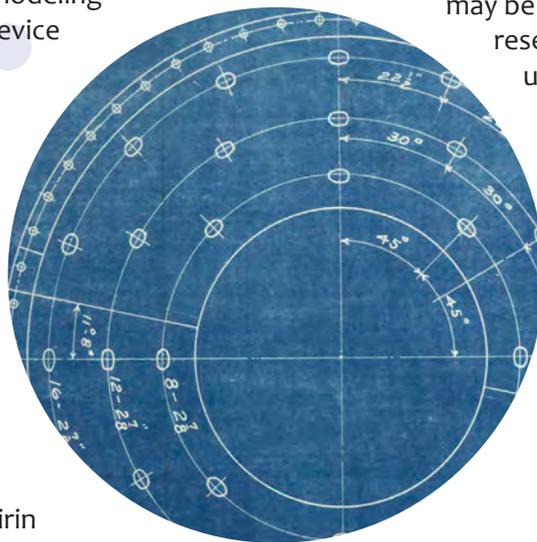
Powerful computers offer in-depth modeling and simulation to predict the behavior of the part, device or system before it is assembled. The relevant DNA instructions are then artificially synthesized and inserted into a biological cell, such as bacteria. The bacterial cell is the “chassis” or vehicle that interprets the DNA instructions. If the synthesized information is read and processed correctly, then the specification and design were appropriately crafted. If not, the original design is modified, continuing the design-modeling-testing cycle. Once complete, the device or system becomes a component created from standard bioparts, rather than constructed each time from scratch.

The rise of synthetic biology has been compared to that of synthetic chemistry, a field that developed and matured during the past century as chemists learned how to synthesize compounds that previously only existed in nature. Early examples such as dyes and medicines like aspirin gave way to the creation of plastics, semiconductors and complex pharmaceuticals.

The concepts behind synthetic biology links to the COS objective 8 for a standard Biology course, particularly as it relates to significant contributions to biotechnology. Discussion of synthetic biology also connects to the AP Biology general theme “Science Technology and Society.” Lastly, the Genetics COS objective 9 and CTE Introduction to Biotechnology COS objective 13 highlight areas of biotechnology that deal with recombinant DNA. This is a natural connection to synthetic biology, which uses recombinant DNA techniques as the cornerstone to creating the artificial bioparts, systems and devices.

Many supporters believe that synthetic biology has the potential to achieve equally important results such as producing inexpensive new drugs, developing environmental biosensors and more efficiently producing biofuels from biomass.

Given that synthetic biology involves creating novel living organisms, it isn’t surprising that security, safety and ethical concerns have been raised. Like many other “dual use technologies,” synthetic biology offers the potential for great good, but also for harm. There are concerns that the increasing accessibility of this technology may spawn a new era of “biohackers” leading to the accidental or deliberate creation of pathogenic biological components. Safety measures taken by the research community include incorporating genetic signals that prevent uncontrolled spreading outside the lab environment. It is worth noting that in many ways, these mechanisms are already in place as part of the guidelines developed for recombinant DNA techniques that are currently in use worldwide. From this perspective, the advances in synthetic biology may be viewed as a natural extension of this research, rather than a great leap into uncharted scientific territory.



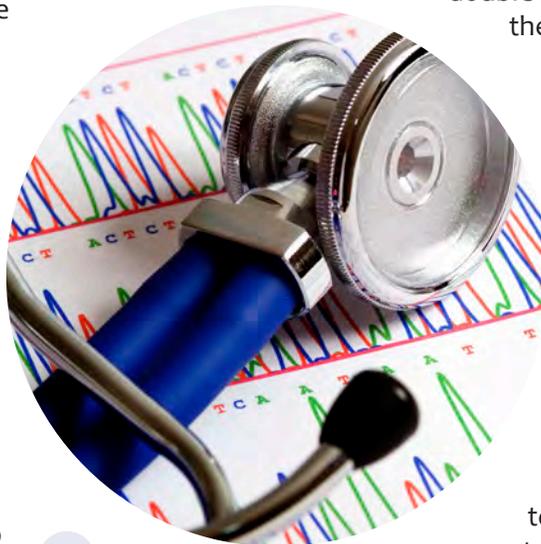
Therapeutic Approaches

Gene Therapy

Gene therapy is defined as the correction of a nonfunctioning gene responsible for causing a disease. For example, a normal (functioning) copy of the gene could be inserted into a cell to replace a nonfunctioning gene. As genes will not enter cells on their own, there must be a mechanism in place to carry the corrected gene into the body's cells. The most common mechanism (vector) is an altered form of a virus. Viruses have the capability of infecting and inserting their genetic information into cells. Researchers are able to exploit this capability of viruses while removing the viral genes responsible for causing illness.

Although the concept of gene therapy is simple in theory there are several technical roadblocks that have to be overcome for these treatments to become a reality. For gene therapy to cure a disorder, the inserted gene must remain active in the body's cells long-term. Currently it is difficult to retain the added gene through multiple rounds of cell division, making it hard to achieve successful gene therapy in actively dividing cells. In addition, it is difficult to ensure that the vector containing the therapeutic gene reaches the organs and body tissues where symptoms occur. Some of the recent successes in gene therapy research have been in ocular (eye) diseases in which the targeted body area is easily accessible.

One of the major setbacks in the gene therapy research occurred in 1999 with the death of 18-year-old Jesse Gelsinger. Jesse had a rare genetic condition called ornithine transcarboxylase deficiency (OTCD) in which a gene mutation causes an enzyme, important for the removal of nitrogen from the body, to be absent. Jesse enrolled in a clinical trial for gene therapy of OTCD aimed at determining a safe dose for treatment and documenting potential side effects. Four days after starting the treatment, Jesse passed away from multiple organ failure thought to have been triggered by an immune response to the viral vector.



Gene Therapy, RNAi and their role in altering/silencing protein synthesis should be discussed in the Genetics course as a part of COS objective 7. The potential as treatment for disease, is described under Genetics COS objective 10 and AP Biology under the general theme "Science, Technology and Society." It could also be incorporated into a discussion about the relationship between DNA, RNA and proteins (COS objective 8) for a Biology class or Introduction to Biotechnology course (COS objective 9).

Researchers are working to overcome many of the roadblocks described above and are making promising strides in developing safe and effective methods for introducing functional genes into the body.

RNAi

Another type of gene therapy currently being researched is RNAi. Much like turning off a light switch, RNA interference (RNAi) offers the ability to selectively silence or "turn off" the activity of a single gene. This technology has the potential to revolutionize our understanding of how genes work and offers new promise in therapy and treatment.

In addition to mRNA and tRNA found in cells, researchers in the 1990s noted an additional form of RNA composed of small double-stranded molecules. These fragments could effectively stop protein production by coordinating the destruction of the single stranded mRNA. In other words, the double stranded RNA "interfered" with the mRNA, effectively silencing the activity of the gene. Researchers have utilized the RNAi pathway to explore the effects of systematically silencing genes. Short synthetic double-stranded RNA molecules can be created in the laboratory and delivered into cells, leading to partial or complete cessation of protein production for specific targeted genes. The ability to target and deplete specific proteins has identified RNAi as a potential therapeutic pathway.

STATUTES AND SESSION LAW

Title 40 REVENUE AND TAXATION.

Chapter 9 EXEMPTIONS FROM TAXATION AND LICENSES.

40-9-34 HudsonAlpha Institute for Biotechnology.

(a) The following is hereby found and declared by the Legislature of Alabama:

(1) The lack of content in natural and bio-science education offered to students in kindergarten through high school is a nationwide problem.

(2) Such lack in curricular offerings to students will be detrimental in the long-term to the economy of the state and the welfare of the citizens during the scientific revolution now engulfing the world.

(3) The biotechnology institute can provide to education leaders of the distance learning program of the state cutting edge biotechnology curriculum recommendations and content for Alabama high schools, by providing information about cutting edge biotechnology curriculum and content to students in kindergarten through high school pursuant to the distance learning program of the state, the state course of study, and state textbooks.

(4) By educating Alabama high school students in the field of biotechnology, such students are more likely to pursue careers in the biological sciences, thereby providing the state with a better educated workforce able to support the growing biotechnology industry, in turn attracting and encouraging biotechnology companies to locate in the state and create additional challenging and rewarding job opportunities for the citizens of the state.

(5) The reputation, economic status, and educational system of the state will be further enhanced by the addition of an internationally renowned biotechnology institute that will support internationally recognized scientists and researchers, with a focus on scientific discoveries that are intended, when possible, to be proven in the state and provided by companies in the state to patients suffering from diseases.

(6) By establishing a biotechnology campus, the biotechnology institute will be in a better position to join with the economic development leaders of the state to attract biotechnology companies to the campus and to the state, thereby creating additional job opportunities for the citizens of the state.

(b) The HudsonAlpha Institute for Biotechnology, a nonprofit corporation, and any real and personal property owned by the corporation, shall be exempt from the payment of any and all state, county, and municipal taxes, licenses, fees, and charges of any nature whatsoever, including any privilege or excise tax heretofore or hereafter levied by the State of Alabama or any county or municipality thereof.

(c)(1) In exchange for the tax exemption granted in subsection (b), beginning October 1, 2008, and for each year thereafter, **the HudsonAlpha Institute for Biotechnology shall make a report to the State Board of Education detailing the curricular content in biotechnology which could enhance the state distance learning program.** This subdivision shall not apply in the event that the distance learning program is discontinued, or is no longer in existence. **Further, the HudsonAlpha Institute for Biotechnology shall report annually to the State Board of Education, the State Course of Study Committee, and the State Textbook Committee all new developments in the field of biotechnology which could be integrated into the curriculum for high school courses in science and health.**

RECOMMENDED READING FOR MORE DETAILS

DNA sequencing

The impact of next-generation sequencing technology on genetics. Mardis E.R., *Trends in Genetics*, Volume 24, Issue 3, March 2008, Pages 133-141

Rise of the machines. Gresham D. Kruglyak L., *PLoS Genetics*. 4(8):e1000134, 2008.

RNA analysis

Mapping and quantifying mammalian transcriptomes by RNA-Seq. Mortazavi A. Williams BA. McCue K. Schaeffer L. Wold B. *Nature Methods*. 5(7):621-8, 2008 Jul.

A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Sultan M. Schulz MH. Richard H. Magen A. Klingenhoff A. Scherf M. Seifert M. Borodina T. Soldatov A. Parkhomchuk D. Schmidt D. O'Keefe S. Haas S. Vingron M. Lehrach H. Yaspo ML. *Science*. 321(5891):956-60, 2008 Aug 15.

Protein analysis

Biologists initiate plan to map human proteome. Pearson H., *Nature* vol 492, 920 (April 2008)

Bioinformatics

Steady progress and recent breakthroughs in the accuracy of automated genome annotation. Brent M., *Nature Reviews Genetics* 9, 62-73 (January 2008).

Agricultural applications

Where's the Super Food? Grant B. *The Scientist*. 23(9):30, September 2009

GM Crops: The First Ten Years - Global Socio-economic and Environmental Impacts. Brookes G, Barfoot P. *Economics*. 2006

Global Status of Commercialized Biotech/GM Crops: 2007. James C. *International Service for the Acquisition of Agri-Biotech Applications (ISAAA)*. February 2008.

Cancer

<http://www.cancerquest.org/> CancerQuest is an excellent online resource that details both normal and cancer biology. It was developed as an educational outreach program by Emory University.

Comparative genomics

Approaches to comparative sequence analysis: towards a functional view of vertebrate genomes. Margulies, E. H. & Birney, E. *Nature Reviews Genetics* 9, 303-313 (2008).

Mammalian karyotype evolution. Ferguson-Smith MA & Trifonov V. *Nature Reviews Genetics* 8, 950-962 (December 2007).

Copy number variation

Structural variation in the human genome. Feuk L, Carson AR, Scherer SW. *Nature Reviews Genetics* 7, 85-97 (February 2006).

Criminal Justice

Forensic DNA Phenotyping: Regulatory Issues, Koops BJ, & Schellekens M. *Columbia Science and Technology Law Review* 158 (2008)

To Sketch a Thief: Genes Draw Likeness of Suspects, Naik G, *The Wall Street Journal*, March 27, 2009 – accessed online at <http://online.wsj.com/article/SB123810863649052551.html>

Diagnosing chromosomal disorders

Pre- and postnatal genetic testing by array-comparative genomic hybridization: genetic counseling perspectives. Darilek S. Ward P. Pursley A. Plunkett K. Furman P. Magoulas P. Patel A. Cheung SW. Eng CM. *Genetics in Medicine*. 10(1):13-8, Jan, 2008.

Epigenetics

Epigenetic mechanisms that underpin metabolic and cardiovascular diseases. Gluckman P, Hanson M, Buklijas T, Lowe F, Beedle A. *Nature Reviews Endocrinology* 5, 401-408 (July 2009)

Environmental epigenomics and disease susceptibility. Jirtle R, Skinner M. *Nature Reviews Genetics* 8, 253-262 (April 2007).

DNA methylation landscapes: provocative insights from epigenomics. Suzuki M, Bird A. *Nature Reviews Genetics* 9, 465-476 (June 2008).

Genetic information nondiscrimination act

A fact sheet and the text of the Legislative Act can be found at <http://www.genome.gov/24519851>

Keeping pace with the times--the Genetic Information Nondiscrimination Act of 2008. Hudson KL. Holohan MK. Collins FS. *New England Journal of Medicine*. 358(25):2661-3, June 19, 2008.

Genetics of Eye Color

Molecular genetics of human pigment diversity. Strum RA. *Human Molecular Genetics*. 18(Review issue 1):R9-R17, April, 2009.

Identifying genetic influence on disease

<http://www.nature.com/scitable/topicpage/Complex-Diseases-Research-and-Application-748>

<http://www.nature.com/scitable/topicpage/Genome-Wide-Association-Studies-GWAS-and-Obesity-752>

These two webpages are from an educational website known as Scitable. Developed by Nature Publishing, Scitable is a free resource for educators, students and the public that is linked to the scientific reports published by the Nature publishing group.

<http://www.genome.gov/20019523> Fact Sheet on Genome Wide Association Studies – developed as an educational resource by the National Institutes of Health National Human Genome Research Institute

Infectious disease

<http://www.nlm.nih.gov/medlineplus/infectiousdiseases.html> This website provides an overview of bacteria, viruses and the body's response to infectious agents. It is developed in partnership with the National Library of Medicine and the National Institutes of Health.

Non-invasive prenatal genetics

An offer you can't refuse? Ethical implications of non-invasive prenatal diagnosis. Schmitz D, Netzer C, Henn W. *Nature Reviews Genetics*. 10:515, August, 2009.

Practical and Ethical Considerations of Noninvasive Prenatal Diagnosis. Benn PA, Chapman AR. *Journal of the American Medical Association* 301, 2154-5, May 27, 2009.

Personal genome analysis

Research ethics and the challenge of whole-genome sequencing. McGuire A, Caulfield T, Cho MK *Nature Reviews Genetics* 9, 152-156 (February 2008).

American College of Medical Genetics Statement on Direct-to-Consumer Genetic Testing (2008) - http://www.acmg.net/StaticContent/StaticPages/DTC_Statement.pdf

Pharmacogenomics

http://www.ornl.gov/sci/techresources/Human_Genome/medicine/pharma.shtml Information on Pharmacogenomics, including links for additional information. Developed by the US Department of Energy as part of their Human Genome Project overview and application pages.

RNA interference

Exploring the Uses of RNAi — Gene Knockdown and the Nobel Prize. Bernards R *New England Journal of Medicine*. 355:2391, December 7, 2006.

Stem cells

<http://stemcells.nih.gov/info/basics/> A primer on stem cells developed by the National Institutes of Health.

Studying the Genome to Understand the Sequence

Comparing whole genomes using DNA microarrays. Gresham D, Dunham M., Botstein D., *Nature Reviews Genetics* 9, 291-302 (April, 2009).

THE SCIENCE OF PROGRESS



Contact

Neil E. Lamb Ph.D.
Director of Educational Outreach
HudsonAlpha Institute for Biotechnology
601 Genome Way
Huntsville, AL 35806
nlamb@hudsonalpha.org
hudsonalpha.org

