Biotechnology Discoveries and Applications

Extensions to high school science curriculum

The 2013 guidebook

HudsonAlpha
INSTITUTE FOR BIOTECHNOLOGY

CONTENTS

Genetic Resources to Empower Alabama Teachers

GREAT

Cell Division

a unique learning opportunity offered at no cost by the HudsonAlpha Institute for Biotechnology

Coming soon to:

Montgomery
Winter, 2014

Mobile
Fall, 2014

Huntsville
Winter, 2015

Birmingham
Fall, 2015

The GREAT Workshop provides opportunities for Alabama public high school life science educators to update genetics knowledge and discover recent scientific findings that are too new for textbooks. This second round of GREAT workshops will dig into cell division by working through the following:

- How do chromosomes behave during cell division?
- What regulates the cell cycle?
- How does cancer occur?
- How is DNA sequencing changing the way cancer is treated?
- How can I teach these concepts to my students?

In two full days of small group concurrent sessions and talks by dynamic speakers, teachers will learn about recent findings and applications that relate to health, agriculture and the environment.

Visit www.hudsonalpha.org/GREAT to register!

The GREAT Workshop is open to Alabama accredited, public high school life science teachers and is made possible through support from the State of Alabama.

Genetic Technologies for Alabama Classrooms

a two-week teacher academy

GTAC® is an intensive two-week professional development academy for high school biology teachers held at the HudsonAlpha Institute for Biotechnology in Huntsville, AL. 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.

June 15-27, 2014
Huntsville, AL

What can I expect?

- Practice using hands on activities
- Hear from scientists involved in cutting edge biotechnology research
- See and use modern biotechnology equipment and laboratories
- Implement learning through individual and group projects
- Create and present a professional poster showcasing GTAC concepts

Walk away with:

- 80+ Professional Learning units
- Stipend
- Toolkit of equipment and resources
- Updated genetic content knowledge
- Modern applications in biotechnology
- Network of teachers from across the state

To apply, visit www.hudsonalpha.org/GTAC

Applications will be accepted November 1, 2013 - January 31, 2014

Digital Resources

iCell®

Why use flat images from a textbook when your students can explore cell structure in 3D? HudsonAlpha iCell®, one of Apple's featured biology apps on the iTunes® Education market, allows students to explore representative plant, animal and bacteria cells with vivid 3D models. iCell is available on multiple platforms and has been downloaded over 300,000 times by students and educators around the world.

iCell is available on Apple® and Android® devices, Windows 8® tablets, as a downloadable program for Mac® and Windows®, and at www.icell.hudsonalpha.org.

The Progress of Science®

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 www.timeline.hudsonalpha.org.

Digital Resources

Build your own genome, or walk ours. GenomeCache® combines the challenge of a scavenger hunt with the human genome.
It allows anyone to create up to 20 walkable paths that explore the human genome with over 150 challenging questions, a leaderboard and themed paths. GenomeCache combines clues, fun facts and trivia questions to create an engaging learning experience.

GenomeCache is available on iPad®, iPhone®
and www.genomecache.hudsonalpha.org.

Touching Triton

Touching Triton is an online educational activity focused on building understanding of common complex disease risk. This serious game uses the engaging storyline of long-term space flight to highlight how risk is influenced by factors from family history, environment and an individual's genomic profile. By engaging students in an interesting storyline and graphical interface, we aim to create an activity that is both enjoyable and educational.

Biotechnology Discoveries and Applications **2013**

How this guide is arranged

Recent research findings are grouped on pages nine through nineteen and provide a quick update on the genetics/genomics/biotechnology field. **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.

Where appropriate, an acknowledgement of research occurring at HudsonAlpha is given:

The  symbol identifies those connections.

EXECUTIVE SUMMARY

A recent podcast featuring Francis Collins, M.D., Ph.D., director of the National Institutes of Health, highlighted the benefits of The Human Genome Project, including the exponential economic return on dollars invested. Collins also noted that since World War II, more than half of the economic growth in this country has been on the basis of science and technology. In this light, he believes society should value teaching as a career that our whole future depends upon.

He's right. Ask a scientist, physician or other biotech professional about his or her path and the majority will tell the story of a teacher who inspired, challenged or otherwise set into motion a career. That is certainly true at HudsonAlpha - many of us can point to one or more educators who helped set our career trajectory. That is also a reason why educational outreach is a key part of HudsonAlpha's mission. Genomic research, educational outreach and economic development are symbiotic, each made stronger in sharing and collaborating. Educational outreach strengthens both research and business when teachers are given resources, provided opportunities to update their knowledge base and are introduced to methods that engage student participation. Together, these resources help reach students who one day may recount their own story about that teacher who made science come alive.

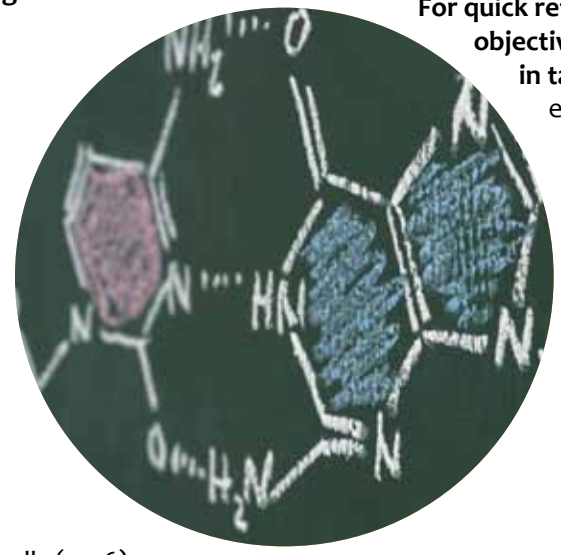
By investing in teachers, the future becomes a more productive, successful place of life-enhancing and enabling technologies. This is part of the aim of our annual *Biotechnology Discoveries and Applications* educator guidebook: **to provide educators with information related to the recent advances in genetics and biotechnology, allowing them to share these findings with their students.**

The guidebook is divided into two sections: research highlights and foundational concepts. This year's research section highlights 43 new discoveries, including articles on:

- understanding how genetic mutations influence cancer development (p. 10)
- genomic studies in wheat that identify natural resistance to rust disease (p. 13)
- the impact of diet and gut microbes on weight (p. 14)
- chromosome conformation in cells (p. 16)
- balancing privacy and scientific advancement (p. 18)

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. Opened in 2008, 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.

These new findings are linked to 23 foundational topics, covered in detail beginning on page 28. Each includes the basic concepts underlying each topic and connections to relevant course of study objectives for Alabama's science, health and career technical education classes. Genomics has become an important tool across the life science landscape, entering classroom discussions ranging from agriculture and human health to bioinformatics and ethics.



For quick reference, the course of study objectives and foundational topics are listed in table format on pages 22-26. For educators who are not in Alabama, many of these course of study objectives will align across states fairly easily.

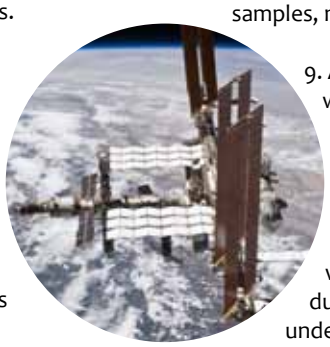
A list of suggested readings for additional exploration of many of the foundational topics can be found on page 54.

SCIENCE SNAPSHOTS

a quick rundown of 10 genetics and biotech stories

1. New types of engineered bacteria can produce important precursors of gasoline-like biofuels. Working with laboratory strains of *Escherichia coli*, scientists modified genes that produce eight-carbon fatty acids known as octanoate. This can be converted into octane, an important component of gasoline. Oil refineries currently produce similar molecules from crude oil.

2. The full genetic sequence of 1,092 healthy individuals has been published as part of the 1,000 Genomes Project. The most comprehensive catalog of human genetic variation to date, it includes 38 million single nucleotide polymorphisms, 1.4 million short insertions/deletions and over 14,000 larger deletions. This collection serves as a reference to analyze DNA changes identified in individuals with genetic disorders.



3. A combination of four genes converts scar tissue from damaged hearts into healthy heart muscle. Performed as a proof of principle study in rats, the scar tissue was pretreated with a form of gene therapy that encourages blood vessel growth to the region. Subsequently, three genes were added that function as transcription factors – encoding proteins that bind DNA sequences to activate transcription in other genes. These activated genes produced proteins that reshaped the cells into healthy heart muscle.

4. In most cases of Burkitts lymphoma, a chromosomal translocation moves the gene *MYC* from its typical position on chromosome 8 to a region that leads to overexpression of the gene. As the myc protein is a transcription factor involved in mitosis, overexpression turns the lymphocytes cancerous. Scientists sequenced the exomes from several Burkitt lymphoma tumors, identifying 70 additional genes that are often mutated during the formation of this cancer. A number of these genes had never been associated with cancer, providing new targets for diagnostics and therapeutics.

HudsonAlpha researcher Dr. Shawn Levy contributed to these findings.

5. The Cancer Genome Atlas, a large collaborative research project, has analyzed the genomes of glioblastoma, ovarian, colorectal and lung adenocarcinomas and invasive breast cancer. This past year, TCGA published findings related to endometrial and kidney cancers, as well as acute myeloid leukemia. These findings improve oncologists' ability to categorize each cancer, predict the severity of disease for individual patients and identify targeted treatments based on the specific mutations present in the tumor.

6. Mice with a genetic mutation in the *Mrap2* gene gain more weight as compared to their unaffected counterparts even when both are fed the same amount of food. The protein encoded by the *Mrap2* gene is part of a large and complex pathway that regulates metabolism, appetite and food consumption. Mice with *Mrap2* mutations sequester fat rather than break it down for energy. When the human equivalent of *Mrap2* gene was analyzed across 500 obese patients, four were found to carry mutations.

7. Circadian patterns are regularly occurring, 24-hour rhythms associated with normal function like sleep-wake cycles, body temperature and hormone release. Dysfunction in these patterns is seen among patients with depression. A set of genes helps establish and maintain circadian rhythms. Research suggests these genes are improperly regulated in the brains of individuals with major depressive disorder. Why the internal clock is disrupted remains elusive, but this cellular link between altered circadian rhythms and depression offers insight for future treatment.

HudsonAlpha researcher Dr. Rick Myers participated in this work.

8. In 2003, scientists unearthed a fossilized fragment of the leg bone from a horse that lived approximately 700,000 years ago. This past year, the DNA from that bone was sequenced to assemble the ancient horse genome. This is the oldest whole genome ever sequenced. It is nearly 10 times older than previous samples, none of which dated back more than 70,000 years.

9. A research project on the International Space Station will explore the microbiomes of the station and its inhabitants. NASA has a long history of studying the microorganisms that are found in closed-environment spacecraft but this is the first systematic study using high-throughput sequencing. Air, surface and water samples from the station will be analyzed, along with multiple samples from crew members before, during and after missions. The results will help scientists understand how an astronaut's personal microbiome interacts with the unique environment of the space station and if space-related travel alters the microbial landscape.

10. The AquAdvantage Salmon, profiled in the 2011 *Biotechnology Guidebook*, moved a step closer to regulatory approval. The Food and Drug Administration concluded the genetically modified Atlantic salmon would not have significant impact on the environment. Because the fish contains a growth hormone gene from the Chinook salmon and a genetic on/off switch from the ocean pout (also a fish) it can reach market size in about 18 months instead of three years. Following a period of public comment, the FDA may decide to approve the AquAdvantage Salmon for introduction into the food supply.

REFERENCES

1. Torella et al., "Tailored fatty acid synthesis via dynamic control of fatty acid elongation," *Proceedings of the National Academy of Sciences* (2013) 110:11290-5 doi: 10.1073/pnas.1307129110.
2. The 1000 Genomes Project Consortium, "An integrated map of genetic variation from 1,092 human genomes," *Nature* (2013) 491:56-65 doi:10.1038/nature11632.
3. Mathison et al., "In Vivo Cardiac Cellular Reprogramming Efficacy is Enhanced by Angiogenic Preconditioning of the Infarcted Myocardium With Vascular Endothelial Growth Factor," *Journal of the American Heart Association* (2012)1:e005652 doi: 10.1161/JAHA.112.005652.
4. Love et al., "The genetic landscape of mutations in Burkitt lymphoma," *Nature Genetics* (2012) 44:1321-5 doi: 10.1038/ng.2468.
5.The Cancer Genome Atlas Research Network, a. "Integrated genomic characterization of endometrial carcinoma," *Nature* (2013) 497:67-73 doi:10.1038/nature12113. b. "Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia," *New England Journal of Medicine* (2011) 368:2059-74 DOI:10.1056/NEJMOA1301689. c. "Integrative analysis of genomic and molecular alterations in clear cell renal carcinoma," *Nature* (2013) 499:43-49 doi:10.1038/nature12222.
6. Asai et al., "Loss of Function of the Melanocortin 2 Receptor Accessory Protein 2 is Associated with Mammalian Obesity," *Science* 19 July 2013:275-78 doi:10.1126/science.1233000.
7. Li et al., "Circadian patterns of gene expression in the human brain and disruption in major depressive disorder," *Proceedings of the National Academy of Sciences* (2013) 110:9950-5 doi:10.1073/pnas.1305814110.
8. Orlando et al., "Recalibrating Equus evolution using the genome of an early Middle Pleistocene horse," *Nature* (2013) 499:74-78 doi:10.1038/nature12323.
9. Barringer, C. and Abadie, L., "Microbial Creatures in Space," http://www.nasa.gov/exploration/humanresearch/Microbial-Creatures-in-Space.html#.UiirmWRVT2k
10. U.S. Food and Drug Administration, (2013, September) "Genetically Engineered Salmon" Retrieved from http://www.fda.gov/AnimalVeterinary/DevelopmentApprovalProcess/GeneticEngineering/GeneticallyEngineeredAnimals/ucm280853.htm.

Diagnostic exome sequencing

a powerful tool to uncover genetic causes for families impacted by rare disease

Exome analysis – sequencing that 2 percent of the genome containing the ~23,000 human genes – has become an established part of research projects to identify disease-causing genetic change. The technique has also gained a place in patient care, especially when symptoms suggest a genetic syndrome but previous genetic tests have been negative. Exome sequencing casts a broad net by examining every gene, rather than a handful of potential candidates. Today's sequencing technology simultaneously analyzes millions of overlapping DNA fragments, producing an enormous text file containing billions of As, Ts, Cs, and Gs. A set of detailed computational screens then identifies relevant DNA changes.

In many cases, exome sequencing has uncovered a definitive (or at least plausible) biological explanation for a patient's disorder. Early estimates suggest around a quarter of patients receive some sort of genetic diagnosis. Discovering the genetic cause of a disorder does not guarantee a treatment will be available. Often, our understanding of the biology behind the disease is incomplete and therapies don't yet exist. Even so, there is great value in providing an explanation for patient symptoms.

Every human genome contains thousands of variants, from rare to common and from inconsequential to harmful. Therefore, exome sequencing may identify so-called incidental findings, DNA changes that predispose to disease separate from the patient's current condition (e.g., mutations for adult-onset cancer or variants that uncover non-paternity within a family). Whether researchers and clinicians have a responsibility to actively search for and disclose secondary findings and if so, what should be assessed is still under discussion.

There are several challenges to exome sequencing, including significant cost and analysis time. Disorders with multiple genetic and environmental risk factors, such as diabetes or cardiovascular disease, are difficult to diagnose; many risks are still unknown and most will increase risk only slightly. Our ability to interpret genetic variants and determine clinical impact is still rudimentary and most healthcare providers lack a solid understanding of genomics. These hurdles will likely be overcome as the technology matures and experience deepens.

HudsonAlpha is leading a clinical exome sequencing project that examines children with unexplained developmental delay.

For more information, see *Identifying Genetic Influence on Disease,* on page 40.

Cancer updates

recent studies provide a better understanding of genetic mutation and its impact on cancer formation

Signatures of mutation

Many cancer-causing mutations change the DNA in characteristic, traceable ways. For example, ultraviolet light damage often results in a C to T transition. Researchers at the Welcome Trust Sanger Institute in the U.K. analyzed nearly 5 million mutations from 7,000 cancers to uncover 21 such mutational signatures. For many of the signatures, the biological process responsible for the mutation was also identified. All of the cancers contained two or more signatures, highlighting the multi-factorial nature of cancer.

Cancer-linked gene may regulate cell division

Nearly 100 years ago, Theodor Boveri suggested that abnormal mitosis could be a cause of cancer. Specifically, Boveri was referring to multipolar mitosis, where the chromosomes are pulled to more than two poles. This past year scientists identified a gene called *FAM190A*, which is mutated in nearly 40 percent of human cancers. Experimentally reducing FAM190 protein levels in cells led to multipolar mitosis and a failure to completely separate at cytokinesis - the last stage of mitosis.

Melanoma

Melanoma originates in the pigment-producing melanocytes found in the deepest layer of skin. The cancer is associated with ultraviolet light exposure, especially for those with fair skin and red hair. Scientists have identified part of the molecular link for this risk factor. The *MC1R* gene encodes a protein that controls pigment production in skin and hair follicles. Specific *MC1R* alleles lead to red hair. When cells are exposed to UV radiation, the MC1R protein also protects a tumor suppressor from destruction. For reasons still unclear, the red-hair causing *MC1R* alleles don't protect the tumor suppressor protein and the activity of cancer-causing proteins increases.

Melanoma has also recently been linked to mutations in a gene called *TERT*. This gene encodes a component of telomerase, which regulates the length of telomeres; those repeating DNA sequences found at the ends of chromosomes. The cancer-associated mutations don't occur in the protein-coding portions of the gene, but are within the promoter and create new transcription factor binding sites that increase *TERT* expression. This produces higher levels of telomerase which extends the length of telomeres, creating cells that can divide indefinitely. Found in over 70 percent of analyzed melanomas, these mutations may be one of the most common drivers of cancer.

For more information, see *Cancer,* on page 32.

REFERENCES

Alexandrov et al., "Signatures of mutational processes in human cancer," *Nature* (2013) 500:415-21 doi:10.1038/nature12477.
Patel et al., "FAM190A Deficiency Creates a Cell Division Defect," *The American Journal of Pathology* (2013) 183:296-303 doi: 10.1015/j.apath.2013.03.020.
Cao et al., "MC1R is a Potent Regulator of PTEN after UV Exposure in Melanocytes," *Molecular Cell* 51:409-22 doi: 10.1016/jmolcel.2013.08.010.
Hwang et al., "Highly recurrent *TERT* promoter mutations in human melanoma," *Science* (2013) 339:957-9 doi: 10.1126/science.1229259.

Sickle cell anemia

countering conventional wisdom on capillary blockages

When discussing genetics, sickle cell anemia is likely one of the most frequently highlighted human disorders. Inherited in an autosomal recessive manner, this disease is most often caused by a single nucleotide change in the first exon of the *HBB* gene on chromosome 11. The change substitutes the hydrophobic amino acid valine in place of the negatively charged glutamic acid that normally resides in the sixth position of the beta globin protein.

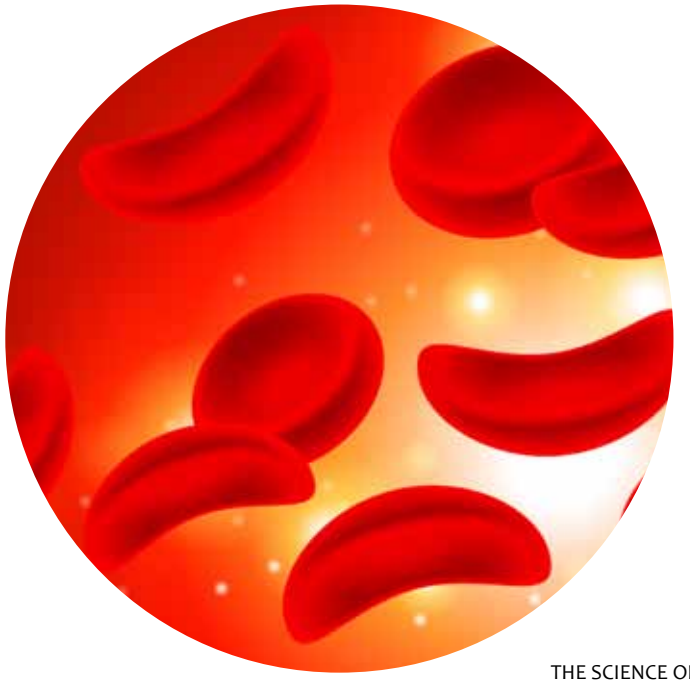
Beta globin is a critical part of hemoglobin, which carries oxygen in red blood cells. Under certain conditions, the mutated beta globin causes the hemoglobin to link into long strands. These strands distort the shape of the red blood cells, changing the properties of their cell membranes. The cells clog the vessels of limbs and organs, blocking blood flow and causing pain and tissue damage.

Historically, it was believed that the rigid, crescent-shaped red blood cells were solely responsible for the blockages. However, recent data suggests a more complex pathway is at work. Thirty years ago, researchers discovered four distinct types of sickle red blood cells. Not all of them are rigid or have the characteristic crescent shape. For example, one type of cell, known as SS2, retains the typical round shape. However, SS2 cells are significantly more sticky due to the presence of protein receptors on the cell membrane that increase adhesiveness.

Computer models, based on experimental data from real cells, allowed the researchers to determine how blockages form. The models showed that the softer SS2 cells initiate the process by sticking to capillary walls. The rigid sickled cells get caught behind the SS2 cells, stacking up like traffic after a car wreck during rush hour and leading to the painful blockages.

REFERENCE

Lei, H., and Karniadakais, G.E., "Probing vasoocclusion phenomena in sickle cell anemia via mesoscopic simulations," *Proceedings of the National Academy of Science* (2013) 110:11326-30 doi:10.1073/pnas.1221297110.



In brief

Genetic testing

Supreme court and patenting

Can you patent human genes? No - according to the recent U.S. Supreme Court decision in the case *Association for Molecular Pathology v. Myriad Genetics*. At the same time, the court reaffirmed the importance of patents as a means of creating incentives for discovery and innovation across the biotechnology industry.

A genetic testing company in Salt Lake City, Myriad Genetics, obtained multiple patents related to *BRCA1* and *BRCA2* - genes that when mutated substantially increase the risk of breast and ovarian cancer. The company commercialized a sequence-based test to determine cancer predisposition in patients. Myriad Genetics claimed these patents provided an exclusive right to isolate the *BRCA1* and 2 gene fragments from patient samples, precluding other companies from developing breast cancer susceptibility tests. Myriad Genetics held a near monopoly on the breast cancer testing market. Women with ambiguous test results were denied options for a second opinion and the $4,000 price tag was prohibitive for many women if insurance didn't cover the cost.

The ruling states, "A naturally occurring DNA segment is a product of nature and not patent eligible merely because it has been isolated." This essentially invalidates Myriad Genetic's patent claims on extracted patient genomic DNA. Within days of the ruling, five companies announced plans to incorporate the genes in their testing pipelines at substantially lower prices than those charged by Myriad Genetics.

The ruling did uphold patent protection for engineered DNA sequences, including so-called complementary DNA - fragments that are reverse-transcribed from messenger RNA. This protection is critical to the growing biotechnology industry as it maintains intellectual property rights associated with inventions and human ingenuity.

REFERENCE: Nathwani et al., U.S. Supreme Court, (2013 September) "Association for Molecular Pathology et al. v. Myriad Genetics, Inc. et al." Retrieved from http://www.supremecourt.gov/opinions/12pdf/12-398_1b7d.pdf

U.S. guidelines on genetic testing of children

As access to genetic testing expands, questions arise about the appropriateness of genetic testing and screening for newborns and children. In response, the American College of Medical Genetics and Genomics and the American Academy of Pediatrics jointly produced a policy statement emphasizing that genetic testing and screening decisions should always be in the best interests of the child as opposed to parental or third party wishes. Predictive testing for adult-onset conditions should not be performed and carrier testing of children should only take place if there are clear health benefits during childhood. There is an emphasis on joint decision-making – including the child's input when appropriate. The guidelines also highlight the importance of genetic counseling and education for the family, but recognizes such counseling may come from health professionals who are not genetic specialists (such as pediatricians), provided they have adequate knowledge.

REFERENCE: Ross L.F., "Technical report: ethical and policy issues in genetic testing and screening of children," *Genetics in Medicine* (2013) 15:234-45 doi: 10.1038/gim.2012.176.

In brief

Genome sequencing

A roundup of recently sequenced organisms

The genomes of several organisms were sequenced during the past 12 months. A sampling, along with genome size in millions of bases and predicted number of genes, includes:

	Genome Size (Million bases)	Number of Genes
Norway spruce	19,600	28,350
Floating bladderwort	82	28,500
Bread wheat	17,100	95,000
Barley	5,100	32,000
Chickpea	738	28,250
Watermelon	425	23,450
Lotus flower	929	26,700
Pacific oyster	823	28,000
Bactrian camel	1,570	23,600
Mallard duck	13,000	15,050
Domestic pig	2,800	21,650
Goat	2,660	22,200

At HudsonAlpha, the Genome Sequencing Center worked on the genomes of several organisms, including:

	Genome Size (Million bases)	Number of Genes
Peach	265	27,850
Button mushroom	30	11,000
Freshwater leech	228	23,400
Cotton D (non-spinable)	748	37,505
Cacao	445	29,400

Livestock

A gene chip for beef quality

DNA testing may one day be used to predict beef quality and tenderness. Current methods to assess beef quality include carcass weight, the method of aging and marbling. A French research lab hopes to add a decidedly more high-tech tool to the process – genetic testing. First, the group identified 3,000 genes involved in muscle quality – regulating fat, connective tissue, protein content and tenderness. This information was used to create a DNA chip – a series of genetic tests that analyze the activity of these genes in beef samples. Comparing the genetic results to expert taste tests showed the gene data was a useful predictor of meat quality.

DNA testing of beef cattle is still in development. It historically has proven useful in identifying carriers of genetic defects and resolving paternity. The development of selection tools for traits like disease resistance, food consumption, age at first calving and meat quality are works in progress. As these traits are controlled by multiple genetic and environmental influences, developing effective assessments are challenging.

REFERENCE: Hocquette, et. al., "The GENOTEND chip: a new tool to analyse gene expression in muscles of beef cattle for beef quality prediction," *BMC Veterinary Research* (2012) 8:135-45 doi: 10.1186/1746-6148-8-135.

Growth under stressful conditions

aluminium tolerance due to copy number variants

In the tropics and subtropics, plants are often grown in acidic soils. These soil types dissolve aluminum from the clays within the ground. This is toxic to growing roots, stunting the growth and development of the plant. Aluminum toxicity is a major threat to food security in food-producing tropical regions.

A few corn varieties have been identified that tolerate aluminum-rich soils. A genomic comparison between aluminum tolerant and susceptible corn strains uncovered key differences at the *MATE1* gene. When aluminum is present, this gene is activated, expressing a protein that exports citric acid from root tip cells out into the soil. The citric acid binds to aluminum, limiting its toxicity.

Intriguingly, aluminum tolerant strains have three copies of the *MATE1* gene, compared to a single copy for the susceptible lines. The extra copies encode more protein, which exports additional citric acid into the soil, providing greater protection to the roots. This is one of the first examples of copy number variation contributing to a desired agricultural trait. A study of

Hybrid vigor

genetic explanations for heterosis

Plants commonly used for agricultural purposes are genetically exclusive strains – obtained from generations of crossbreeding for larger yields, easier cultivation and efficient harvesting. Essentially, farmers cross-pollinate two strains of a plant in hope of producing offspring with the best qualities from both parents. Many of the offspring from these genetically distinct crosses show evidence of heterosis - they are almost always more prolific and have an increased overall health than either of the parent strains. Also known as hybrid vigor, this phenomenon was first described by Charles Darwin in 1876 and has been utilized extensively by farmers and breeders. In corn, heterosis is particularly pronounced and nearly all of the corn grown in developed nations is some type of hybrid.

Varying theories attempt to explain the mechanisms leading to hybrid vigor. The dominance hypothesis argues that harmful recessive changes from one parent are suppressed in hybrids by dominant beneficial DNA variants from the second parent. The overdominance hypothesis assumes that at key genes, heterozygous alleles are superior to two copies of the same allele. Scientific evidence suggests that both mechanisms may be at work .

The availability of genomic technologies provides a clearer understanding of the molecular differences between parent strains and how those differences are transmitted to the hybrid

126 corn strains identified the triplicated gene in only three lines, all of which originated in the acidic soils of South America. This suggests the triplication is a recent event that gave these plants a growth advantage, leading to their inclusion in farmers' fields.

Similar types of comparative studies have been performed for Sorghum, a drought-tolerant grain. Scientists have sequenced 971 different sorghum strains. Of particular interest were DNA changes associated with the panicle, where the grains form and grow. Dense rows of grain maximize crop yield, but looser spacing allows the grains to dry more efficiently, reducing crop loss from moisture-associated diseases.

Recognizing a specific trait-based DNA change linked to soil tolerance or grain formation allows breeders to track trait inheritance as it is crossed into other lines. Hundreds of seedlings can be rapidly screened, offering considerable cost and time savings over traditional methods that require direct analysis of the plant throughout the growing season.

For more information, refer to *Agriculture* on page 31 and *Copy Number Variation* on page 34.

REFERENCE

Maron et al., "Aluminum tolerance in maize is associated with higher *MATE1* gene copy number," *Proceedings of the National Academy of Sciences* (2013) 110:5241-6 doi:10.1073/pnas.1220766110
Morris, et al. "Population genomic and genome-wide association studies of agroclimatic traits in sorghum." *Proceedings of the National Academy of Sciences* (2012) doi: 10.1073/pnas.1215985110

offspring. For example, genome sequencing of multiple corn varieties indicates as many as 10 percent of the genes present in the reference strain are absent in other varieties. A recent analysis of the transcriptome (that portion of the genome that is transcribed into RNA) from the primary roots of seedlings compared two inbred corn lines and their hybrid progeny. Nearly 70 percent of expressed genes were transcribed at different rates between the parents, and more than one-third were expressed at two fold or greater differences. However these expression differences were dramatically muted in the hybrids.

Epigenetic differences – small chemical structures that mark the DNA in parent-specific ways - may serve as the biological pathway behind these altered transcription levels. Scientists are exploring how hybrids could interpret these epigenetic patterns to produce healthier, higher-yielding crops.

REFERENCE

Schnable, P.S., and Springer, N.M., "Progress towards Understanding Heterosis in Crop Plants," *Annual Review of Plant Biology* (2013) 64:71-88 doi: 10.1146/annurev-arplant-042110-103827.

Protected by the ancestors

rust resistance from the genes of ancestral wheat

Over a third of the global population depends on bread wheat for survival. Bread wheat is part of a larger family of plants that traces its ancestry to an ancient strain of wild wheat. Over thousands of years, this strain was domesticated by farmers, crossed with other grain-like plants and cultivated to express specific characteristics and traits, leading to the many types of wheat today. During this cultivation process the wheat genome diverged - some strains duplicated entire sets of chromosomes while others lost large swaths of genetic information.

Genomics is making it easier to untangle this complex family tree. Working in partnership, farmers and scientists have gathered seeds from domesticated strains, heirloom strains no longer commercially grown, and wild relatives of wheat. These are stored in "seed banks" across the globe. Using these resources, the genomes of related wheat strains and the ancestral plants have recently been sequenced. The data traces the flow of genetic information throughout the process of domestication and is proving especially useful to identify genes that confer disease resistance. For example, wheat stem rust is a devastating disease that can quickly turn a healthy field of wheat into a blackened mass of twisted stems and shriveled grain. Caused by a fungus, the stem rust is carried on the wind, allowing it to spread rapidly from field to field.

Fifty years ago, the introduction of rust-resistant wheat strains helped reduce the global incidence of famine. Unfortunately, the fungal DNA frequently mutates, and some of these random changes allow it to infect what were previously resistant strains of wheat. One mutated version, known as Ug99, currently threatens the wheat-growing regions of Africa and Central Asia.

As a first step to developing new strains of bread wheat, scientists identified ancestral wheat species resistant to Ug99 and compared the information in their genomes to those from modern varieties. Although wild relatives of modern crops are usually less productive than their domesticated cousins, they have survived thousands of years of extreme environmental conditions, often due to subtle changes in genetic code. The ancestral wheat species owes its Ug99 resistance to a gene called *Sr35*. When the genomes of modern bread wheat and ancestral wheat were compared, the *Sr35* gene was found to be missing in the modern variety.

Once these genetic variants are located and characterized, the traits can be incorporated into modern crops using conventional plant breeding techniques. Initial studies suggest that transferring this gene into bread wheat will confer resistance to Ug99, protecting worldwide food supplies.

For more information, refer to *Agriculture* on page 31.

REFERENCE

Saintenac et al., "Identification of Wheat Gene *Sr35* That Confers Resistance to Ug99 Stem Rust Race Group," *Science* (2013) 341:783-6 doi: 10.1126/science.1239022

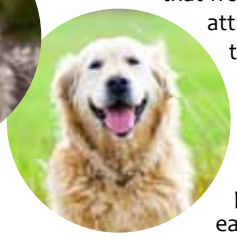
NEW FINDINGS - THE IMPACT OF LARGE-SCALE SEQUENCING

Comparative genomics

gaining insight into evolutionary mechanisms by comparing different genomes

Broadly stated, evolution is the concept that organisms experience change over time. The three main types of evolution are divergent, parallel and convergent. Recent genomic studies helped explain the mechanisms behind all three types.

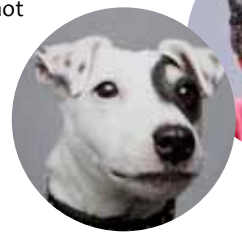
Divergent evolution in wolves and dogs



Divergent evolution is what individuals most commonly think of when they hear the word evolution - two species that share a common ancestor but become increasingly different over time. A comparison between the genomes of dogs and wolves identified regions of the genome that have undergone divergent evolution during dog domestication. Not surprisingly, half of the regions included brain genes that probably influence temperament differences between dogs and wolves.

Intriguingly, other regions contained genes involved in digestion. Dog domestication likely coincided with the start of the agricultural revolution, 10,000 years ago. Scientists have hypothesized that wolves were attracted to trash heaps containing vegetable scraps and cereal plants near early human settlements. Compared to wolves, dogs have genetic changes that increase their ability to digest starch. These genetic changes would allow the ancestors of dogs to thrive on a diet rich in starch, compared to the meat-based diet of wolves – an important step in the domestication process.

Parallel evolution between dogs and Humans



Dogs were not the only creatures that gained the ability to digest starch. A similar process occurred in humans at the *AMY1* gene, which codes for salivary amylase. Found on chromosome 1, this gene is part of a copy number variant; humans have between 2 and 15 copies. Greater copy number means increased enzyme production and additional ability to digest starch. Individuals from populations with high-starch diets tend to have more *AMY1* copies than those from low-starch diets. This suggests that as humans transitioned from hunter-gatherers to farmers, they experienced diet-driven evolutionary changes similar to those in dogs. This is an example of parallel evolution - when different species experience comparable adaptive pressures and undergo similar changes in response to that pressure.

Covergent evolution for bats and dolphins

Convergent evolution occurs when species from different ancestors independently evolve similar traits. For

Highlights from the human microbiome

a roundup of stories about the impact of our personal community of microbes

A flurry of research has been published in the past year regarding the bacteria that live in and on the bodies of humans and other organisms. A sampling includes:

"Repoopulating" the gut to treat disease: Transplanting bacteria from healthy colons into the large intestine of individuals suffering from *Clostridium difficile* has shown promising results in treating this often life-threatening bacterial infection. However, there is a strong, "Ick!" factor in transplanting bacteria obtained from the stool of another individual. Scientists identified 33 strains of bacteria commonly present in healthy colons, grew the bacteria under sterile conditions and created a stool-substitute mixture to infuse into patients' colons. Symptoms disappeared within days and follow-up found many of the infused bacterial strains still residing in the intestines, suggesting the treatment not only resolves infection, but may have long-lasting impacts on the digestive microbiome.

Tracking human diets through dental plaque: Throughout history, the human diet has undergone two major shifts: The consumption of wheat and barley during the agricultural revolution and the increased use of processed flour and sugar beginning around 1850. Dietary changes can alter the population of microbes present in the mouth. Researchers have traced the changing oral microbiome by studying the calcified remains of dental plaque on ancient skeletons. A dense biofilm of bacteria that accumulates at the base of teeth, dental plaque, mineralizes over time into a bone-like matrix called tartar. Bacteria DNA trapped in the tartar can be analyzed, shedding light on the oral microbiome's history. Tartar was extracted from European skeletons, ranging from Mesolithic hunter-gathers living 7,500 years ago to people from the late 1600s. Modern tartar samples were also analyzed.

Hunter-gatherers had a diverse oral microbiome with little evidence of bacteria that cause cavities or periodontal disease. The rise of a diet dominated by cultivated grains, which pack down around the base of teeth, selected for specific types of bacteria. Many of these types negatively impact oral health. This trend accelerated with the industrial revolution, as both processed flour and refined sugars became commonplace. Bacteria that cause tooth decay and gum disease dominated the oral microbiome. Other health issues arose with these dietary and bacterial changes to include diabetes, heart disease and stroke.

Diet, weight gain and the gut microbiome: Multiple genetic and environmental factors influence obesity. Previous studies suggest that differences in gut microbiomes may also play a role, but it was unknown whether these differences cause obesity or are a consequence of it. To address this question, the microbiomes from obese and lean humans were

example, birds, butterflies and bats all have wings, but none of these organisms inherited the genes involved in wing formation from any of the others. Each wing evolved independently. In the same way, bats and dolphins separately developed echolocation – the ability to emit and receive sound waves to locate objects and/or prey.

Building on previous findings, scientists compared the genomes of various bat species and the bottlenose dolphin. Evidence of convergence was found in nearly 200 regions of the genome – a surprisingly large number. As expected, many genes were linked to hearing, but others were involved in vision and brain development. In both creatures the genes underwent similar types of change even though the process of echolocation evolved independently.

For more details, see *Comparative Genomics* on page 33.

REFERENCES

Axelsson et al., "The genomic signature of dog domestication reveals adaptation to a starch-rich diet," *Nature* (2013) 495:360-4 doi: 10.1038/nature11837
Perry et al., "Diet and the evolution of human amylase gene copy number variation," *Nature Genetics* (2007) 39:1256-60 doi: 10.1038/ng2123.
Parker et al., "Genome-wide signatures of convergent evolution in echolocating mammals," *Nature* (2013) doi: 10.1038/nature12511.

transplanted into germ-free mice (who lack their own microbiome). After receiving the microbiome from the obese person, the mice gained weight. Mice given gut microbes from the lean person did not gain weight.

Intriguingly, if the two groups of mice were housed in the same cage, the lean microbiome could infiltrate and displace the obese one – but only if the mice were eating food high in fruits and vegetables. Because mice normally eat each other's feces, their neighbor's gut microbes are introduced into their own digestive system. Bacterioidetes – a group of bacteria that break down otherwise undigestible plant fibers – are present in the guts of lean individuals, but absent from obese microbiomes. When the mice ate a plant-rich diet, the Bacteriodetes from the lean mice were able to fill a vacant role and colonize the obese gut. If the mice ate a plant-poor, high fat diet, there wasn't an opportunity for Bacteriodetes to become established and the obese microbiome remained.

REFERENCES

Petrof et al., "Stool substitute transplant therapy for the eradication of Clostridium difficile infection: 'RePOOPulating' the gut," *Microbiome* (2013) 1:3 doi: 10.1186/2049-2618-1-3.
Adler et al., "Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions," *Nature Genetics* (2013) 45:450-455 doi: 10.1038/ng.2536.
Ridaura et al., "Gut microbiota from twins discordant for obesity modulate metabolism in mice," *Science* (2013) 341: 1241214 doi: 10.1126/science.1241214.

In brief

Infectious disease

Mycobacteria are a class of aerobic rod-shaped bacteria with waxy cell walls. This genus contains two species, *M. tuberculosis* and *M. leprae* that cause two of the world's oldest diseases – tuberculosis and leprosy. The genomes of both species were sequenced this past year, providing new insight into these harmful and difficult to treat bacteria.

Characterizing drug-resistant tubercolosis

Mycobacteria tuberculosis kills approximately two million people each year. Like most bacteria, it accumulates mutations as it spreads from person to person. Some mutations alter its virulence or ease of transmission. Other mutations provide resistance to anti-tuberculosis drugs. Drug resistance has become a major challenge to treating tuberculosis, with several strains now showing multidrug resistance. Researchers recently sequenced the genomes of hundreds of genetically diverse strains of *M. tuberculosis* in an effort to understand the specific mutations that regulate drug resistance. Many resistance markers affect the structure or permeability of the waxy cell wall, reducing the amount of drug that enters the bacterium. Others increase the overall rate of DNA mutation so the cell more quickly acquires mutation, some of which help increase resistance. In one study, nearly half of the mutations had no known function. The findings suggest drug resistance is much more complex than previously thought, emerging from an accumulation of multiple DNA changes working in concert. A catalog of resistance markers will one day help physicians track the acquisition of drug-resistant changes and identify strains on the verge of multidrug resistance.

Comparing leprosy throughout history

Leprosy is a skin, nerve and bone infection caused by the bacteria *Mycobacterium leprae*. Over 200,000 cases occur annually, primarily in developing nations. Although leprosy was widespread throughout medieval Europe, its incidence nearly disappeared during the 16th century. Medical historians have speculated the bacteria acquired mutations that dramatically reduced its virulence. The genomic sequence of the bacteria was determined using samples from the skeletons of five medieval lepers (dating from the 10th -14th centuries) as well as biopsies from 11 modern-day cases. Remarkably, the genomes of the different samples were nearly identical. Leprosy has remained almost unchanged over the past 1000 years - the decline of leprosy in Europe was not due to mutations that reduced virulence. Scientists speculate that improved resistance within the European population as well as improvements in public health likely led to the declining incidence.

REFERENCES: Young E., "Genomes reveal roots of TB drug resistance," *Nature News* (01 September 2013) doi: 10.1038/nature.2013.13645.
Schuenamann et al., "Genome-Wide Comparison of Medieval and Modern Mycobacterium leprae," *Science* 341:179-83 doi: 10.1126/science.1238286.

Genes & jeans

epigenetic exercise links

A recent study suggests one way that epigenetic alterations to our *genes* may impact the fit of our *jeans*. Environmental factors can alter gene activity through the process of epigenetic modification, especially DNA methylation - the binding of methyl groups to specific cytosine bases in the DNA sequence. If certain regions of a gene are methylated, its transcription is often repressed.

Previous research suggests DNA methylation patterns change in response to long-term exercise. As a next step, researchers examined DNA methylation in adipose (fat) tissue from 23 healthy, but sedentary men before and after a 6-month exercise program. There were significant alterations in DNA methylation patterns, which often correlated with changes in transcription levels of nearby genes. Some of these genes were in regions associated with type 2 diabetes or obesity.

Functional studies of two such genes showed that increased methylation and the resulting decrease in transcription led to greater fat storage. This work provides insight into part of the biological pathway linking exercise, epigenetic modifications and metabolism.

For additional information about epigenetics, refer to the infographic on page 20 and the article *Epigenetics* on page 37.

REFERENCE

Ronn et al., "A Six Months Exercise Intervention Influences the Genome-wide DNA Methylation Pattern in Human Adipose Tissue," *PLOS Genetics* (2013) 9:e1003572 doi: 10.1371/journal.pgen.1003572.

Age-related memory loss

keeping memories strong by regulating gene expression pathways

Clinicians believe age-related memory loss and Alzheimer's disease are two distinct disorders. Age-related memory loss is tied to neurons in the dentate gyrus of the brain, while Alzheimer's seems to begin in the nearby entorhinal cortex. To better understand the origins of age-related memory loss, gene activity was compared in human brains from different ages. The study focused on proteins that changed with age in the dentate gyrus but not the entorhinal cortex.

In particular, an age-related decline was identified for a protein known as RbAp48. When the level of this protein was experimentally reduced, young mice exhibited memory lapses like those seen naturally among more senior mice. In comparison, when RbAp48 levels were increased, older mice had improved memory performance comparable to youthful mice.

RbAp48 binds to histones – large protein complexes wrapped in DNA. Specifically, it regulates the addition of acetyl molecules on lysine amino acids present in the histone. This relaxes the interaction between the DNA and the histone, making the DNA more accessible for transcription. These findings imply that RbAp48 plays a role in keeping specific genes transcriptionally active. As levels decline with age it becomes more difficult for neurons in the dentate gyrus to respond and interact. The study also suggests drugs that increase RbAp48 activity may one day be used to reverse memory loss among older individuals.

REFERENCE

Pavlopoulos et al., "Molecular Mechanism for Age-Related Memory Loss: The Histone-Binding Protein RbAp48," *Science Translational Medicine* (2013) 5:200ra115 doi: 10.1126/scitranslmed.3006373.

How chromosomes fold up inside our cells

unraveling the science of chromatin conformation

Biology students are familiar with the images of condensed chromosomes in the nuclei of cells undergoing mitosis, but what do chromosomes look like during the rest of the cell cycle? Do the decondensed chromosomes move about the nucleus or remain in a fixed location? Do the DNA strands of different chromosomes overlap, like a big bowl of spaghetti or do they remain relatively distinct? Improved microscopic and molecular technologies are beginning to answer these questions, offering important insights into how DNA is organized within cells.

The large size of mammalian chromosomes prevents them from freely mingling. Once they begin to decondense in early interphase, their location is relatively set and within the nucleus, the DNA is organized into distinct chromosome territories (see figure at right). As a result, sequences located on the same chromosome will interact far more frequently than those from different chromosomes. Even still, live-cell imaging shows some movement within chromatin loops of a chromosome. As an analogy, think of a bed of kelp on the ocean floor – each plant is anchored in place, but the strands bend and sway with the current. These movements allow DNA sequences from neighboring chromosomes to overlap and chromatin loops to intermingle and interact.

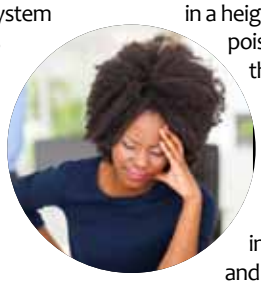
The spatial organization of the genome varies greatly from cell to cell, meaning no two nuclei will have the exact same genomic arrangement. However, chromosome placement within the nucleus is not completely random. For both mouse and human cells, short, gene-dense chromosomes tend to group near the center of the nucleus, while longer chromosomes often are found around the periphery of the nucleus.

The position of a gene within the nucleus can influence silencing or activation. Heterochromatic regions of the genome are often associated with the nuclear lamina, a filamentous network

Epigenetic links to lupus

identifying regulators of autoimmune disease

Lupus is an autoimmune disease that primarily occurs in women and affects several parts of the body, including the skin, joints, blood and kidneys. Autoimmune disorders like lupus occur when the immune system mistakenly identifies healthy cells as invaders that need to be destroyed. The disorder likely occurs when genetically predisposed individuals interact with environmental triggers. Lupus alternates between an active stage when symptoms worsen and a quiet stage where they subside.



Recent findings explored methylation levels for more than 95 percent of human genes across multiple types of immune cells. A significant lack of DNA methylation (associated with increased transcription) was identified in genes that function when exposed to interferons. Interferons comprise a group of proteins produced by cells that help trigger the immune system. The epigenetic results suggest these immune cells are in a heightened state of alert - poised for activation even if the disease is currently in a quiet phase.

For additional information about epigenetics, refer to the infographic on page 20 and the article on page 37.

The lab of HudsonAlpha researcher Dr. Devin Absher participated in this work.

REFERENCE

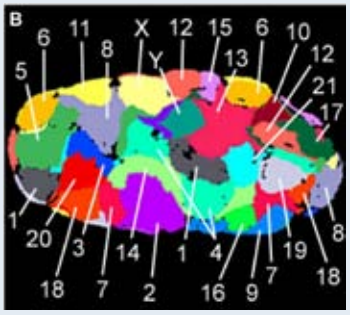
Absher et al., "Genome-Wide DNA Methylation Analysis of Systemic Lupus Erythematosus Reveals Persistent Hypomethylation of Interferon Genes and Compositional Changes to CD4+ T-cell Populations," *PLOS Genetics* 9:e1003678 doi: 10.1371/journal.pgen.1003678

of proteins lining the inner nuclear membrane. In contrast, transcriptionally active genes cluster around regions called transcription factories, which are enriched for RNA polymerase and transcription factors. These findings represent the early steps in understanding how the spatial organization and folding patterns of the genome regulate gene expression.

For additional information about this topic, see *Studying the Genome to Understand the Sequence* on page 50.

REFERENCE

Gibcus JH and Dekker J, "The Hierarchy of the 3D Genome," *Molecular Cell* (2013) 49:773-82 doi: 10.1016/j.molcel/2013.02.011.



The figure above shows the arrangment of chromosome territories in a skin cell from a human male. Different fluorescently labeled probes specifically bind to each chromosome, identifying their locations within the nucleus.

figure reproduced with permission from Bolzer et al., "Three-Dimensional Maps of Chromosomes in Human Male Fibroblast Nuclei and Prometaphase Rosettes," PLoS Biology (2005) 3:5:e157 doi:10.1371/journal.pbio.0030157.

In brief

Identifying genetic influence on disease

Gene variant + environmental trauma = PTSD

Psychiatric disorders are thought to arise from a combination of genetic and environmental risk factors, although how the inherited and experiential risks interact remains largely a mystery. Scientists previously identified a genetic polymorphism in the *FKBP5* gene that increases the risk of adult-onset post-traumatic stress disorder or major depression - but only if these individuals experienced childhood trauma. New research has implicated an epigenetic link. Apparently the trauma decreases DNA methylation on the risk allele of *FKBP5*, which increases transcription of *FKBP5* and alters how the body responds to stress.

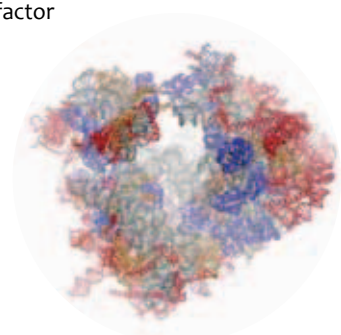
The risk of PTSD or depression does not increase when the genetic polymorphism is present but there was no childhood trauma or among patients who experienced young-life trauma but lacked the genetic change. The decresed methylation on the *FKBP5* gene only occur if both risk factors are present. In addition, the interaction seems to be age-dependent, as trauma during adulthood does not alter methylation in people carrying the high-risk genetic variant. This finding identifies a biological pathway linking childhood experiences, epigenetic modification and long-term behavioral development.

REFERENCE: Klengel et al., "Allele-specific *FKBP5* DNA demethylation mediates gene-childhood trauma interactions," *Nature Neuroscience* 16:33-41 doi: 10.1038/nn.3275.

One gene, one mRNA..... two distinct proteins?

For the first time, scientists have identified a human gene that encodes a single strand of mRNA that can be translated into two structurally distinct proteins. The *CACNA1A* gene is found on chromosome 19 and encodes the α1A protein, a channel protein that allows calcium ions to flow into cells. Mutations in this gene are associated with dominantly inherited migraine, epilepsy and ataxia.

While studying a different neurodegenerative disease also associated with *CACNA1A*, researchers discovered something totally unexpected: a completely separate protein - generated from the identical mRNA sequence of the α1A protein - was responsible for the disorder. This is due to the presence of a specific nucleotide sequence known as an internal ribosomal entry site. IRES sites provide a place for ribosomes to bind mRNA and initiate translation. Usually, IRES sites are found at the beginning of mRNA sequences, but the *CACNA1A* mRNA contained a second IRES site in the middle of the transcript. The second protein is a transcription factor that binds and activates genes involved in neuron growth. It is not known if other dual-function genes are present across the human genome.



REFERENCE: Du et al., "Second Cistron in *CACNA1A* Gene Encodes a Transcription Factor Mediating Cerebellar Development and SCA6," *Cell* (2013) 154:118-33 doi:10.1016/j.cell.2013.05.059.

In brief

Protein translation

Tracking ribosomal movement

Proteins are made when genetic instructions are copied from the DNA sequence to an intermediary molecule known as messenger RNA. Small molecular machines known as ribosomes read the RNA instructions in three letter segments called codons. The codons match specific amino acid building blocks, allowing for the assembly of amino acids into protein. This process also requires specialized molecules called transfer RNA. One end of the transfer RNA aligns with the specific messenger RNA codon, while the other end carries the corresponding amino acid into the ribosome for attachment to the growing protein chain. Although this general process of translation has been known for nearly 60 years, translocation - the stepwise movement of transfer RNA and messenger RNA within the ribosome - has remained a mystery. Thanks to X-ray crystallography, the details of translocation were recently identified. An internal rotation of the ribosome ratchets the mRNA forward and holds it in place to prevent slippage. In addition to shedding light on a previously murky process, understanding the structural specifics of translocation may speed the development of new antibiotics, which are often based on disrupting translation by bacterial ribosomes.

REFERENCE: Tourigny et al., "Elongation Factor G Bound to the Ribosome in an Intermediate State of Translocation," *Science* (2013) 340:1235490 doi:10.1126/science.1235490.
Pulk, B., and Cate, H.D., "Control of Ribosomal Subunit Rotation by Elongation Factor G," *Science* (2013) 340:1235970 doi:10.1126/science.1235970.
Zhou et al., "Crystal Structures of EF-G-Ribosome Complexes Trapped in Intermediate States of Translocation," *Science* (2013) 340:1236086 doi:10.1126/science.1236086.

Fungal genomics

White-nose syndrome in bats

White-nose syndrome is a disease that affects hibernating bats across eastern North America, including those in northeastern Alabama. Caused by a fungus, little is known about why the symptoms appear, how the disease spreads or ways to treat it. In some hibernation locations, 90 to 100 percent of the bats have died from white-nose syndrome. More than 5 million bats have died during the last seven years. Symptomatically, bats with white-nose syndrome show atypical behavior during the winter hibernation months, including daytime flight.

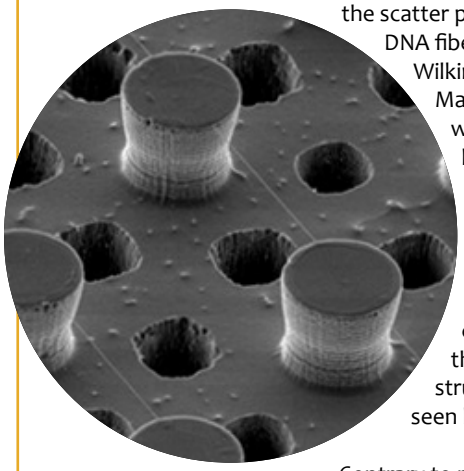
In an attempt to better understand the disease, genetic sequences of the associated fungus were compared to other regional fungi sequences. The findings show the fungus has no close relatives in North America, suggesting it is an invasive species. Early genetic studies suggested the fungus was a member of the genus Geomyces, but more detailed sequencing shows it should actually be in the genus Pseudogymnoascus. The fungi likely came from Europe. Scientists have identified European bats infected with the fungus that leads to white-nose syndrome, but the symptoms seem to be much milder. It is possible that European bats have evolved ways to resist many of the harmful effects of the fungus, potentially pointing toward treatment and prevention avenues for North American bats.

REFERENCE: Minnins, A.M. and Linder, D.L., "Phylogenetic evaluation of *Geomyces* and allies reveals no close relatives of *Pseudogymnoascuc destrucans*, comb. nov., in bat hibernacula of eastern North America," *Fungal Biology* (2013)117:638-49 doi: 10.1016/j.funbio.2013.07.001.

Seeing DNA clearly.....almost

microscopic visualization of DNA structure

In 1953 James Watson and Francis Crick determined the molecular structure of DNA. This was accomplished in part by analyzing the scatter patterns of X-rays aimed at DNA fibers obtained by Maurice Wilkins and Rosalind Franklin. Mathematical analysis of the way DNA scatters an X-ray beam showed that DNA has a double helical shape. This was not determined by looking at DNA fibers under a microscope, but inferred from indirect crystallography data. In the 1950s, the molecular structure of DNA could not be seen by microscopy.



Contrary to popular belief, that frustrating fact is still true today – science is unable to microscopically visualize single base pairs of DNA. However, researchers have recently come a little closer to that goal, producing high-contrast images of a bundle of seven DNA double helices. The secret is in the preparation of DNA for imaging. As shown in the image above, the technique essentially stretches strands of DNA between nanoscopic silicon pillars, much like stringing a clothesline across two poles. This raises the DNA strands above the surface, allowing for greater contrast during

The genome of Henrietta Lacks

balancing privacy and scientific advancement

In 1951, a 31-year-old African American woman named Henrietta Lacks was being treated for cervical cancer at Johns Hopkins Hospital in Baltimore. Without her knowledge or consent, cells taken from a biopsy of her cancer were used for research purposes. Note that at that time, permission to use samples for research was neither required nor generally sought.

Henrietta Lacks died from the aggressive cancer only a few months after initiating treatment. The cells from her biopsy, ultimately known as HeLa cells, became the first continuously cultured cancer cell line. They have been used extensively in groundbreaking biomedical research and are the most widely used human cell line in existence, mentioned in over 76,000 scientific papers. They have allowed scientists to perform experiments without using a living human and helped lead to drugs for polio, leukemia, hemophilia and Parkinson's disease.

In March of 2013, German researchers published the genomic sequence of HeLa cells. There was a strong response from scientists, patient advocates and bioethicists who objected

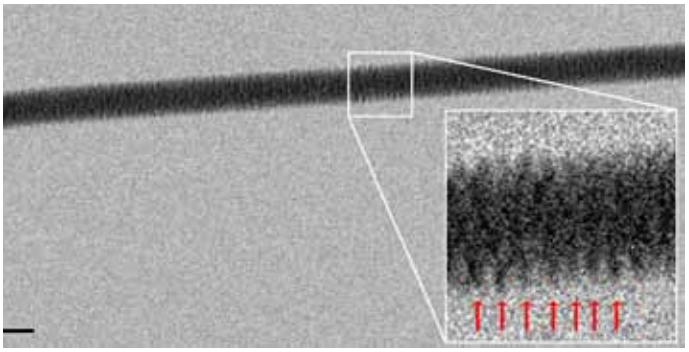
imaging. Transmission electron microscopy of the DNA bundle shows the repetitive features associated with the helical structure (figure below).

The researchers note their method of DNA preparation is strikingly similar to that used by Wilkins and Franklin, although the DNA fibers obtained for the current experiment are about a thousand times finer than those used in the 1953 experiment.

A detailed interpretation of Rosalind Franklin's famous X-ray image can be found at http://www.pbs.org/wgbh/nova/body/DNA-photograph.html.

REFERENCE

Gentile et al., "Direct Imaging of DNA Fibers: The Visage of Double Helix" *Nano Letters* (2012) 12:6453-8 doi:10.1021/nl3039162.



to the publication because it violated the privacy of the Lacks family due to the potential to identify disease risk. After the Lacks family expressed concern, the data was removed from public view.

In response, the National Institutes of Health reached an agreement with the Lacks family, giving them a role in determining what future research projects will use HeLa genomic data. In addition, any research using this data will include in their publications an acknowledgement and expression of appreciation to the Lacks family for their contributions.

A subsequent genomewide analysis determined the complex set of chromosomal rearrangements in one strain of the HeLa cells. It identified an insertion of the human papillomavirus into chromosome 8 near the MYC gene, a known oncogene. The papillomavirus is a major risk factor for cervical cancer, in part because the virus contains its own set of cancer-associated genes. This insertion is the likely reason Henrietta Lacks' cancer was so aggressive - it led to an overexpression of the MYC gene, which dramatically increases cell growth.

For more details concerning Henrietta Lacks and HeLa cells, see *The Immortal Life of Henrietta Lacks* (Crown, 2010), by Rebecca Skloot.

REFERENCE

Hudson K.L. and Collins F.S. "Biospecimen policy: Family matters," *Nature* (2013) 500:141-2 doi: 10.1038/500141a.

In brief

Identifying genetic influence on disease

The Y-chromosome inheritance pattern that wasn't

Because the human Y chromosome is inherited directly from father to son, traits arising from Y-linked genes should be identified through male-line inheritance in a pedigree. While many such traits have been proposed, including hairy ears, it was the 2004 publication of a Y-linked hearing impairment in a Chinese family that seemingly provided validation. Going back seven generations, all adult males were affected. Recently, the Y chromosome from these individuals was sequenced and found to contain a 160,000 base-pair insertion from chromosome 1. This chromosome 1 segment is associated with hearing loss, although the causative mutation is unknown. It seems likely, therefore, that the Y-linked hearing loss is due to the chromosome 1 sequence rather than a true mutation in a Y chromosome gene.

REFERENCE: Wang et al., "Genetic Basis of Y-Linked Hearing Impairment," *American Journal of Human Genetics* (2013) 92:301-6 doi:10.1016/j.ajhg.2012.12.015.

Lifestyle changes may lengthen telomeres

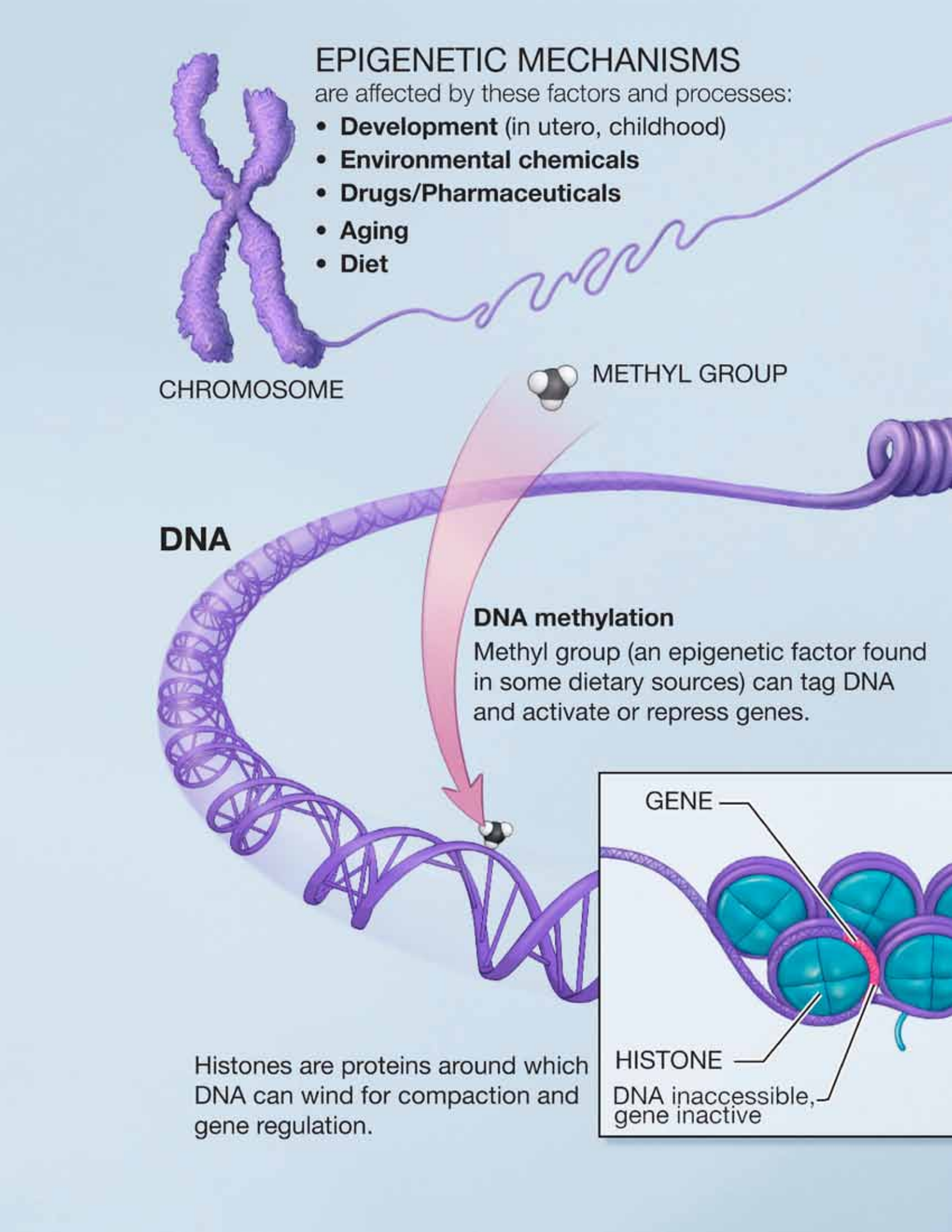
A pilot study suggests that telomere length may be altered by diet, exercise, stress management and social support networks. The study was very small and requires validation, but men who engaged in healthy lifestyles (balanced diets, moderate exercise, daily yoga, breathing or meditation and weekly social group time) experienced an average of 10 percent increases in telomere length across a five year span. In contrast, men who did not participate in these lifestyle changes saw a 3 percent average decrease in telomere length. Telomere shortening increases the risk of many chronic diseases and increasing telomere length may help prevent these conditions.

REFERENCE: Ornish et al., "Effects of comprehensive lifestyle changes on telomerase activity and telomere length in men with biopsy proven low-risk prostate cancer: 5-year follow-up of a descriptive pilot study" *The Lancet Oncology*, published online Sept 17, 2013.

Preventing the pain from a day in the sun

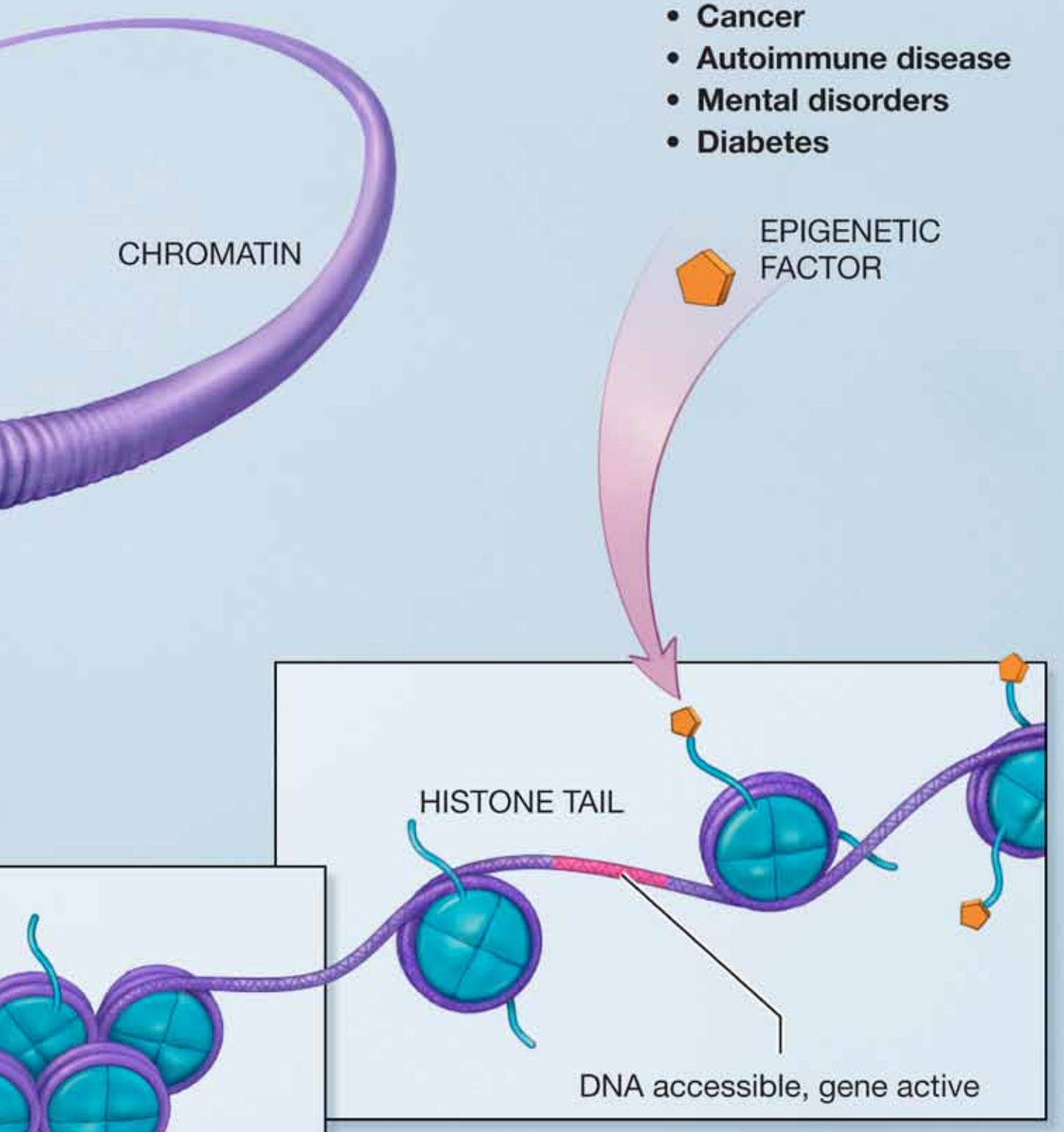
Scientists have uncovered the pathway leading to sunburn, identifying a potential approach to prevent the painful red skin associated with too much time in the sun. Research in mouse models and human skin samples suggests a molecule known as TRPV4 is to blame for the pain and tissue inflammation of sunburn. TRPV4 is an ion channel, meaning it controls the entry of positively charged ions like calcium into the cell. When skin cells absorb UVB radiation from the sun, the channel is activated and calcium floods into cells. This process also brings endothelin-1, a molecule known to cause pain, into the cell. Applying a drug that blocks TRPV4's action prevented the pain and redness associated with sunburn. It is not known if the drug protected the cell's DNA from damage associated with UVB rays. Once this is determined, TRPV4 blockers may become an important part of sunblock or skin creams.

REFERENCE: Moore et al., "UVB radiation generates sunburn pain and affects skin by activating epidermal TRPV4 ion channels and triggering endothelin-1 signaling" *Proceedings of the National Academy of Sciences* (2013) published online August 8.


EPIGENETIC MECHANISMS
are affected by these factors and processes:
• Development (in utero, childhood)
• Environmental chemicals
• Drugs/Pharmaceuticals
• Aging
• Diet
CHROMOSOME
METHYL GROUP
DNA
DNA methylation
Methyl group (an epigenetic factor found in some dietary sources) can tag DNA and activate or repress genes.
GENE
HISTONE
DNA inaccessible, gene inactive
Histones are proteins around which DNA can wind for compaction and gene regulation.

HEALTH ENDPOINTS

- **Cancer**
- **Autoimmune disease**
- **Mental disorders**
- **Diabetes**



Histone modification

The binding of epigenetic factors to histone "tails" alters the extent to which DNA is wrapped around histones and the availability of genes in the DNA to be activated.

COURSE OF STUDY CONNECTED TO GUIDEBOOK TOPICS

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

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.	
		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.	
		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

Course		Objective and Applicable Subheading	Linking Scientific Concept
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.	Agricultural Applications, Bioinformatics, Cancer, Criminal Justice and Forensics, DNA Sequencing, Personalized Medicine, Pharmacogenomics, Recombinant DNA and Genetic Engineering, RNA and Protein Analysis
		Describing techniques used with recombinant DNA	Agricultural Applications, Recombinant DNA and Genetic Engineering, RNA and Protein Analyses
	10	Explain the development and purpose of the Human Genome Project.	Bioinformatics, Criminal Justice and Forensics, DNA Sequencing, Identifying Genetic Influence on Disease, Studying the Genome to Understand the Sequence; See also Biotechnology Timeline (pg 4)
		Analyzing results of the Human Genome Project to predict ethical, social, and legal implications.	Cancer, Copy Number Variation, Criminal Justice and Forensics, Genetic Information Nondiscrimination Act, Personal Genomic Analysis, Personalized Medicine, Pharmacogenomics, Therapeutic Approaches
		Describing medical uses of gene therapy, including vaccines and tissue and antibody engineering.	DNA Sequencing, Infectious Disease, RNA and Protein Analysis, Therapeutic Approaches
AP Biology	II	Evolution	Agricultural Applications, Comparative Genomics
	IV	Continuity and Change	Agricultural Applications, Bioinformatics, Cancer, Comparative Genmoics, 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
Health	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
	6	Discuss valid and essential information for the safe use of consumer goods and health products.	Agricultural Applications, Cancer, Noninvasive Prenatal Diagnosis, Personal Genomic Analysis, Pharmacogenomics
	10	Determine the causes of disability and premature loss of life across life stages.	Cancer, Identifying Genetic Influence on Disease

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 Accountabilitiy 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 Accountabilitiy 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 4) Agricultural Applications, Bioinformatics, Criminal Justice and Forensics, Diagnosing Chromosome Disorders, DNA Sequencing, Pharmacogenomics, See also Biotechnology Timeline (pg 4)
	4	Correlate key cellular components to function.	See HudsonAlpha iCell (pg 4)
	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

Course		Objective and Applicable Subheading	Linking Scientific Concept
Introduction to Biotechnology	9	Describe methods cells use to regulate gene expression.	Cancer, Comparative Genomics, Epigenetics, RNA and Protein Analysis, Therapeutic Approaches
		Defining the role of ribonucleic acid (RNA) in protein synthesis	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.	Agricultural Applications, Bioinformatics, Criminal Justice and Forensics, DNA Sequencing, Infectious Disease
		Describing techniques used with recombinant DNA	Agricultural Applications, DNA Sequencing, Synthetic Biology
	14	Explain the development, purpose, findings, and applications of the Human Genome Project.	Comparative Genomics, Copy Number Variation, DNA Sequencing, Identifying Genetic Influence in Disease, Personalized Medicine, Pharmacogenomics, Studying the Genome to Understand the Sequence
		Analyzing results of the Human Genome project to predict ethical, social and legal implications	Criminal Justice and Forensics, Genetic Information Nondiscrimination Act, Personalized Genomic Analysis
		Describing medical uses of gene therapy, including vaccines and tissue and antibody engineering.	Cancer, DNA Sequencing, Infectious Disease, Recombinant DNA and Genetic Engineering, RNA and Protein Analysis
		Using computer bioinformatics resources to provide information regarding DNA, protein, and human genetic diseases	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
		Identify medical advancements in plant biotechnology	Agricultural Applications, Comparative Genomics
		Describing environmental advancements in plant biotechnology	Agricultural Applications; See also Biotechnology Timeline (pg 4)
	17	Describe methods of genetic engineering.	Agricultural Applications

FOUNDATIONAL CONCEPTS AND APPLICATIONS

DNA Sequencing

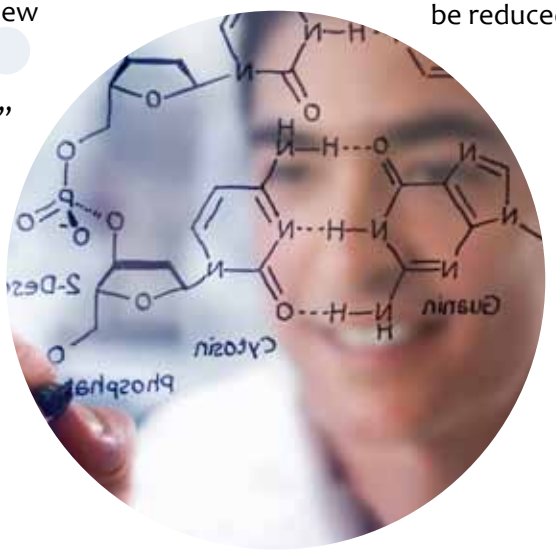
In 1977 Fred Sanger and Alan Coulson published a method to rapidly determine the specific order of the adenine, thymine, cytosine and guanine nucleotides in any DNA sequence. This technology ultimately transformed biology by providing a tool for deciphering complete genes and later entire genomes. Improvements 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 $10,000.

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 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.



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 8) 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 26 for more details). The lab has been incorporated into the AMSTI high school program (Science in Motion) across the state.

RNA and Protein Analyses

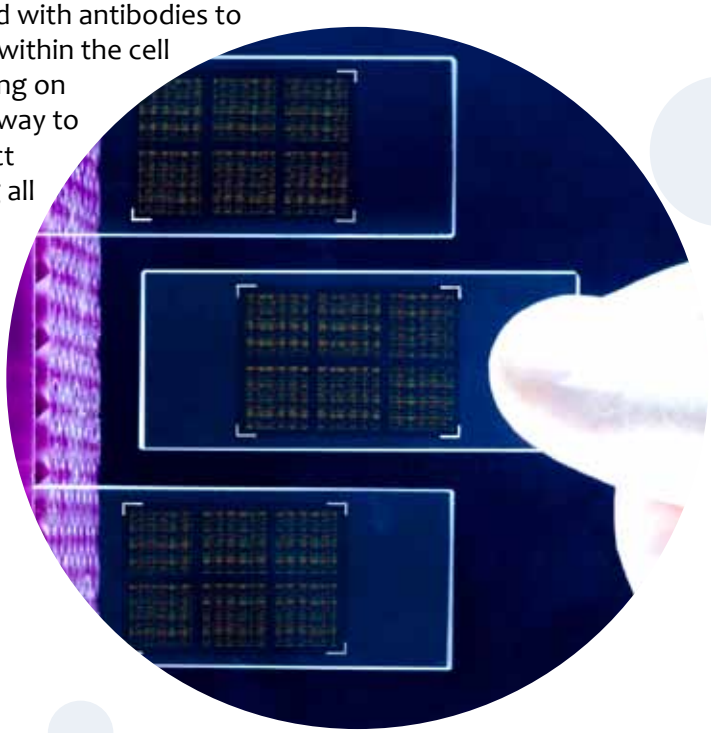
As sequencing techniques identify the genetic recipes of an organism, understanding the function of those genes becomes increasingly important. Many of the fundamental measurements used by molecular biologists seek to determine the presence, absence or relative amounts of RNA produced by a gene. Initially, 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.

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.

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.

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

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.

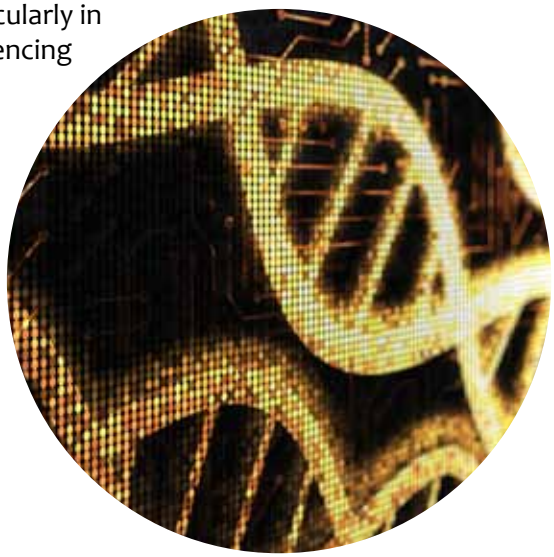


Bioinformatics

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 is incorporated into the AMSTI Science in Motion program statewide.

Agriculture

The demand for crop production is rising due to increased human population, greater worldwide meat and dairy consumption, and the expanding role of biofuels. Studies suggest that agricultural production must double between 2005 and 2050 to meet this growing need. Increasing crop yields, rather than clearing additional farmland, is believed to be the more sustainable path. However, crop yields are not increasing fast enough to keep up with projected demands. The additional challenges of drought, temperature change and poor soil quality further strain the productivity of agricultural systems.

Developing new high-yield seeds adapted for our present and future environmental conditions is a cornerstone of increased food production. This begins with the ability to locate and characterize agriculturally important versions of specific genes. These discoveries can then be shared with the farmers and commercial plant breeders who are developing new varieties of crops. Such a collaborative approach blends the emerging field of genomics with the ancient practice of agriculture, increasing yields and ensuring global food security.

Sequencing Plant Genomes for Food and Bioenergy Needs

Over the last decade, genome sequencing projects have been completed for a number of plants, including rice, corn, soybean, canola, and orange. These efforts provide a better understanding of the 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, and switchgrass. For example, soybean not only accounts for 70 percent 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, allows researchers to select for specific type high-yield strain as well as develop soybean plants that are more resistant to drought or disease.

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).

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 are in commercial production and 15 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 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.

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 percent 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.

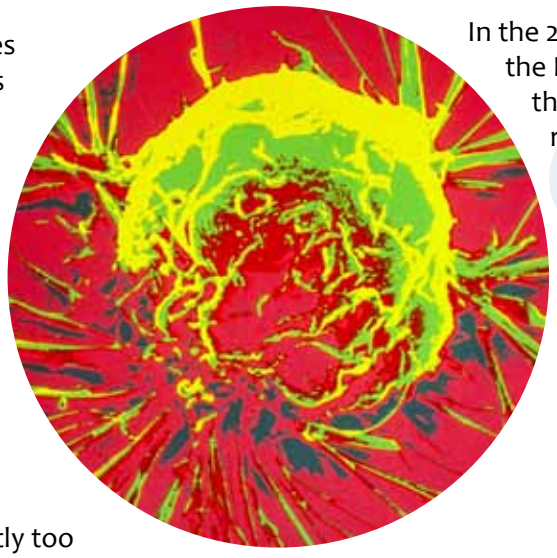
The idea that all cancers are genetic in nature and occur as a stepwise addition of mutations, many of which are initiated by environmental factors, is a useful addition to a discussion on common causes of disability and premature loss of life in a Health class (COS objective 10). These concepts should 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.

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.

However, once a small subset of the genes 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 42-43 for specific genetic tests used in this manner.

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.

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

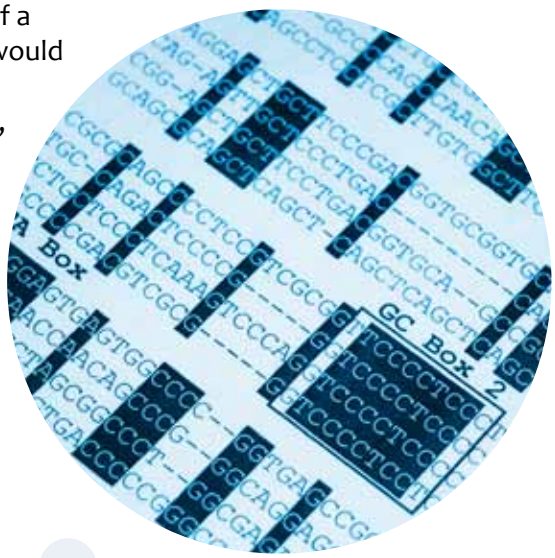
Comparative genomics provides evidence for the molecular process that underlies evolutionary theory and explains the nature and diversity of organisms, as outlined in the Biology COS objectives 5, 8 and 12 as well as in the Genetics COS objectives 2 and 7. Comparative genomics and its relationship to 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).

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 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 percent 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.



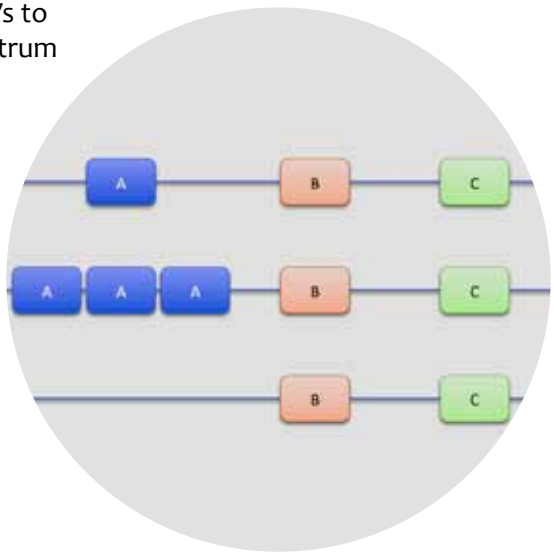
Copy Number Variation

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 percent. CNVs revise that estimate: the two genomes differ by at least 1.0 percent. 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 percent 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 49).

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.

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".

For many mammals (humans included), differences in diet and level of stress during fetal 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

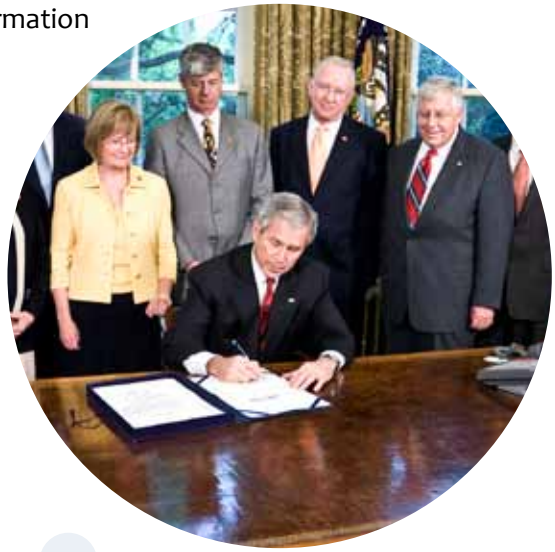
While most Americans are optimistic about the use of genetic information to improve health, many have been concerned that genetic information may be used by insurers to deny, limit or cancel health insurance and by employers to discriminate in the workplace. There has also been concern that some insurers 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 three-fourths 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 percent of individuals with two copies of the blue eyed *OCA2* allele actually had blue eyes. Blue eye color was also found among 7.5 percent of the individuals with the brown-eyed *OCA2* alleles. A number of other genes (such as *TYRP1*, *ASIP*, and *SLC45A2*) 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.

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.

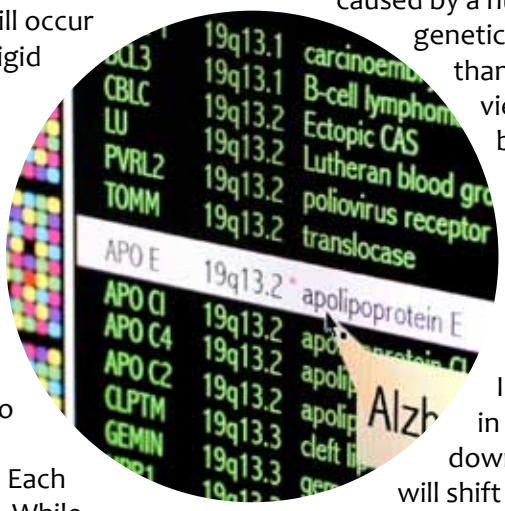
Identifying Genetic Influence on Disease

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 percent 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 percent. 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 have used 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 believe 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 percent 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?

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.

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 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-predicable 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.

Non-invasive 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 36) - 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 non-invasive 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 chose to undergo prenatal diagnosis. In 2012, three companies introduced this form of non-invasive prenatal diagnosis into the clinic. Initially only the most common trisomies 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 non-invasive 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 components of many disorders become better understood, would non-invasive diagnostic testing allow parents, with only a blood test to identify mild, adult-onset disorders, as well as nonmedical traits such as eye color?

Prenatal diagnosis is a standard part of discussions around egg and sperm formation and the abnormalities that can occur during meiosis. The advent of non-invasive 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 non-invasive 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.

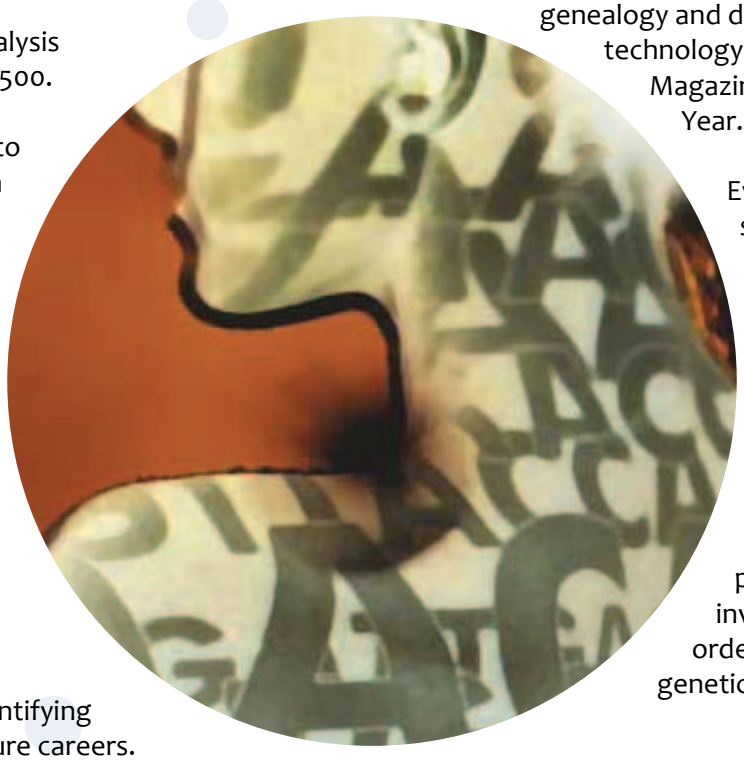
Personal Genome Analysis

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 $100 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."

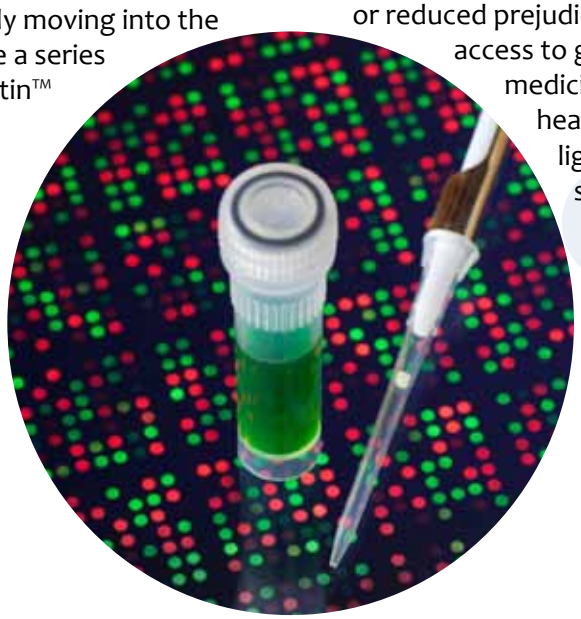
Personalized Medicine

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 43).

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 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 46-47.



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 Introducation to Pharmacy (COS objectives 9 and 11) and Intro to Biotechnology (COS objectives 11 and 14).

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 "Genome Sequencing in the Clinic" on page 10 and discussed in detail on page 28) 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 46-47. 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 percent of breast cancer patients will see any benefit from Herceptin™ -- the rest should be given another treatment. In a similar manner, Gleevac™ 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 overexpress the HER2 protein and who have received one or more chemotherapy regimens for their metastatic disease."
Pharmaceutical and surgical prevention options and surveillance	*BRCA 1,2*	**Breast cancer**: Guides surveillance and preventive treatment based on susceptibility risk for breast and ovarian cancer.
Tamoxifen	Aviara Breast Cancer Index℠ (*HOXB13, IL17BR*)	**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)	*CYP2C9*	**Cardiovascular disease**: "an increased bleeding risk for patients carrying either the *CYP2C9*2* or *CYP2C9*3* alleles."
Coumadin® (warfarin)	*VKORC1*	**Cardiovascular disease**: "Certain single nucleotide polymorphisms in the *VKORC1* gene (especially the -1639G>A allele) have been associated with lower dose requirements for warfarin."
Coumadin® (warfarin)	PGx Predict™: Warfarin	**Cardiovascular disease**: Determines *CYP2C9* and *VKORC1* 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	PhyzioType SINM	**Cardiovascular disease**: Predicts risk of statin-induced neuro-myopathy, based on a patient's combinatorial genotype for 50 genes.
Atorvastatin	*LDLR*	**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)	*UGTIA1*	**Colon cancer**: "Variations in the *UGT1A1* 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."
Erbitux® (cetuximab) Gefitinib Vectibix® (panitumab)	*EGFR* expression	**Colon cancer**: "Patients enrolled in the clinical studies were required to have…evidence of positive *EGFR* expression using the DakoCytomation *EGFR* pharmDx™ test kit." *EGFR* positive individuals are more likely to respond to the drug than those with reduced *EGFR* expression.
Erbitux® (cetuximab) Gefitinib Vectibix® (panitumab)	*KRAS*	**Colon cancer**: Certain *KRAS* mutations lead to unresponsiveness to the drug.
Erbitux® (cetuximab) and Vectibix® (panitumab) Fluorouracil Camptosar® (irinotecan)	Target GI™	**Colon cancer**: Provides information of the expression of key molecular targets—*KRAS*, *TS*, and *TOPO1*—to guide therapy.
Tagretol (carbamazepine)	*HLA-B*1502*	**Epilepsy and bipolar disorder**: Serious dermatologic reactions are associated with the *HLA-B*1502* allele in patients treated with carbamazepine. "Prior to initiating Tegretol therapy, testing for *HLA-B*1502* should be performed in patients with ancestry in populations in which *HLA-B*1502* may be present."
Immunosuppressive drugs	AlloMap® gene profile	**Heart transplantation**: Monitors patient's immune response to heart transplant to guide immunosuppressive therapy.
Ziagen® (abacavir)	*HLA-B*5701*	**HIV**: "Patients who carry the *HLA-B*5701* allele are at high risk for experiencing a hypersensitivity reaction to abacavir. Prior to initiating therapy with abacavir, screening for the *HLA-B*5701* 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)	*BCR-ABL*	**Leukemia**: "Gleevec® (imatinib mesylate) is indicated for the treatment of newly diagnosed adult and pediatric patients with Philadelphia chromosome positive [indicated by presence of *BCR-ABL*] 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)	*EGFR* 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	*MLH1, MSH2, MSH6*	**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 **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, Tolterodine, Tramadol, Venlafaxine	Amplichip® *CYP2D6/CYP2C19*	**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."
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 *FcgRIIIa*) to predict response to cancer drug rituximab.
Celebrex® (celecoxib)	*CYP2C9*	**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® (resperidone) 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)	*c-KIT*	**Stomach cancer**: "Gleevec® is also indicated for the treatment of patients with Kit (*CD117*) 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.

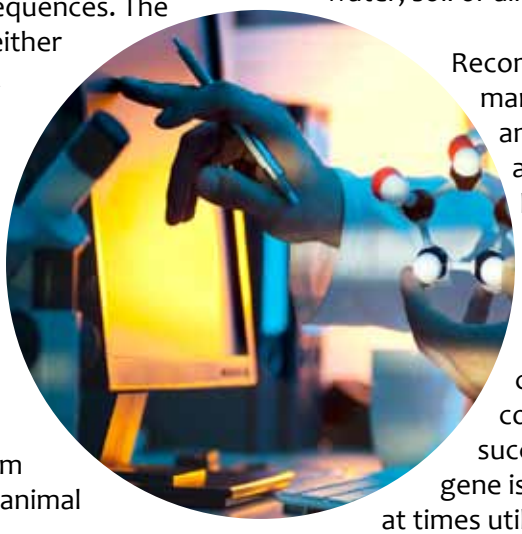
BCR-ABL = breakpoint cluster region – Abelson
BRCA 1,2 = breast cancer susceptibility gene 1 or 2
c-KIT = tyrosine kinase receptor
CYP = cytochrome P450 enzyme
DPD = dihydropyrimidine dehydrogenase
G6PD = glucose 6 phosphate dehydrogenase
HER2 = human epidermal growth factor receptor 2
NAT = N-acetyltransferase
TOPO1 = 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 100 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 31)
- 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 52.

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 pluriopotency - 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

In 2001 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 data released in 2012, the majority of DNA in the human genome appears to have 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 firmly puts to rest any 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.

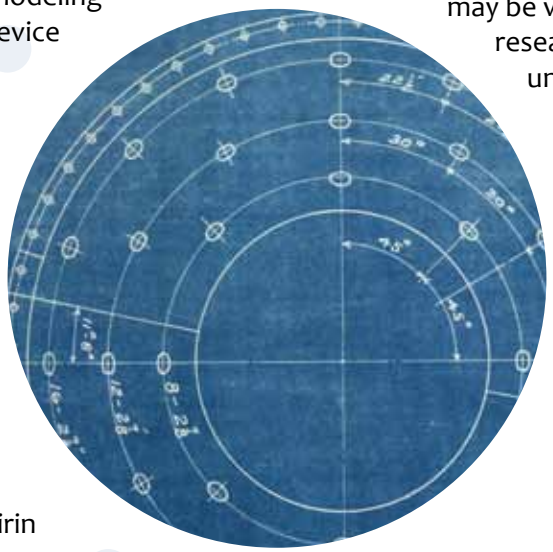
Powerful computers offer in-depth modeling and simulation to predict the behavior of the part, device or system before it is assembled. The relevant DNA instructions are then artificially synthesized and inserted into a biological cell, such as bacteria. The bacterial cell is the "chassis" or vehicle 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 link 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 unchartered scientific territory.



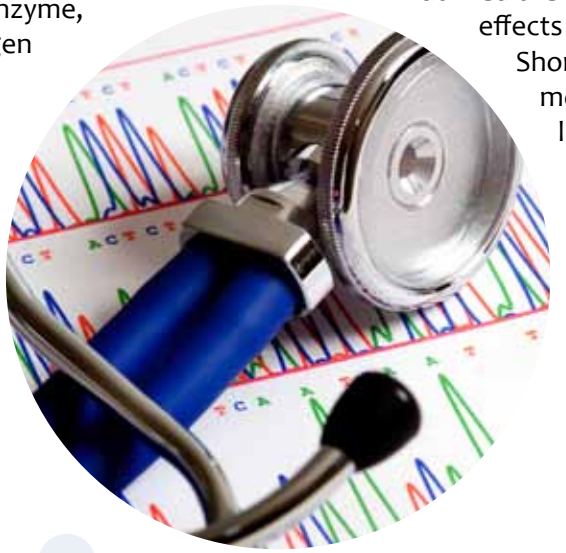
Therapeutic Approaches

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 MG. Richard H. Magen A. Klingenhoff A. Scherf M. Seifert M. Borodina T. Soldatov A. Parkhomchuk D. Schmidt D. O'Keeffe 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 Genetcs* 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 Theif: 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. Suuki 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_SWAS-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 studeies - 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 Rviews 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 U.S. 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://www.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).

Image Credits:

Page 9

International Space Station – NASA

Page 16

Typical Human Chromosomes – NHGRI Digital Media Database
http://www.genome.gov

Page 18

Henrietta Lacks – image courtesy of the Lacks family

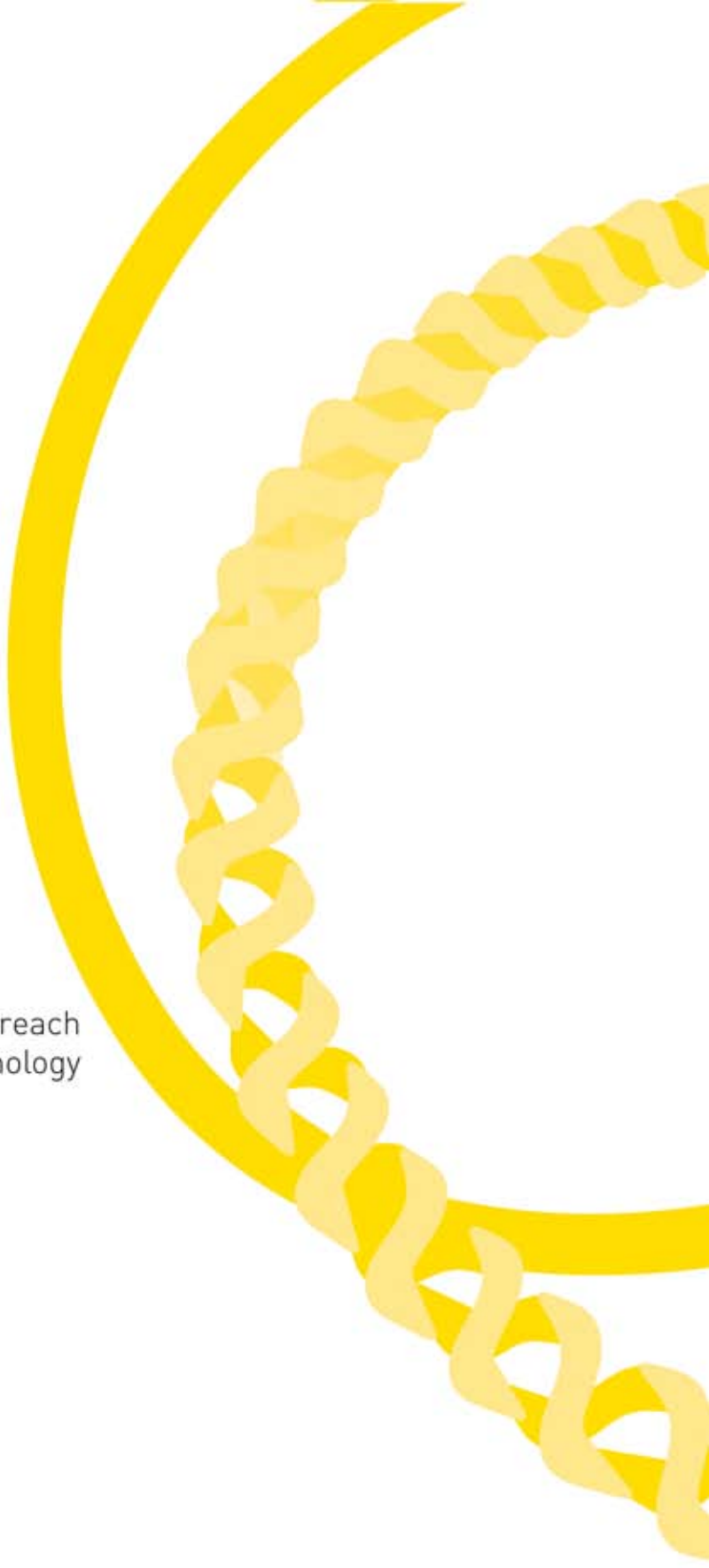
Page 18-19

Images of DNA 7-mer bundles – reprinted with permission from Gentile et.al., Nanoletters Dec 1, 2012. Copyright American Chemical Society

Pages 20-21

Epigenetics – NHGRI
http://commonfund.nih.gov/epigenomics/epigeneticmechanisms.aspx

NOTES



Contact

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