Biotechnology Discoveries and Applications

Extensions to high school science curriculum

The 2016-2017 guidebook

HudsonAlpha
INSTITUTE FOR BIOTECHNOLOGY

2016-2017 GUIDEBOOK CONTENTS

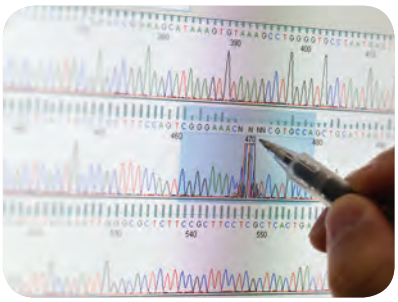
HUDSONALPHA
INSTITUTE FOR BIOTECHNOLOGY

science for life

About HudsonAlpha

HudsonAlpha is a nonprofit research institute committed to improving human health and quality of life through a unique four-fold mission of genomic research, genomic medicine, economic development and educational outreach. A collaborative environment hastens the process from discoveries made in research laboratories into the lives of individuals, whether it be through patient care or improved agriculture.

Genomic Research



HudsonAlpha scientists are adding to the world's body of knowledge about the basis of life, health, disease and bio-diversity and seeking to enable:

- Earlier and/or less invasive diagnostics
- Better, more customized treatments for disease
- Improved food and energy sources

Current research focus areas are:

- Multiple forms of cancer, including breast, ovarian, prostate, kidney and pancreatic
- Neurological and psychological disorders, including Alzheimer and Parkinson disease, ALS, bipolar disorder and autism
- Childhood genetic disorders, affecting 2 out of every 100 children born
- Immunogenomics, which is using genomics technology to understand the human immune system and related diseases such as lupus, rheumatoid arthritis, pancreatitis and psoriasis
- Agriculture and bioenergy

Biotech Enterprises

HudsonAlpha strengthens and diversifies North Alabama's economy by attracting new and growing existing life science companies to the Tennessee Valley. Through industry recruitment, retention and expansion of the Institute's associate companies or encouraging entrepreneurship, HudsonAlpha takes a leading role in building a biotech hub in Alabama. HudsonAlpha's flagship building features 270,000 square feet of laboratory, office and collaboration areas and is the cornerstone of the 152-acre HudsonAlpha Biotechnology Campus in Cummings Research Park. More than 30 associate life sciences companies do business on the HudsonAlpha campus, developing new diagnostics and therapeutics, creating health-related products and offering varied services. McMillian Park, with its signature double helix walkway, runs the length of the campus and is the backbone for future expansion.

Educational Outreach

HudsonAlpha's Educational Programs

HudsonAlpha's educational outreach team inspires the next generation of researchers, while building a more biotech-literate community. The Institute's dynamic educators are preparing future scientists through hands-on classroom modules, in-depth school and summer camp experiences, and digital learning opportunities. Additionally, the team builds awareness through community outreach classes and events. Over one million individuals were impacted through HudsonAlpha educational outreach during the 2015-16 academic year.

Teacher Professional Development

Besides providing this guidebook, HudsonAlpha has several opportunities for teacher professional development, ranging from single-day workshops to a two-week academy. These increase an educator's comfort in discussing genetic concepts and terminology and the associated ethical, social and legal issues. As part of all professional development activities, educators receive a genetics and biotechnology "toolkit" of laboratory activities, video clips, animations and online resources.

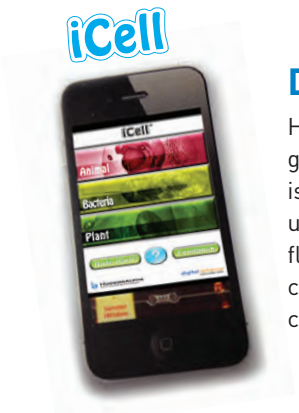
Student Experiences

Activities based on direct experience are some of the most powerful learning tools available to students. They provide a context that connects knowledge to relevancy. At HudsonAlpha, experiential learning includes field trips, classroom visits by industry leaders, summer camp sessions, in-depth internship opportunities and college-level laboratory courses. These activities engage students in biotechnology-related fields, increase exposure to career options, provide mentoring opportunities and equip students with a toolbox of content-specific skills. Communities looking to recruit science and technology occupations need to build a population of workers who can thrive in a knowledge-based economy. HudsonAlpha has crafted a pipeline of programs that blend conceptual understanding and skill acquisition to identify and engage our future workforce.

Classroom Kits and Activities

In 2007, HudsonAlpha began a partnership with the Alabama Department of Education to develop an eight-lesson module for seventh grade students matching state curriculum requirements related to DNA, how proteins are made and how genetic information is copied and segregated when cells divide. These activities have been incorporated into seventh grade classrooms across the state. Preliminary evidence on the module's impact is promising. Several teachers have shared that the percentage of their students achieving mastery on content standards addressed by the module has increased by 20 percentage points or more.

HudsonAlpha has also developed six laboratory activities for students in grades 9-12. Each activity meets state-mandated requirements for a range of courses. Activities highlight topics such as extracting DNA, exploring chromosome behavior in cells, diagnosing genetic disorders and using bioinformatics databases. Feedback has been overwhelmingly positive, with teachers expressing appreciation for the ability to expose their students to these hands-on activities.



Digital Resources

HudsonAlpha has crafted a suite of digital activities to showcase the history of genetics and biotechnology and explore the content of the human genome. iCell® is an interactive simulation that allows students and teachers to explore and understand the inner workings of a typical animal, plant or bacterial cell. Unlike flat, static images from a textbook, iCell® offers a full 3-D representation of cellular components and their dynamic interrelationships, giving students a context for learning fundamental cell structure and function.

genome cache

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®, through GooglePlay™ and at **genomecache.hudsonalpha.org**.

FREE Digital Activity

Want to enhance the way your students learn about the genetics of disease?

TOUCHING TRITON®

With this online interactive game, your students work together to ensure the health and safety of a deep space crew while learning the genomics of common disease. **Touching Triton** teaches the complexity of common disease risks from family history, environment and individual genomic profiles. Students begin to understand how genetics and lifestyle choices affect their health. Learn more at **bit.ly/touching-triton**.

Touching Triton engages students in a long-term splace flight storyline while helping them build an understanding of common complex disease risk.

Made possible by:

LOCKHEED MARTIN

SEPA SCIENCE EDUCATION PARTNERSHIP AWARD
Supported by National Institutes of Health

HUDSONALPHA INSTITUTE FOR BIOTECHNOLOGY

Grant Number 8R25 OD010981-02

triton.hudsonalpha.org

iCell®

Why use flat images from a textbook when your students can explore cell structure in 3D?

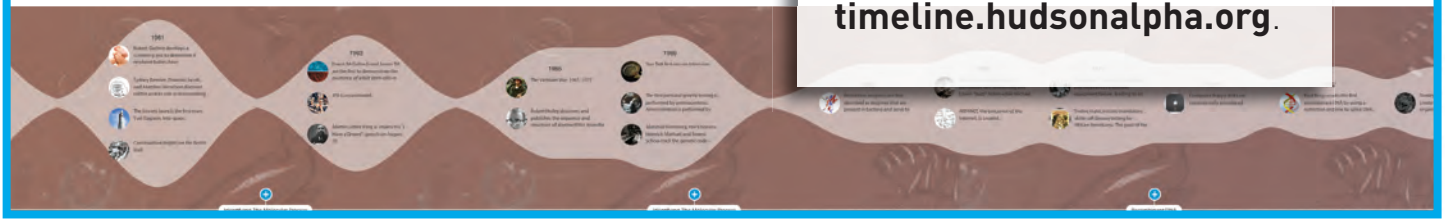
HudsonAlpha iCell®, one of Apple's featured biology apps on the iTunes® Education market, allows students to explore representative plant, animal and bacteria cells with vivid 3D models. iCell® is available on multiple platforms and has been downloaded over 1 million times by students and educators around the world.

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



Biotechnology Discoveries and Applications 2016-2017

HOW THIS GUIDE IS ARRANGED

Recent research findings are grouped on pages 10 through 23 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 teal.

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

These are identified by the  symbol in green.

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

The  symbol identifies those connections.

EXECUTIVE SUMMARY

Every year, the Biotechnology Discoveries and Applications guidebook provides educators with information about recent advances in genetics and biotechnology, allowing them to share those findings with their students. It is divided into two sections: research highlights and foundational concepts. This edition of the Guidebook contains more than 40 new discoveries published from August 2015 to July 2016.

The collection including articles describing:

- changes in the Zika viral genome that may alter its virulence and clinical impact (pg.10)
- the first eukaryote found to completely lack mitochondria (pg.11)
- a new law requiring foods containing genetically engineered ingredients to be labeled (pg.18)
- why exome and genome sequencing are moving closer to becoming standard of care tests (pg.12)
- the intriguing discovery behind why elephants almost never get cancer (pg.15)
- the impact of the CRISPR-Cas9 gene editing technique on plants, livestock and humans (pg.16 and pg.18).
- a textbook-changing realization that many lichens are composed of not two, but three different types of organisms (p.19).

The new findings are linked to one or more foundational topics, covered in detail beginning on page 30. Each topic links to course of study objectives for science, health and relevant career technical education classes in Alabama.

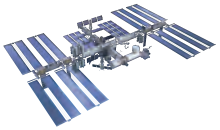
For quick reference, these linkages are also shown in table form on pgs. 24-28. Educators from states other than Alabama will find that these foundational topics align to their own state objectives fairly easily.

Neil Lamb, PhD
Vice President for Educational Outreach
HudsonAlpha Institute for Biotechnology

This past year, there were an especially large number of research publications focused on either the growing impact of human genetic testing or the emerging application of CRISPR-Cas9 gene editing. The Guidebook has given these topics additional emphasis with a set of infographics for deeper understanding (pgs. 20-23).

Lastly, I'd like to give special thanks to the talented and dedicated individuals on HudsonAlpha's Educational Outreach and Marketing and Communications teams. They have spent numerous hours helping identify, edit, design and review the content that makes this edition of the Biotechnology Guidebook an engaging and visually beautiful publication. It is a privilege to work alongside each of them on a daily basis.

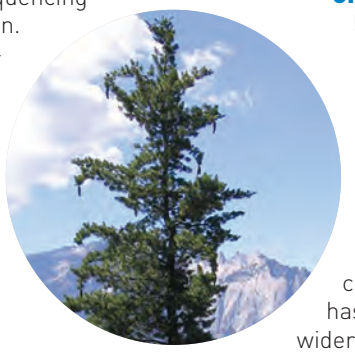
Neil E. Lamb, Ph.D.
Vice President for
Educational Outreach
HudsonAlpha Institute
for Biotechnology
nlamb@hudsonalpha.org



SCIENCE SNAPSHOTS

a quick rundown of 10 genetics and biotech stories

1. On July 20, 2016, the first-ever-in-space DNA sequencing experiment arrived at the International Space Station. The experiment is designed to determine if DNA sequencing can be performed in a non-Earth environment. A MinION® handheld sequencer will attempt to detect and sequence DNA from three previously characterized samples of viral, bacterial and mammalian sources. If the experiment is successful, it opens the door for future opportunities to study the environment aboard a spacecraft, to assess astronaut health and perhaps one day to search for life on Mars and beyond.



2. Worldwide, migraines affect one out of every seven people. In the largest study of its kind, scientists from the International Headache Genetics Consortium linked 38 regions of the human genome to migraine susceptibility. The results were obtained by pooling data from 22 previous genetic studies, collectively including almost 60,000 migraine suffers and 316,000 controls. Many of the associated regions include genes actively transcribed in brain, vascular and smooth muscle tissue, supporting the idea that migraine may be connected to abnormally functioning blood vessels in the brain.

3. The *2010 Guidebook* highlighted the first "man-made" genome, a 1.08 million basepair genome of the bacteria Mycoplasma mycoides that was synthesized entirely in the lab. The scientific team responsible for this work has spent the last five years methodically paring down the bacterium's 901 genes in search of the smallest number of genes needed to support life. They recently reported a bare-bones minimal genome of 531,000 basepairs with 438 protein-coding genes and 35 RNA genes. About a third of these genes have unknown functions but are required to produce viable cells, suggesting the existence of yet-unidentified functions essential for life.

4. Studies of stickleback fish identified a region of the genome that impacts bone size and shape. The regulatory region controls transcription of the GDF6 gene, which is important for bone development. Although a similar GDF6 regulatory region is present across nearly all mammals, it is absent in humans. Mouse models lacking this regulatory region have significantly shorter toes on their hind feet. The loss of this enhancer region may explain some of the differences in foot shape and size that occurred as early humans transitioned to walking upright on two feet.

HudsonAlpha researchers Richard Myers, PhD, Jane Grimwood, PhD, and Jeremy Schmutz contributed to this research.

5. A team at Children's Mercy Hospital in Kansas City has been noted by the *Guinness World Records*® as having the fastest time to reach a patient diagnosis through whole genome sequencing. The team went from sample acquisition to genomic answer in 26 hours, nearly half the time of their previous record, set four years ago.

6. The *2011* and *2015 Guidebooks* profiled Oxitec, a British company that developed a genetically modified male mosquito that when mated to wild mosquitoes, prevents fertilized eggs from developing into larvae. In April 2015, as part of an initiative to reduce the incidence of mosquito-borne diseases such as dengue fever, Oxitec's GM mosquitoes were released in CECAP/Eldorado, a neighborhood of 5,000 residents in the city of Piracicaba, Brazil. Over the next twelve months, 12 cases of dengue fever were reported, compared to 133 cases the previous year. A contract has been signed to release GM mosquitoes across a wider region of the city during the coming year.

7. The genome of the endangered California sugar pine (*Pinus lambertiana*) was recently sequenced. The genome of this cone-bearing tree is the largest ever sequenced for any organism – ten times the size of a human's. One of the tallest tree species in the world, the sugar pine's survival is threatened by a fungal pathogen, the ongoing California drought and damage from bark beetles. It is hoped that sequence information may speed the development of hardier, disease-resistant varieties.

8. BabySeq is a best-practice pilot study to explore the process and impact of newborn genomic sequencing. A collaboration between Brigham and Women's Hospital and Boston Children's Hospital, this randomized clinical trial is enrolling two groups of infants and their parents: 240 healthy newborns and 240 from the neonatal intensive care unit. All babies will undergo standard newborn screening, but half will additionally receive exome sequencing. The project seeks to determine whether newborn sequencing benefits children's health without negatively affecting their parents' attitudes or parent-child relationships.

9. Domestic horses have undergone thousands of years of selection for traits such as speed, strength and appearance. Scientists recently identified the genetic player behind the ancestral color pattern known as Dun: pale hair with a thin dark stripe along the back. While the hairs from the darker stripe are uniformly pigmented, the pale coloration results from an uneven distribution of a pigment-controlling transcription factor active within the nucleus of hair follicles. The inner half of the hair lacks pigment-producing melanocytes - a finding very different from that observed in other animals.

HudsonAlpha researcher, Greg Barsh, MD, PhD contributed to this research.

10. Compared to other primates, humans have lost most of their body hair. They have, however, retained a significant amount of hair on their face and head, although there is considerable variation in the texture, color and placement of this hair. A recent genome-wide association study of over 6,000 Latin Americans identified 18 regions of the genome linked to features of scalp and facial hair. New regions were identified for hair shape and balding, as well as the first reported variants associated with greying, eyebrow and beard thickness, and unibrow.

See image references on p.54

NEW FINDINGS

Genetics and the peppered moth

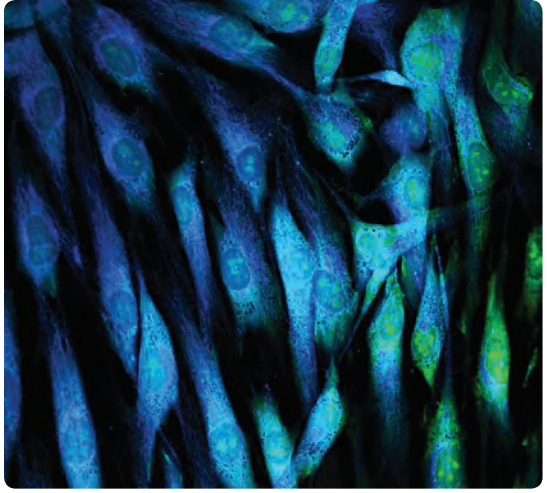
The peppered moth *(Biston betularia)* is a textbook example of adaptation to changing conditions. During the Industrial Revolution, as soot from English factories blackened tree trunks, naturalists noticed black-winged varieties – a change from the normal lighter ones. Within a few years, black moths predominated in urban areas, presumably because they could blend into darkened tree trunks and escape from predators.

The *2011 Guidebook* noted the genetic mutation had been traced to moth chromosome 17. This past year, researchers announced the first intron of the black moth *cortex* gene contained several copies of a transposable element - a "jumping gene" composed of virus-like DNA pieces that copy and insert themselves throughout the genome. The insertion increased *cortex* transcription, resulting in an overabundance of protein. The protein is involved in wing scale growth, although researchers are uncertain of the link between its overabundance and the pigment change.

The mutation is estimated to have occurred around 1819, allowing 20-30 moth generations for the frequency of black moths to increase throughout the population before their first recorded sighting in 1848.

REFERENCE: Hof, A. E. V. T., et al. The industrial melanism mutation in British peppered moths is a transposable element. Nature (2016) 534: 102-105. doi:10.1038/nature17951.

Turning skin into muscles and nerves



In 2006, adult human skin cells were "reprogrammed" into pluripotent stem (iPS) cells by initiating transcription of genes typically active only during early embryonic development. Theoretically, these iPS cells can be coaxed to form other cell types such as neural or muscle tissue. The process is laborious, time consuming and may lead to increased risk of tumor formation. Direct reprogramming that enables one cell type (such as skin) to directly express the genes active in the target cell type (such as nerve or cardiac tissue) would skip the pluripotent step and its drawbacks.

In 2016, researchers did just that – chemically reprogramming human skin cells into functional cardiac cells using a cocktail of nine small molecules. These molecules mimic transcription factors that suppress the expression of skin genes or enhance the expression of cardiac genes. Treated cells show expression patterns similar to cardiac muscle cells and synchronized "beating" in culture. When transplanted into damaged heart tissue in mice, the reprogrammed human cells integrated and began to form new cardiac muscle in the damaged area.

Separately, the same researchers reprogrammed mouse skin cells into neural stem cell-like cells, using a different nine-molecule combination. The resulting cells functioned like nerve cells and showed the chemical signaling of neural tissue. Still in the early stages, cellular reprogramming may offer the potential to regenerate damaged heart muscle or repair nerve tissue damage by treating a patients' own cells with a drug cocktail targeted to produce the cell of choice.

REFERENCE: Cao, N. et al. Conversion of human fibroblasts into functional cardiomyocytes by small molecules. Science (2016) 352:1216-1220 doi:10.11126/science.aaf1502.

Zhang, M. et al. Pharmacological Reprogramming of Fibroblasts into Neural Stem Cells by Signaling-Directed Transcriptional Activation. Cell Stem Cell (2016) 18:653-667 doi:10.1016/j.stem.2016.03.020.

Analyzing the genome of Zika

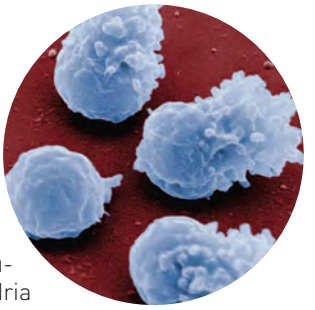
On February 1, 2016, Zika virus was declared a global public health emergency by the World Health Organization. First isolated in 1947 from a monkey living in Uganda's Zika forest, less than 20 cases of human infection were reported across African and Southeast Asian until outbreaks in the Pacific Islands in 2007 and 2013. From there, the virus traveled to Brazil (with an estimated 500,000-1,500,000 cases in 2015) before explosively expanding across Latin America and the Caribbean.

Zika is spread to people primarily through the bite of an infected mosquito, although person-to-person transmission has recently been documented. Previously thought to result in mild and short-lasting symptoms, recent evidence suggests Zika may also cause significant diseases of the adult nervous system and fetal brain defects such as microcephaly (a condition in which infants have abnormally small heads).

Zika is a flavivirus with a single-stranded RNA genome inside a lipid membrane that is wrapped in a protein shell. Its 10,800-base genome codes for a single protein precursor that is subsequently cleaved into 10 proteins. Researchers have analyzed the sequences obtained from humans, animals and mosquitoes between 1947 and 2015 as well as Latin American samples from the current epidemic. Based on the results, the viral sequences cluster into two lineages, African and Asian. The strains obtained from the current Latin American outbreak show greater similarity to the Asian lineage, suggesting the current Zika cases have evolved from that common ancestor.

Scientists compared the sequences between current and historical outbreaks in hopes of identifying genetic changes that could account for Zika's increased viru-

Degrading dad's mitochondria

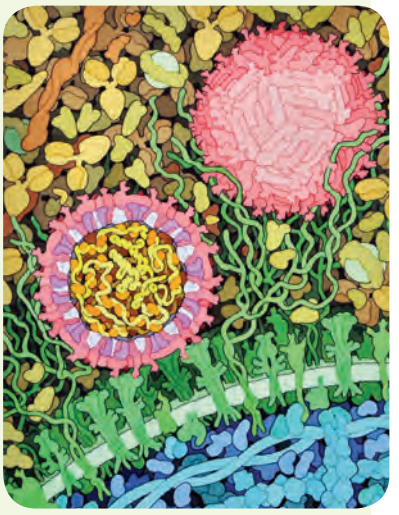


In most sexually reproducing organisms, mitochondria are inherited exclusively from the mother. The mitochondria present in sperm are selectively eliminated shortly after fertilization, although the mechanisms behind mitochondrial removal are poorly understood. Mitochondria in sperm must produce enormous amounts of energy during the journey to fertilization. It is thought this may cause significant damage to the mitochondrial DNA (mtDNA), marking it for destruction.

Working in *C. elegans*, researchers sequentially silenced genes for mitochondrial proteins and tracked the fate of paternal mtDNA after fertilization. Paternal mtDNA persistence was linked to loss of *CPS-6*, which encodes an endonuclease involved in cell death. Typical worms eliminated paternal mitochondria by the four-celled embryo, but paternal mitochondria remained until the 64-cell stage when *CPS-6* was silenced. These embryos also had slower rates of cell division and increased spontaneous cell death. Further work showed that after fertilization, CPS-6 protein travels from the inner membrane to the matrix of paternal mitochondria. Once there, it degrades mtDNA and destabilizes the inner mitochondrial membrane, which likely triggers the maternal cell to destroy the paternal mitochondria

REFERENCE: Zhou, Q. et al. Mitochondrial endonuclease G mediates breakdown of paternal mitochondria upon fertilization. Science (2015) 348:873 doi:10.1126/science.aac5605

lence and pathogenicity. They detected more than three dozen genetic changes that led to amino acid differences in the recent viral strains. Several of the substitutions mapped to the prM protein, believed to be important in assembly and maturation of the virus. Computational modeling suggests these amino acid differences would lead to dramatic structural changes, although the consequences of these alterations were still unknown. Further studies are required to link the biological significance of the genomic variation from pre-epidemic to epidemic Zika.



REFERENCES: Zhu Z. et al. Comparative genomic analysis of pre-epidemic and epidemic Zika virus strains for virological factors potentially associated with the rapidly expanding epidemic. Emerging Microbes and Infection (2016) 5, e22 doi:10.1038/emi.2016.48

Wang L. et al. From Mosquitos to Humans: Genetic Evolution of Zika Virus. Cell Host & Microbe (2016) 19:561-5 doi:10.1038/emi.2016.48

In brief

A eukaryote that lacks mitochondria

Researchers have identified an organism that defies textbook definitions: a eukaryote that lacks mitochondria! Typically, eukaryotes contain a nucleus and other membrane-bound organelles, the most prominent being the energy-generating mitochondria. However, a recent analysis of the eukaryote *Monocercomonoides* finds it has no mitochondrial structures, and lacks the genes and proteins typically associated with mitochondria.

Monocercomonoides are members of the phylum Metamonada, a diverse group of flagella-possessing, single celled protozoa that live in very low oxygen conditions as parasites or symbiotes. Without abundant oxygen, their mitochondria cannot use oxidative phosphorylation to generate ATP, the molecular currency of energy. Instead, most members of this phylum produce ATP through glycolysis and anaerobic fermentation. As there is no selective pressure to keep the mitochondrial oxidative phosphorylation genes, they have been lost from the Metamonoda genomes. In many cases, a vestige of the mitochondrion remains, known as a mitochondrion-related organelle (MRO). Although this remnant lacks the folded cristae structures and circular mitochondrial genome, it still performs some important cellular functions such as assembling the iron-sulfur clusters that are critical components of certain eukaryotic proteins.

Monocercomonoides, however, doesn't even contain an MRO. A search of the 16,000 genes in its 70 million base pair genome found no similarity to known mitochondrial genes, including the iron-sulfur assembly genes whose proteins function in the MRO. Surprisingly, *Monocercomonoides* uses a completely different iron-sulfur assembly process, encoded by genes that were likely acquired from bacteria.

The data suggest that ancestors of *Monocercomonoides* once had mitochondria but shed the components in stages. As the organism evolved to thrive in an area lacking oxygen, the energy-producing mitochondrial genes were no longer needed. The mitochondrion was streamlined to an MRO, containing the iron-sulfur assembly system. At some point, the organism acquired a second iron-sulfur assembly mechanism, likely by lateral gene transfer from bacteria. Ultimately, this secondary system became primary and the mitochondrial components became dispensable, leading to the complete loss of all mitochondrial-associated structures from this eukaryote.

REFERENCE: Karnkowska A. et al. A Eukaryote without a Mitochondrial Organelle. Current Biology (2016) 26:1274-84 doi:10.1016/j.cub.2016.03.053

In brief

DNA variation - seeking consensus

Determining whether a DNA change has clinical implications is a challenging task, yet the benefit of genetic testing relies on the ability to correctly interpret sequence findings. In 2015 the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) developed recommendations for classifying DNA changes into five categories (pathogenic, likely pathogenic, uncertain significance, likely benign or benign). The recommendations use multiple types of evidence (variant frequency, computational predictions, functional studies, etc.) for prioritizing and categorizing the variants. To test the recommendations, nine sequencing laboratories independently classified a common set of variants and compared their results. Initially, the labs agreed on the classification only 34% of the time, which was no different than the level of agreement when each lab used its own internal classification guidelines. However, when the laboratories discussed their decision-making process and clarified their understanding of the recommendations, agreement increased to 71%. Among the remaining differences, only 5% varied in a way that could impact clinical management. The disagreements were often due to individual interpretations of the recommendations, which is not surprising given the complexity of the data and the subjective nature of deciding if specific criteria were met. The findings underscore the importance of a standardized approach for variant assessment and highlight the need for detailed guidelines and significant training to ensure consistency in classification.

REFERENCES: Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine (2015) 17:405-423 doi:10.1038/gim.2015.30.

Amendola, L.M. et al. Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. The American Journal of Human Genetics (2016) 98:1067-76 doi:10.1016/j.ajhg.2016.03.024.

HudsonAlpha researcher Greg Cooper, PhD, and members of his lab contributed to this work.

HudsonAlpha researcher David Bick, MD, is a clinical geneticist at the Smith Family Clinic for Genomic Medicine. The new clinic (powered by HudsonAlpha, Children's of Alabama and UAB Medicine), located on the campus of HudsonAlpha, is the first clinic devoted to genomic medicine. Whole genome sequencing provides patients and their families better treatment and monitoring options.

Exomes and genomes galore

Determining the population frequency of a genetic variant is one way to put its biological significance into context – DNA changes commonly found in seemingly healthy populations are generally not harmful. This type of analysis depends on access to large sets of comparison genomes. Fortunately, two such datasets have recently been described in scientific publications.

1000 Genomes – www.1000genomes.org: Initiated in 2008, and profiled in the *2011 Guidebook*, this project analyzed the genomes of 2,504 individuals pulled from 26 global populations. It provides the most comprehensive view of the entire genome (coding and noncoding sequences) to date.

Exome Aggregation Consortium (ExAC) – exac.broadinstitute.org: An international coalition of researchers gathered exome sequence data from 60,706 individuals who were sequenced as part of various genetic studies (including the exome portion of the 1000 Genomes Project). This is almost ten times larger than previous exome databases and includes around ten million rare variants, most of which appear in only a single individual.

REFERENCE: The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature (2015) 526:68-74 doi:10.1038/nature15393.

Exome Aggregation Consortium. (2016) Analysis of protein-coding genetic variation in 60,706 humans. bioRxiv doi:10.1101/030338 (preprint)

Using the best test first

Moving genomic sequencing into the clinic

Individuals with a potential genetic condition who follow the current standard of care in clinical practice often undergo a series of diagnostic tests that stretches over months or even years. The process is costly and doesn't always provide an answer to patients and their families. New research, however, is shining light on the potential for large-scale genomic sequencing in the clinic as a first-line tool instead of a technique of last resort. These studies propose that exome sequencing or perhaps even better, whole genome sequencing, could provide individuals with an answer more quickly, more reliably, and at a lower cost than conventional diagnostic tests.

A group of Australian scientists conducted exome sequencing – looking at all of the expressed genes in a genome – on 80 infants from newborn up to age two who all likely had a genetic disorder. They were able to diagnose 46 of the children compared to 11 children who had been diagnosed via the standard testing methods. The authors of the paper that resulted from this research say the study is evidence that when exome sequencing is the first test administered for patients with a likely genetic disorder, the diagnostic process is considerably shorter and more successful. With a shorter diagnostic journey, patients and their families can benefit from better treatment and monitoring options.

Identifying genes that protect us

Mendelian diseases are caused by a change in a single gene. Many genetic studies of Mendelian diseases highlight new disease-causing mutations identified in patients. In contrast, the Resilience Project identifies healthy individuals who carry the mutations for Mendelian conditions yet have no symptoms of the disease. Researchers analyzed 874 genes in 589,306 genomes looking for genetic disorders that are known to appear in childhood. They identified 13 adults who harbor mutations for eight severe childhood conditions but display none of the symptoms for those diseases. Previous studies noted these Mendelian disorders are completely penetrant – an individual with a mutation will develop the disease. The Resilience Project results suggest these diseases may instead show incomplete penetrance, possibly through a buffering effect from DNA variants elsewhere in the genome.

Resilience Project researchers suggest that understanding the protective mechanism behind this so-called superhero DNA might lead to better ways to treat the diseases. In other words, instead of looking for disease-causing genetic changes, they are looking for the genes that may protect individuals from inherited conditions.

The researchers who led this project note that their retrospective study did not allow them to contact any of the individuals whose genomes were studied in order to gather additional information to validate the resilience. The group plans a prospective study to expand the search for resilient individuals. They hope that by analyzing massive numbers of genomes for these rare superhero DNA traits, they can gain new insights into Mendelian diseases.

REFERENCE: Chen R. et al. Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases. Nature Biotechnology (2016) doi:10.1038/nbt.3514.

A Dutch research project concluded that whole-exome sequencing, when used as the first test in difficult-to-diagnose patients with intellectual disability, could save the health care system substantial costs by preventing other procedures and tests. Their analysis suggested that exome sequencing as a first option to diagnose intellectual disability would save an average of $3,500 for patients who receive a diagnosis and up to $1,700 for patients for whom exome sequencing does not provide a diagnosis.

One step beyond exome sequencing is whole genome sequencing, which looks at an individual's entire genome rather than just the portion that is expressed. Research by a Canadian team indicates that whole genome sequencing for individuals with developmental delay or congenital abnormalities may be the best diagnostic tool available. They conclude that implementing clinical whole genome sequencing as the first test, rather than a last resort, could provide patients with a diagnosis more often and more quickly than any other available diagnostic test.

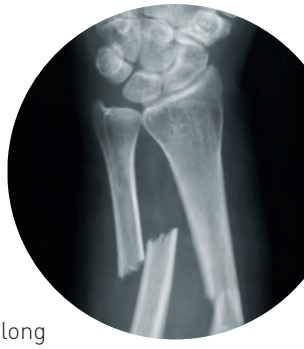
REFERENCES: Monroe, G.R. et al. Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. Genetics in Medicine (2016) doi:10.1038/gim.2015.200.

Stark, Z. et al. A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. Genetics in Medicine (2016) doi:10.1038/gim.2016.1.

Stavropoulos, D. et al. Whole-genome sequencing expands diagnostic utility and improves clinical management in paediatric medicine. Genomic Medicine (2016) doi:10.1038/npjgenmed.2015.12.

In brief

Building better bones before birth

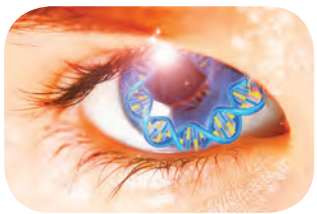


Osteogenesis imperfecta (OI) is a genetic disorder that causes unusually brittle bones. Children with OI suffer from frequent and severe factures beginning before birth. Most OI patients have mutations in a gene coding for type I collagen, an important part of bone framework. When this framework is disturbed, bones break easily, leading to life-long physical disabilities and stunted growth.

A new clinical trial aims to treat OI children across Europe with collagen-producing donor stem cells before they are even born. As part of the *Boost Brittle Bones Before Birth* (BOOSTB4) project, scientists will use a specific type of stem cell that targets and strengthens the bone by producing normal type I collagen. The researchers believe administering the cells in utero increases the likelihood of developing strong bones for life. During the first two years, additional stem cells would be provided every six months. Although this procedure has been performed a small number of times in the past, no well-designed clinical trial has confirmed the treatment truly makes a difference. Scientists plan on enrolling the first set of families from across Europe during 2016.

REFERENCE: University of Leicester (2015) Stem cell study to help fight brittle-bone disease [Press release]. Retrieved from http://www2.le.ac.uk/offices/press/press-releases/2015/october/stem-cell-study-to-help-fight-brittle-bone-disease.

RPE65 phase 3 clinical trial gene therapy success



Gene therapy is a type of treatment that involves getting healthy genes into the body to replace an individual's mutated genes that cause disease. Despite hundreds of trials of gene therapies in humans, none are approved for use in the United States. However, a phase III clinical trial has successfully used the technique to correct a rare, inherited form of blindness in 21 patients. The patients all have a type of Leber's congenital amaurosis, which causes night blindness and loss of peripheral vision, and can eventually lead to total blindness. In the trial, patients who had received the treatment were able to more easily maneuver in dimmer light than they could before the therapy. In early tests of the treatment, a boy who had relied on canes and a classroom aide had enough vision restored that he was able to start playing baseball and read the blackboard. The therapy developer, Spark Therapeutics, plans to apply to the FDA for approval to sell the product.

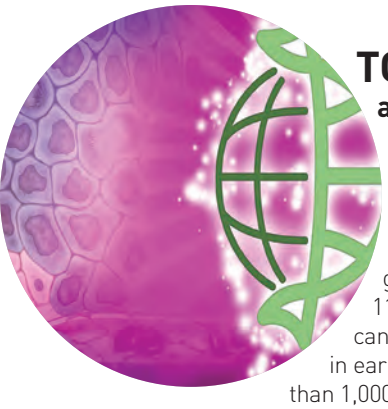
REFERENCE: Spark Therapeutics Announces Positive Top-line Results From Pivotal Phase 3 Trial of SPK-RPE65 for Genetic Blinding Conditions (October 5, 2015). Retrieved July 28, 2016 from http://ir.sparktx.com/phoenix.zhtml?c=253900&p=irol-newsArticle&ID=2093863

In brief

Cancer Moonshot

During the 2016 State of the Union address, the President announced the formation of a broad initiative to accelerate cancer research. Known as the "Cancer Moonshot," the project seeks to double the pace of advances to improve our understanding of cancer, develop more effective techniques for cancer prevention and early detection, and increase patient access to cancer therapies. The initiative is chaired by Vice President Joe Biden and involves increasing collaboration across multiple federal agencies as well as non-profit advocacy groups, research universities and institutions, and for-profit companies. As part of the Cancer Moonshot, the National Cancer Institute has announced several new activities and programs, including a public-private partnership with over two dozen pharmaceutical and biotechnology companies to provide a "one-stop shop" for obtaining approved and investigational cancer drugs. This eliminates the need for scientists and clinicians to negotiate individually with each company for research projects and clinical trials – a process the NCI notes can take up to 18 months.

REFERENCES: Cancer Moonshot. (n.d.). Retrieved July 27, 2016, from http://www.cancer.gov/research/key-initiatives/moonshot-cancer-initiative



TCGA
and the Genomic Data Commons

During the last decade, The Cancer Genome Atlas (TCGA) has comprehensively analyzed the genomic changes of more than 11,000 patients across 33 types of cancer. The project will come to a close in early 2017, having contributed to more than 1,000 cancer studies. It has improved understanding of how cancer occurs, revolutionized the way cancer is classified, and identified genomic alterations in tumors that can be targeted with existing therapies and clinical trials. TCGA has generated 2.5 petabytes of genomic data. This is the equivalent of nearly 530,000 DVDs of information. Fortunately, this data has now been centralized and reorganized in the Genomic Data Commons (GDC), part of the National Cancer Moonshot and the President's Precision Medicine Initiative. Cancer researchers around the world will be able to add their own findings to the database and the GDC will standardize the results so they can be analyzed online and in a consistent way. The GDC will store information from the DNA of patients participating in clinical trials for cancer therapies, becoming a data-sharing platform doctors can use to better understand their patient's cancer and identify appropriate treatments to give.

REFERENCES: The Cancer Genome Atlas – http://cancergenome.nih.gov
The Genomic Data Commons – https://gdc.nci.nih.gov

How genetic testing impacts cancer treatment

Scientists recently used a combination of genomic analysis techniques to determine whether multiple testing methods provided better information for shaping treatment options for adult patients with several types of cancer. Tumor and normal cells underwent whole exome and targeted cancer panel sequencing, single nucleotide microarray genotyping and RNA-Seq. This combination of tests allowed scientists to search for gene mutations, changes in transcription rate, fusions between two genes and copy number variations that could influence treatment approaches and outcomes.

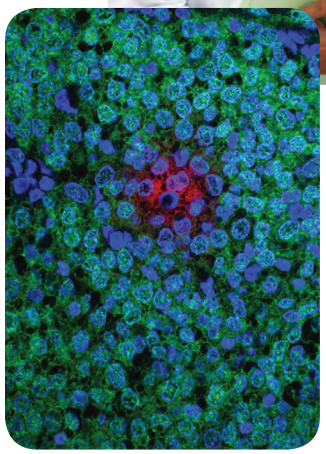
The combinatorial approach identified two to eight-fold more cancer-related mutations and actionable variants than would have been identified using commercially available cancer panels. Treatment-related recommendations were identified for 42 of 46 patients (91%). In several cases, panel testing identified no clinically actionable mutations, but the comprehensive system did. Additionally, as the RNA-Seq assay examined gene expression patterns, certain cancers (such as breast) could be further classified into subtypes – often with treatment implications.

Although the comprehensive analysis clearly provides more cancer-relevant information, there are several caveats to consider. Greater quantities of DNA are required, the analysis and interpretation times are significantly longer, and variants present in only a subset of cancer cells are more difficult to detect. The costs are substantially higher than for common targeted cancer gene panels

Liquid biopsies gain early approval

The FDA has approved the first so-called "liquid biopsy" test – a blood-based genetic test to diagnose non-small cell lung cancer that one day may decrease the need for expensive and painful tumor biopsies. Tumors shed small fragments of their DNA into the bloodstream. A couple of teaspoons of blood could contain as many as hundreds of genomes worth of these fragments. Liquid biopsies collect and analyze these fragments, searching for specific known genetic changes ("hot spot" testing) or sequencing the entire gene to gain a broader picture of potential variation.

Liquid biopsies are less invasive than tissue-based biopsies and may be particularly useful in cases where tissue testing isn't possible due to tumor location or the health of the patient. A recent study analyzed over 15,000 cancer patients using a commercially available liquid biopsy. Researchers were able to access tissue biopsy data on 386 of these patients. The tests matched 87% of the time, increasing to 98% if the blood and tissue samples were collected less than six months apart.



and few insurance companies currently reimburse laboratories for exome sequencing and RNA-Seq. It is likely these challenges will be overcome as technologies improve, but the authors suggest that current clinical pipelines should begin with a targeted panel and only proceed to more comprehensive approaches if actionable genetic mutations are not initially found.

REFERENCES: Uzilov, A. et al. Development and clinical application of an integrative genomic approach to personalized cancer therapy. Genome Medicine (2016) 8:62 doi:10.1186/s13073-016-0313-0.

In June 2016, the FDA approved the first liquid biopsy test to aid in clinical decision-making. The cobas® EGFR Mutation Test v2 detects specific mutations in the epidermal growth factor receptor (*EGFR*) gene. Developed by Roche, it is a "companion diagnostic test" for non-small cell lung cancer. Patients with certain *EGFR* mutations are candidates to receive a drug developed by Genentech that disrupts the cell growth activity of *EGFR*. The test was previously approved for DNA directly obtained from tumor and has been extended to cover circulating tumor DNA extracted from blood.

A number of challenges must be overcome before liquid biopsies become widespread. The process obtains less DNA meaning some samples may not have enough DNA for accurate testing. It's more difficult to analyze the entire genome, and there are hints that the technology may have a higher than expected rate of false positives. Additional clinical studies are needed to assess the sensitivity and specificity of this approach.

REFERENCES: Zill O.A. et al. Somatic genomic landscape of over 15,000 patients with advanced stage cancer from clinical next-generation sequencing analysis of circulating tumor DNA [abstract]. In: American Society of Clinical Oncology Annual Meeting 2016 June 3-7, Chicago (IL). Abstract nr LBA11501.

Cobas EGFR Mutation Test v2. (2016, June 2). Retrieved July 27, 2016, from http://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm504540.htm

In brief

Why elephants don't get cancer

Conventional wisdom suggests larger and longer-lived organisms would have higher rates of cancer, since having more cells and a greater number of cell divisions increases the risk to accumulate cancer-causing mutations. However, Utah researchers comparing cancer rates between zoo animals and humans found no correlation between cancer risk and body size, metabolic rate or life span. Of interest, elephants had a particularly low incidence of cancer, despite a 60+ year lifespan and being among the largest land mammals. An elephant's lifetime risk of cancer mortality was less than 5%, compared to almost 20% for humans.

Intrigued, researchers investigated the pachyderm genome, focusing on *TP53* – a gene frequently mutated in human cancers. The gene's protein product, known as p53, is a tumor suppressor that triggers apoptosis (cell death) in response to DNA damage. Humans carry two copies of the *TP53* gene, one on each chromosome. Strikingly, elephants have more than 40 copies of *TP53* and produce much more p53 protein. In the lab, when exposed to DNA-damaging radiation, elephant cells have a greater p53 response and trigger apoptosis at a lower threshold than human. It appears elephants have developed an efficient cancer prevention strategy by up-regulating their DNA damage response.

REFERENCES: Abegglen, L.M. et al. Potential Mechanisms for Cancer Resistance in Elephants and Comparative Cellular Response to DNA Damange in Humans. JAMA (2015) 314:1850-60. doi:10.1001/jama.2015.13134.

Molecular hallmarks of pancreatic cancer

Pancreatic cancer statistics are sobering: 7% of patients survive to five years, and the majority of patients die within a year of initial diagnosis. In over half of patients, the cancer is not identified until after the tumor has metastasized. The five-year survival rate increases to 26% when the tumor is diagnosed early and can be surgically removed, followed by chemotherapy and radiation therapy. Even so, many of these patients experience a relapse within two years. To identify cellular markers that may predict survival success, scientists recently analyzed gene expression patterns on a cohort of 51 pancreatic cancers. They identified a panel of 19 RNA transcripts that could distinguish between patients with short (average 7.2 months) and long (average 51.5 months) survival times after surgery. A study of pancreatic cancer cell lines found that some of the transcripts were differentially expressed between cells killed by the chemotherapy drug gemcitabine and those that were resistant to its effects. These results may provide a method to identify patients who would benefit a gemcitabine-based therapy.

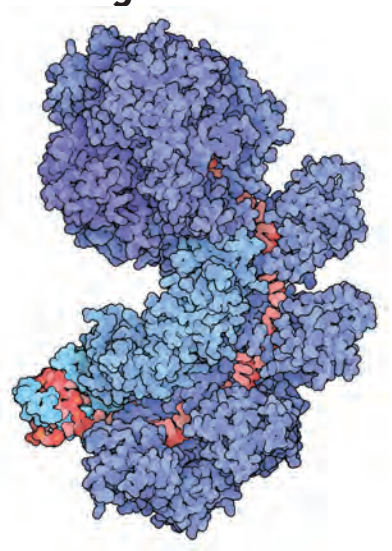
HudsonAlpha researchers and the labs of Richard Myers, PhD, and Sara Cooper, PhD contributed to this research

REFERENCES: Kirby, M.K., et al. RNA sequencing of pancreatic adenocarcinoma tumors yields novel expression patterns associated with long-term survival and reveals a role for ANGPTL4, Molecular Oncology (2016) dx.doi.org/10.1016/j.molonc.2016.05.004

NEW FINDINGS — ETHICAL, LEGAL AND SOCIAL IMPLICATIONS

Human genome editing

CRISPR-Cas9 gene editing moved closer to human application this past year with a summit to discuss its appropriate use, several proof of principles studies in model organisms and human cells lines, and the first approval for human clinical trials.



Gene Editing Summit

In December 2015, the United States National Academy of Sciences hosted a three-day summit on human genome editing. Discussions ranged from the philosophical (should editing even be allowed in humans) to the practical (how to measure off-target effects). At the conclusion of the meeting, the organizing committee released a statement voicing support for existing regulations that oversee basic and preclinical research as well as clinical applications involving gene editing in somatic cells. The committee noted that germline editing (impacting future generations) for clinical use would be "irresponsible," but they acknowledged gene editing in eggs and sperm may be useful for basic research, as long as the cells are not used to establish a pregnancy. The committee called for an international forum to solicit additional comments from various stakeholders including scientists, clinicians, patients, faith leaders, policymakers and the public.

REFERENCES: On Human Gene Editing: International Summit Statement. (2015, December 3). Retrieved July 26, 2016, from http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=12032015a

CRISPR and Muscular Dystrophy

Duchenne Muscular Dystrophy (DMD) is an X-linked muscle wasting disorder affecting 1 in 3,500-5,000 boys. It is caused by inactivating mutations in the dystrophin protein, which maintains muscle integrity. The dystrophin gene has 79 exons, and approximately 10% of patients have extra copies of one or more exons. Working with cells obtained from a DMD patient, scientists recently used CRISPR-Cas9 to remove a duplication of exons 18-30. The edited gene produced normal dystrophin. Separately, three research groups independently announced successful results using CRISPR-Cas9 in a mouse model of DMD. In each case, a virus transported the CRISPR-Cas9 system into muscle cells, where it removed a mutation-containing exon within the dystrophin gene. The cells transcribed and translated a shortened form of dystrophin, partially restoring function in the skeletal muscle of the mice. If such an approach became clinically feasible, an estimated 60% of patients with DMD might benefit from exon-skipping.

REFERENCES: Long, C. et al. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science (2016) 351:400-3 doi: 10.1126/science.aad5725.

Nelson, C.E. et al. In vivo genome editing improves muscle function in a mouse model of Duchene muscular dystrophy. Science (2016) 351:403-7 doi: 10.1126/science.aad5143.

Tabebordbar, M. et al. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science (2016) 351:407-11 doi:10.1126/science.aad5177.

Disclosing genetic information in a school setting

A recent California lawsuit highlights the challenge of protecting and sharing genetic information in an education setting.

The Background: Cystic fibrosis is a genetic disorder caused by the presence of two mutated copies of the *CFTR* gene. It affects the ability to clear mucus from the lungs, resulting in recurrent life-threatening bacterial lung infections. Although the disease itself is not contagious, individuals with CF can cross-infect each other with their specific lung bacteria. Because such cross-infection can be harmful, affected individuals are generally kept at least six feet apart and often are not allowed in the same room.

The Case: A teenage boy carries a known mutation for cystic fibrosis on one copy of his *CFTR* gene and an additional milder DNA change on the other copy. Although the boy has two *CFTR* mutations, his lung function is completely normal and he has no clinical symptoms of CF.

In 2012, the boy's mother enrolls him in middle school, noting his *CFTR* genetic and clinical results on the school's medical forms. Four weeks into the school year, a teacher discloses the boy's genetic information to the parents ("Mr. and Mrs. X") of a classmate during a parent-teacher conference. The X's child has active CF, and they are concerned about the risk of cross-infection. With a physician's support, the X's request the boy be transferred to another school. Over the objections of the boy's own parents, the school system agrees. Court action blocks the transfer and the case is settled before the transfer occurs.

Subsequently, a lawsuit is brought to address the school's disclosure of the boy's medical information and the decision to transfer. The trial court dismisses the case, finding that the school system had a reasonable basis for believing the boy posed a cross-infection risk. In early 2016, the boy's parents file an appeal with the Ninth Circuit Court of Appeals.

The Law: This case pits the genetic nondiscrimination rights of the boy against the perceived health and safety rights of the other students. It also demonstrates the limitations

CRISPR clinical trial approval

On June 22, 2016, the federal Recombinant DNA Advisory Committee gave approval for the first in-human use of CRISPR to modify the genes inside T-cells as a form of cancer immunotherapy. University of Pennsylvania researchers plan to knock out three genes, including the *PD-1* gene that suppresses a cell's ability to attack tumors. The goal is to edit the T-cells so they recognize the patient's cancer cells. When reinjected into the patient, the modified cells would hopefully mount an immune response and destroy the cancer. Assuming the trial receives approval from the University of Pennsylvania and the US Food and Drug Administration, it will enroll 15 patients to determine the safety and feasibility of the technology.

REFERENCES: Reardon, S. (2016, June 22). First CRISPR clinical trial gets green light from US panel. Retrieved July 26, 2016, from http://www.nature.com/news/first-crispr-clinical-trial-gets-green-light-from-us-panel-1.20137

of the Genetic Information Nondiscrimination Act of 2008 (GINA), which protects genetic information in the context of health insurance and employment but provides no protection under an educational setting. California has a state law (CalGINA) that broadens genetic nondiscrimination rights in the context of programs that involve state funding (such as education), but the family's claim isn't under this law.

Rather, they allege the school violated both the Americans with Disabilities Act and the Rehabilitation Act, which require public education and federally funded programs to be free from discrimination for those who have, have had, or are perceived to have a disability. The family is arguing that genetic information warrants protection as a "perceived disability," although to date no court cases have directly ruled on this point.

REFERENCES: Wagner, J. K. (2016, February 2). Genetic Discrimination Case Against School District is Appealed to Ninth Circuit. Retrieved July 25, 2016, from http://www.genomicslawreport.com/index.php/2016/02/02/genetic-discrimination-case-against-school-district-is-appealed-to-ninth-circuit/

In brief

LawSeqSM

As groups around the world explore how to incorporate genomic information into the clinic, federal and state laws regarding the use of genomics and the protections given to genomic information are often unclear and poorly understood. Accordingly, the NIH has awarded the first-ever grant focused on clarifying the existing state of genomic law, identifying policy gaps, and providing recommendations for the legal issues arising as genomic medicine becomes a standard component of patient care. The project, called LawSeqSM, will bring together a team of 25 legal, ethical and scientific experts from across the United States.

REFERENCES: Consortium on Law and Values in Health, Environment & the Life Sciences (2016). Pioneering Study will Establish the Legal Framework for Genomic Medicine [Press release]. Retrieved from https://consortium.umn.edu/sites/consortium.umn.edu/files/lawseq_umnvanderbilt_finalrev.pdf.

23andMe® relaunches testing

In 2013, citing concern that consumers would have difficulty understanding genetic results without assistance from a physician or genetic counselor, the FDA ordered personal genetics company 23andMe to stop selling their genetic tests directly to customers. This past October, after two years of working with the FDA, 23andMe was given the green light to once again offer certain direct-to-consumer genetic tests to tell customers whether they carry genetic variants for 36 rare diseases. The tests also include information about ancestry and wellness traits. Although 23andMe previously provided information about drug response and the risk for developing disorders such as Alzheimer disease, those findings are not included in the current test.

REFERENCES: 23andMe® Launches New Customer Experience – Reports Include Carrier Status that Meet FDA Standards, Wellness, Traits, and Ancestry [Press Release]. Retrieved from http://mediacenter.23andme.com/blog/new-23andme/.

Expanding the circle for inherited cancer screening

Historically, genetic screening for inherited forms of breast, ovarian and colorectal cancer risk have been limited to individuals with a significant personal or family history. Two groups recently introduced initiatives providing access to this information to a wider circle of individuals. HudsonAlpha's "Information is Power" initiative provides free inherited cancer risk screening to 30-year-old women in Madison County, Alabama, while Counsyl's "Get Ahead of Cancer" free screening program targets all women in the San Francisco Bay Area. Both groups offer the screening test in connection with the participant's physician, and genetic counseling is available once results are returned. Although neither group has published their experiences, the initiatives advocate for expanding access to groups traditionally excluded from these genetic tests.

REFERENCES: http://hudsonalpha.org/information-is-power and http://getaheadofcancer.counsyl.com/

In brief

Genomics for better bread

A team of breeders and geneticists has developed a series of DNA markers for bread wheat to predict the presence of traits that are important for breadmaking. Historically, bread wheat breeding programs emphasized grain yield and plant hardiness. Because quality traits like dough strength and loaf volume weren't examined until late in the breeding cycle, years of work could be wasted developing a variety of wheat that didn't actually produce a good loaf of bread. The new DNA markers evaluate bread quality earlier in the selection process, identifying which varieties should be discarded and which are worthy of further development.

REFERENCES: Battenfield, S.D. et al. Genomic Selection for Processing and End-Use Quality Traits in the CIMMYT Spring Bread Wheat Breeding Program. The Plant Genome (2016) doi: 10.3835/plantgenome2016.01.0005.

Genome Editing knocks the horns off dairy cattle

Unlike beef cattle, most dairy cattle have horns. These are often cauterized or removed early in life for economic and safety reasons. Hornless cattle require less space and are less likely to injure other cattle or human handlers.

Attempts have been made to develop hornless dairy cattle by breeding them with naturally hornless beef cattle. Unfortunately, the beef cattle introduce unwanted traits such as lower milk production or slower growing calves. Breeding these traits out of a population is a lengthy and expensive process.

New genome editing techniques may provide a comparatively faster and less costly method to silence the horn-producing gene. Using a process known as TALEN, researchers focused on a section of DNA known to regulate horn formation. An embryonic cell from a horned Holstein dairy cow was edited to mimic the DNA sequence found in hornless Angus beef cattle. This resulted in the birth of two healthy, hornless male Holstein calves. These calves are not transgenic and have only Holstein DNA. Their offspring will display all Holstein traits – minus the horns. The regulatory issues surrounding gene editing are somewhat unclear, but study suggests gene editing may offer a powerful way to introduce genetically important traits into livestock.

REFERENCES: Carlson D.F. et al. Production of hornless diary cattle from genome-edited cell lines. Nature Biotechnology (2016) 34:479-481. doi:10.1038/nbt.3560.

GMO labeling bill becomes law

A new law will establish national standards for labeling of foods containing genetically engineered (GE) ingredients. The law, passed by Congress and signed by President Obama in July 2016, describes "bioengineered" foods as those that *"contain genetic material that has been modified through in vitro recombinant…DNA techniques; and for which the modification could not otherwise be obtained through conventional breeding or found in nature."*

Over the next two years, the US Department of Agriculture will determine which foods require labeling. The agency faces a number of decisions around this issue including whether to include processed foods – such as high fructose corn syrup made from the sugars found in genetically modified corn – that are derived from GE ingredients but contain no "genetic material." It's also not clear whether labeling extends to plants modified through new gene editing techniques that do not insert foreign DNA sequences or to plants with genetic changes that mimic those found naturally in other living organisms.

Food manufacturers have one year after the guidelines are finalized to implement labeling. This can be through a symbol or label, but also a 1-800 number, website address or QR code printed on the package that directs consumers to more details about the ingredients. This is a point of contention for many pro-labeling groups, who argue that technology-dependent options fail to provide consumers with clear, on-package information. The law preempts all existing state-level labeling laws – such as those in Vermont, Connecticut and Maine – even if the state laws are more restrictive than the federal standards.

REFERENCES: Full text of the National Bioengineered Food Disclosure Law: https://www.congress.gov/bill/114th-congress/senate-bill/764/text?resultIndex=1

Gene editing for allergy-free eggs

Up to 3% of children are allergic to eggs, with reactions varying from skin hives to life-threatening anaphylaxis. In addition to food-based concerns, these individual must be cautious about receiving certain influenza and other vaccines containing trace amounts of egg proteins. The two major allergenic egg proteins are known as *OVA* and *OVM*, both found in egg whites. Although most children eventually outgrow their egg allergy, early evidence suggests individuals allergic to *OVM* are less likely to do so.

In an attempt to produce eggs without allergic reactions, researchers have turned to gene editing techniques. In 2014, the *OVA* gene was silenced using a process known as TALEN. This year a Japanese team silenced *OVM* through CRISPR-Cas9. Scientists began with primordial germ cells (PGC), embryonic cells that develop into eggs or sperm. In the lab, PCGs underwent CRISPR-Cas9 editing, producing frame-shift deletions in the *OVM* gene that disrupted OVM protein function. The modified PGCs were injected into male chicken embryos, where they migrated to the developing testes and developed alongside pre-existing PGCs. After hatching and maturing to adulthood, roosters were tested to determine their percentage of *OVM*-edited sperm. High-frequency roosters were mated to typical hens, producing offspring with a single copy of the *OVM*-edited gene. Additional rounds of mating resulted in chickens homozygous for edited *OVM*.

Seagrass Genome Analysis

Despite their name, seagrasses are not "grasses" but underwater flowering plants descended from land-dwelling angiosperms. Scientists recently sequenced the genome of the seagrass *Zostera marina*, uncovering clues about how it evolved to live in shallow saltwater. The plant has lost genes that code for a number of no-longer required land-based traits: defending against insects, responding to UV damage and producing stomata – the pores on leaves that prevent excessive water loss. Simultaneously, the plant has modified existing genes, allowing it to regulate ion exchange under high salt conditions, directly exchange oxygen and carbon dioxide through the epidermis of the leaves, and better capture light to conduct photosynthesize under dimly lit conditions.

HudsonAlpha researchers Jane Grimwood, PhD, Jeremy Schmutz and members of the GSC lab contributed to this work.

REFERENCES: Olsen, J. et al. The genome of the seagrass Zostera marina reveals angiosperm adaptations to the sea. Nature (2016) 530:331-335. doi:10.1138/nature16548.

Scientists exploited differences in feather color to trace edited *OVM* inheritance. The edited donor PGCs came from a black-feathered Barred Plymouth Rock but the recipient male embryos and the hens for subsequent mating were white-feathered Leghorns. Any chickens born carrying the edited *OVM* gene also had black feathers, while those with the typical *OVM* gene were white.

This was a proof-of-concept for the use of CRISPR-Cas9 in chickens. The next step is to determine whether the eggs from *OVM*-edited hens induce an allergic reaction. The researcher's eventually hope to produce chickens silenced at both *OVA* and *OVM*, hopefully resulting in low-allergenicity eggs.

REFERENCES: Oishi, I. et al. Targeted mutagenesis in chicken using CRISPR/Cas9 system. Scientific Reports (2016) 6, 23980; doi: 10.1038/srep23980.

In brief

Soil inoculations for land restoration

Erosion, over-farming and poor land management can deplete soil resources, leaving barren fields that require decades to return to health. A team of ecologists from the Netherlands has identified soil inoculations as a way to potentially accelerate the restoration process. Working with abandoned farmland, some plots were spread with a thin (1 cm) layer of donor topsoil. Other plots removed the top 60 cm of existing dirt before adding the donor soil. The transplanted soil came from either a grassland or a heath – a tract of land characterized by low woody shrubs. The team added seeds from 30 different plants and monitored the plots for six years.

All the plots that received donor soil fared better than untouched controls, but the plots where the topsoil had been removed showed even greater improvement. The source of the soil steered the restoration towards the donor's ecosystem (grassland vs. heath). The authors believe microbes in the donor soil drive this transformation – analogous to the way human microbiome transplants can reestablish a population of healthy gut microbes.

REFERENCES: Wubs, E.R. et al. Soil inoculation steer restoration of terrestrial ecosystems. Nature Plants (2016) 11:16107- doi:10.1126/science.aad6253.

Yeast: the silent partner in lichens

Almost 150 years ago, scientists recognized that lichens occur when a fungus enters into a symbiotic relationship with an algae or cyanobacterium. The fungus forms the structural component of the lichen, providing protection from the elements. The algae or cyanobacterium uses photosynthesis to produce food. Scientists recently made the unexpected discovery that in many instances this symbiotic twosome is actually a trio. mRNA transcripts from various lichens identified the presence of a previously unknown species of yeast. Fluorescent visualization techniques showed the yeast are embedded in the cortex of these lichens. This finding could explain why lichens composed of the same fungal and phytosynthetic partner have very different physical features – the yeast contributes metabolites that impact fungal traits.

REFERENCES: Spribille, T. et al. Basidiomycete yeasts in the cortex of ascomycete macrolichens. Science (2016) doi:10.1126/science.aaf8287.

Types of Genetic Testing

Genetic tests use various laboratory methods to examine your DNA. They range from studying a single letter of the DNA sequence to analyzing the entire genome.

Prenatal:

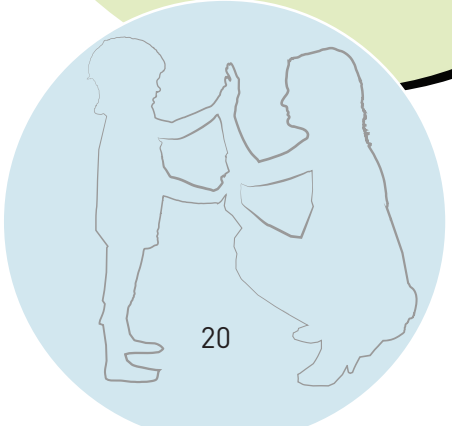
There are two types of prenatal tests related to genetic conditions: screening and diagnostic. **Prenatal screening** tests generally measure the concentration of specific proteins or hormones in the mother's bloodstream to identify the risk of having a child with certain genetic disorders, such as Down syndrome. Recent noninvasive techniques are even able to collect and test pieces of fetal DNA that circulate in the mother's blood. These approaches do not diagnose a disorder, but signal that further testing should be considered. **Diagnostic tests** directly analyze fetal DNA, often obtained through invasive procedures such as amniocentesis or chorionic villus sampling (CVS). Prenatal diagnostic tests may study the number and structure of the chromosomes in fetal cells, or identify the sequence of a specific gene or region of the genome.

Lifestyle:

Genetic testing to identify lifestyle and wellness-related traits is an emerging field. These tests provide information on topics ranging from earwax type and personality style to nicotine dependence and muscle performance. Since many of these traits are influenced by multiple genetic and environmental factors, the accuracy and utility of these tests is unclear.

Pediatric:

Between two and three percent of all children have a physical birth defect or clinical disorder. These may be seen at birth, or become evident during childhood. All infants born in the United States undergo **newborn screening** to identify disorders that can affect a child's long-term health. Using a few drops of blood from a baby's heel, clinical laboratories test for at least 29 diseases, most of which are genetic in nature. Children who have symptoms of a genetic disorder or do not meet developmental milestones may undergo **diagnostic genetic testing** to identify or rule out a specific condition. This may be a targeted test for a specific mutation, a test of a single gene or a handful of genes known to be associated with the child's symptoms, or a genome-wide analysis to more broadly search for answers.



Adult:

Genetic testing in adults generally falls into one of three categories:

Diagnostic testing seeks to identify disease-causing mutations to explain a patient's existing set of symptoms.

Predictive/Presymptomatic testing detects mutations for disorders that often appear later in life. These tests are usually ordered for individuals who have a family history of a disease but have no signs of that disease at the time of testing.

Carrier testing identifies people who carry a single copy of a mutation that - when present in two copies - causes disease. The individual is healthy, but could pass along the mutation to a child. Couples may decide to have carrier testing to determine their risk of having a child with certain genetic conditions.

Pharmacogenomic:

Some genes are responsible for how the body processes medications. Pharmacogenomic testing looks for changes in those genes and seeks to correlate that information to a person's response to medications. It seeks to predict the most effective drug at the right dose, as well as identify those drugs that may cause harmful side effects. For example, warfarin is a drug that helps prevent blood clots, strokes and heart attacks. Individuals who have specific genetic variants require lower doses for therapy. Similar variants are associated with medications for depression and chemotherapy. While these types of tests are currently used for only a few health issues, they will become increasingly important in the years ahead. At the moment, no single pharmacogenomics test can predict an individual's response to all medications. In addition, no such tests are available for most over the counter medications.

Ancestry:

Because certain patterns of DNA variation are more commonly found among individuals of specific backgrounds, DNA analysis can shed light on where an individual's ancestors likely came from. The more of these patterns that two people share, the more closely related they are. Genetic ancestry testing usually examines DNA variation on the Y chromosome (to study the male line), the mitochondria (for details about the female line) or single letter changes throughout the genome (to estimate the overall ethnic background). This type of testing does not reveal any medical or health-related information and while it may provide geographic origins for distant ancestors, it cannot provide the names of those ancestors.

Cancer:

Genetic testing may be useful as a **predictive test** for individuals with a family history of certain types of cancers (such as breast, ovarian and colorectal). A positive test result indicates the person has inherited a genetic mutation that significantly increases his or her lifetime risk of developing cancer. These individuals may have more frequent cancer screenings or chose to undergo surgery to reduce the cancer risk. In some cases, this type of genetic testing for inherited mutations may also be appropriate when cancer has already been detected. In addition, **genetic testing of the tumor cells** may be requested to determine which cancer-causing mutations have been acquired by the tumor. This knowledge may aid in diagnosis and shape a physician's choice of therapy to treat the cancer.

DNA Profiling:

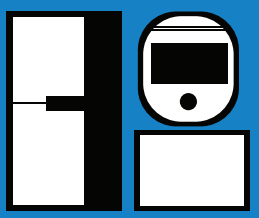
DNA profiling identifies an individual's unique pattern of DNA variation and is often used in parentage testing and criminal investigations. A parent shares 50% of his or her genetic variation with a child. A paternity or maternity test compares the DNA patterns between the child and alleged parents to look for evidence of genetic sharing. Forensic DNA testing can link a perpetrator or victim to a crime scene as well as exonerate individuals convicted of crimes they did not commit. There are limitations to this type of testing, including the inability to distinguish between identical twins and the challenge of assessing samples with degraded or low amounts of DNA.

How does a human genome get sequenced?

Your genome is your unique sequence of DNA, **3 billion letters** long. It's found in almost every cell in your body.

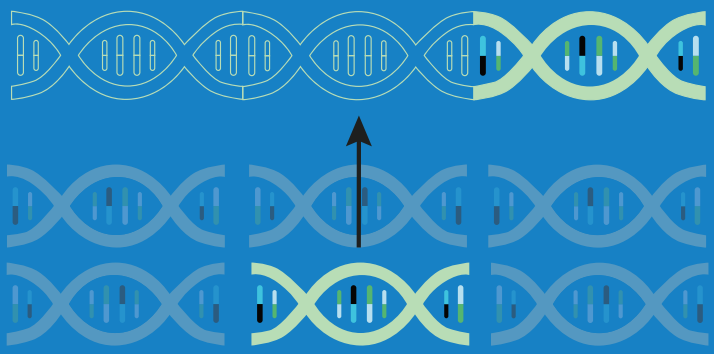
The letters **A**, **T**, **C** and **G** represent the chemical elements, or bases, of DNA.

1 DNA is extracted from a sample and loaded on to a **sequencing machine**.



2 The machine determines the sequence of short pieces of DNA, 150 letters long. These are called **reads**.

3 The 'reads' from the sequencing machine are matched to a reference sequence. This is called **mapping**.



4 Analysis

Within the 3 billion letters in your genome are 20,000 genes. These make up about 2% of the sequence. The position of most of our genes is known, and is marked on the **reference sequence**.

Every person has millions of differences (called **variants**) from the reference sequence.

Most of these difference are harmless – they are the reason we are different from each other. Some differences could be causing a disease.

Bioinformatics specialists use a variety of tools and techniques to filter these differences down from millions to just a handful that could be harmful.

If it is not clear which difference is causing disease, researchers anaylze the genome further.

CRISPR-Cas9

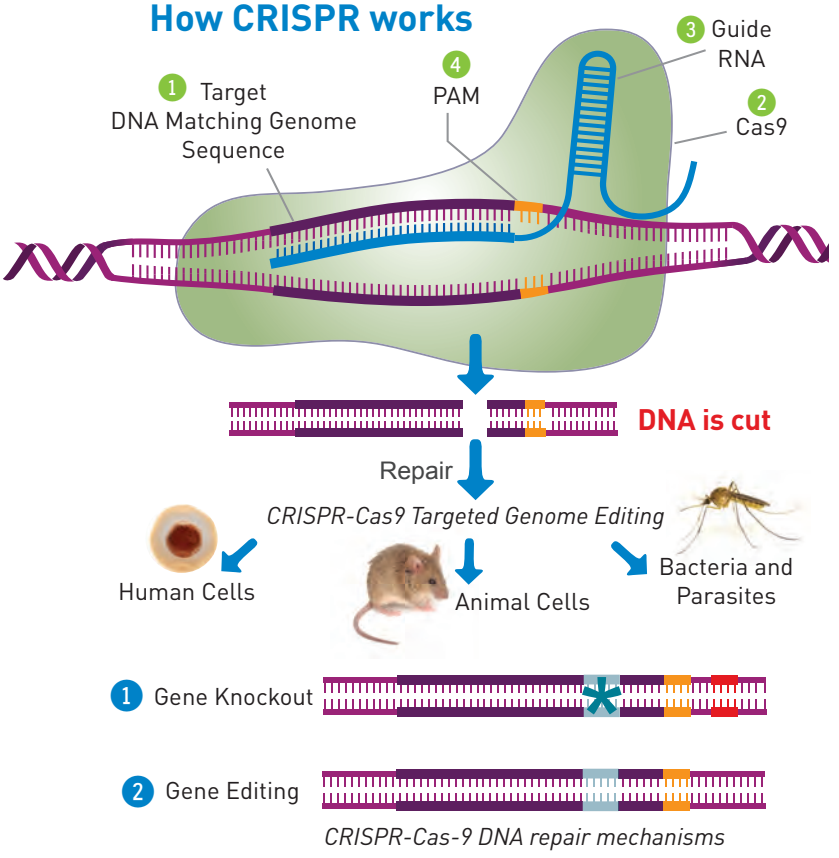
A powerful new tool for genome editing

Overview

Tools that modify the genome allow scientists to explore the impact of DNA mutations and engineer changes that create drug-producing bacteria, disease-resistant crops or life-saving genetic therapies. A technological breakthrough known as CRISPR-Cas9 is revolutionizing this process of genome editing. ***CRISPR*** stands for ***Clustered Regularly Interspaced Palindromic Repeats*** and is part of the bacterial immune system. The CRISPR-Cas9 tool is analogous to the "find and replace" function in word processing software. First developed in 2012, **the process is based on naturally occuring bacterial defense mechanisms that recognize and cut up genetic sequence from invading viruses.**

Because CRISPR-Cas9 is easy to implement and relatively inexpensive, it is quickly showing widespread use. If some key challenges can be overcome, it holds great promise to transform basic science, medicine and biotechnology.

How CRISPR works



CRISPR-Cas-9 DNA repair mechanisms

There are 4 key components of a CRISPR-Cas9 system:

Target DNA: This is the region of the genome to be modified.

Cas9: This bacterial enzyme unzips and cuts the target DNA. To date, most approaches use the Cas9 protein found in the bacteria *Streptococcus pyogenes.*

PAM sequence: PAM stands for Protospacer Adjacent Motif. It is part of the target sequence DNA and is one of the factors that is required to define the cutting site.

Guide RNA: A short fragment of RNA binds to Cas9 and contains a recognition sequence that matches the target. With different guide RNA sequences, the Cas9 enzyme can be directed to recognize almost any DNA sequence.

The guide RNA leads Cas9 to the desired location in the genome, binds the target sequence and triggers Cas9 to cut both strands of target DNA. This double stranded break provides the opportunity to edit the genome. *(See the various applications on the next page).*

Researchers use 2 existing cellular DNA repair mechanisms, depending on the goal of editing:

Gene Knockout: An error-prone repair system often inadvertently inactivates the gene. By observing the impact of a gene's absence, scientists can better understand its function.

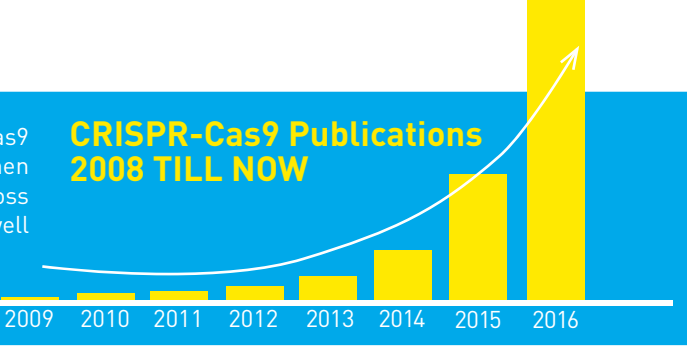
Gene Editing: An alternative repair system replaces the broken DNA, using an intact copy as a template. By providing a template of their own making, researchers can introduce specific new DNA changes or insert entirely novel sections of DNA.

CRISPR-Cas9 Usage at HudsonAlpha

1 **The Myers Lab is analyzing how certain proteins bind to DNA and alter the activity of various genes. As part of this research, the proteins are attached to antibodies. Unfortunately, many proteins do not form strong antibody linkages, making analysis difficult. Using CRISPR-Cas9, instructions for a stronger attachment site are inserted into the protein's genetic recipe. This slight modification doesn't impact the protein's function, but ensures reliable antibody binding.**

2 **TransOMIC technologies, a HudsonAlpha associate company, sells CRISPR-Cas9 reagents. They have used their experience with vector design to advance reagents so that scientists can select from a wide array of options to customize their genome editing experiments.**

The first successful application for CRISPR-Cas9 gene editing was published in late 2012. Since then the technology has gained widespread use across several biotechnology applications, resulting in well over 1,000 publications.



Potential Applications for Genome Editing

The CRISPR-Cas9 method has many applications in basic research, agriculture, drug development and eventually treating humans with genetic diseases.

Drug Development

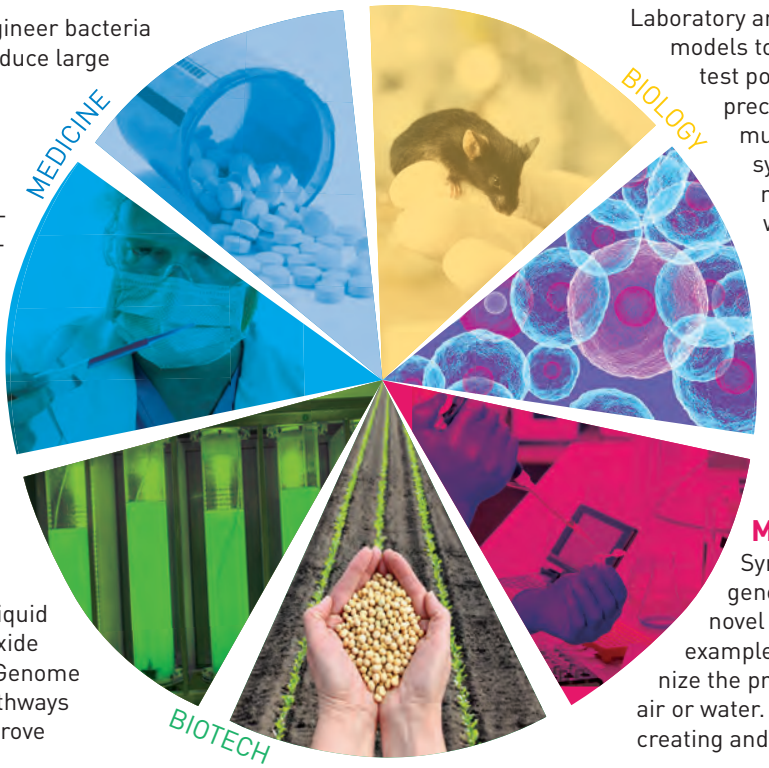
CRISPR-Cas9 technology can engineer bacteria to rapidly and cost-effectively produce large quantities of drugs.

Gene Therapy

Using the template-based repair system, CRISPR-Cas9 has experimentally been used to correct genetic mutations responsible for inherited disease. For example, a cataract-causing mutation in the mouse crygc gene was successfully corrected by injecting the CRISPR-Cas9 components and a normal copy of the gene into fertilized mouse eggs.

Fuel

A number of microbes generate liquid biofuels by converting carbon dioxide or by fermenting plant material. Genome editing can modify the genetic pathways controlling fuel production to improve efficiency and yield.



Animal Models

Laboratory animals are often used as models to understand disease and to test potential therapies. Gene editing precisely reproduces human mutations in these model systems. These specific animal models can be developed in weeks, rather than years.

Genetic Variation

Using CRISPR-Cas9, researchers have sequentially knocked out nearly every gene in human laboratory-grown cells to better understand their function.

Materials

Synthetic biology combines genetic instructions to build novel biological tools – for example living biosensors that recognize the presence of pollutants in the soil, air or water. Genome editing speeds the creating and assembling process.

Food

CRISPR-Cas9 can modify plant genomes more quickly and efficiently than classical plant breeding. This can increase yields and deliver protection from pests and disease, without inserting the foreign DNA sequences required by genetically modified organisms (GMOs).

References:
Doudna, J.A. and Charpentier, E., *Science* 346:1258096-1-9 (2014).
Hsu, P.D. et al. *Cell* 157:1262-78 (2014).
Belhaj K., et al. *Plant Methods* 9:39 (2013)

Existing Challenges of CRISPR-Cas9

Off-targeting

Under certain conditions the Cas9 enzyme may tolerate mismatches between the guide and genomic sequences, cutting at sites outside the target region. Informatics tools can predict off-target effects for a specific guide RNA molecule and techniques to increase Cas9 specificity and minimize off-target cleavage are being developed.

Effective Delivery

The CRISPR-Cas9 editing technology must be correctly delivered to the appropriate cells and tissues. To correct a disease-causing mutation, the template-based repair method must be used. The efficiency of this type of repair is generally low and, successful gene editing therapies must identify ways to significantly boost repair rates.

Regulatory and Safety

Although CRISPR-Cas9 genome editing has shown immense promise in the laboratory, it is still a relatively new technique. The long-term implications must be thoroughly characterized. A number of regulatory and safety hurdles must be cleared before a disease therapy or edited agricultural crop will win approval.
Lastly, detailed conversations are clearly needed regarding the ethical, social and policy implications of altering human genomes.

COURSE OF STUDY CONNECTED TO GUIDEBOOK TOPICS

Course	Objective and Applicable Subheading		Linking Scientific Concept
Biology	**1**	Use models to compare and contrast how the structural characteristics of carbohydrates, nucleic acids, proteins and lipids define their function in organisms.	DNA Sequencing, RNA and Protein Analysis, Recombinant DNA and Genetic Engineering, Synthetic Biology, Pharmacogenomics
	2	Obtain, evaluate and communicate information to describe the function and diversity of organelles and structures in various types of cells (e.g., muscle cells having a large amount of mitochondria, plasmids in bacteria, chloroplasts in plant cells).	See HudsonAlpha iCell (pg 5) RNA and Protein Analysis, Gene Therapy and RNAi, Stem Cells
	3	Formulate an evidence-based explanation regarding how the composition of deoxyribonucleic acid (DNA) determines the structural organization of proteins.	DNA Sequencing RNA and Protein Analysis RNA and Protein Analysis, Bioinformatics, Application, Recombinant DNA and Genetic Engineering, Synthetic Biology, Therapeutic Approaches: Gene Therapy, Copy Number Variation, Personal Genome Analysis
	3a	Obtain and evaluate experiments of major scientists and communicate their contributions to the development of the structure of DNA and to the development of the central dogma of molecular biology.	DNA Sequencing, Bioinformatics, Genetic Information Nondiscrimination Act, Personalized Medicine, Pharmacogenomics See also Biotechnology Timeline (pg 5)
	3b	Obtain, evaluate, and communicate information that explains how advancements in genetic technology (e.g., Human Genome Project, Encyclopedia of DNA Elements [ENCODE] project, 1000 Genomes Project) have contributed to the understanding as to how a genetic change at the DNA level may affect proteins and, in turn, influence the appearance of traits.	DNA Sequencing, RNA and Protein Analysis, Bioinformatics, Copy Number Variation, Genetics of Eye Color, Personalized Medicine, Personal Genome Analysis, Studying the Genome to Understand the Sequence, Synthetic Biology, Therapeutic Approaches: RNAi
	3c	Obtain information to identify errors that occur during DNA replication (e.g., deletion, insertion, translocation, substitution, inversion, frame-shift, point mutations).	Diagnosing Chromosome Disorders, Personal Genome Analysis, Noninvasive Prenatal Diagnosis
	4	Develop and use models to explain the role of the cell cycle during growth and maintenance in multicellular organisms (e.g., normal growth and/or uncontrolled growth resulting in tumors).	Cancer, Stem Cells, Diagnosing Chromosome Disorders
	10	Construct an explanation and design a real-world solution to address changing conditions and ecological succession caused by density-dependent and/or density-independent factors.	Agriculture - Sequencing Plant Genomes for Food and Bioenergy Needs, Genetically Modified Crops (biofuels and GM crops may play in a role in student developed real world solutions to ecological problems)
	11	Analyze and interpret data collected from probability calculations to explain the variation of expressed traits within a population.	Copy Number Variation, Criminal Justice and Forensics, Epigenetics, Genetics of Eye Color, Noninvasive Prenatal Diagnosis, Therapeutic Approaches: Gene Therapy and RNAi, Noninvasive Prenatal Diagnosis, Diagnosing Chromosome Disorders, Epigenetics
	11a	Use mathematics and computation to predict phenotypic and genotypic ratios and percentages by constructing Punnett squares, including using both homozygous and heterozygous allele pairs.	Genetics of Eye Color, Agriculture, Criminal Justice and forensics, Identifying Genetic Influence on Disease, Epigenetics
	11b	Develop and use models to demonstrate co-dominance, incomplete dominance and Mendel's laws of segregation and independent assortment.	Genetics of Eye Color, Diagnosing chromosome Disorders, Identifying Genetic Influence on Disease, Epigenetics

Course	Objective and Applicable Subheading	Linking Scientific Concept
Biology	**11c** Analyze and interpret data (e.g., pedigree charts, family and population studies) regarding Mendelian and complex genetic disorders (e.g., sickle-cell anemia, cystic fibrosis, type 2 diabetes) to determine patterns of genetic inheritance and disease risks from both genetic and environmental factors.	Copy Number Variation, Criminal Justice and Forensics, Genetic Information Nondiscrimination Act, Identifying Genetic Influence on Disease, Infectious Disease, Personal Genome Analysis, Personalized Medicine, Pharmacogenomics, Epigenetics, Genetics of Eye Color, Studying the Genome to Understand the Sequence
	12 Develop and use a model to analyze the structure of chromosomes and how new genetic combinations occur through the process of meiosis.	Diagnosing Chromosome Disorders, Epigenetics, Noninvasive Prenatal Diagnosis
	12a Analyze data to draw conclusions about genetic disorders caused by errors in meiosis (e.g., Down syndrome, Turner syndrome).	Diagnosing Chromosome Disorders, Non-invasive Prenatal Diagnosis, Personal Genome Analysis, Therapeutic Approaches: Gene Therapy and RNAi, Copy Number Variation, Personalized Medicine, Pharmacogenomics
	13a Engage in argument to justify the grouping of viruses in a category separate form living things.	Infectious Disease
	14 Analyze and interpret data to evaluate adaptations resulting from natural and artificial selection that may cause changes in populations over time (e.g., antibiotic-resistant bacteria, beak types, peppered moths, pest-resistant crops).	Comparative Genomics, Infectious Disease
	15 Engage in argument from evidence (e.g., mathematical models such as distribution graphs) to explain how the diversity of organisms is affected by overpopulation of species, variation due to genetic mutations, and competition for limited resources.	Comparative Genomics
	16 Analyze scientific evidence (e.g., DNA, fossil records, cladograms, biogeography) to support hypotheses of common ancestry and biological evolution.	Comparative Genomics, Studying the Genome to Understand the Sequence
Anatomy and Physiology	**3a** Analyze the effects of pathological conditions (e.g. burns, skin cancer, bacterial and viral infections, chemical dermatitis) to determine the body's attempt to maintain homeostasis.	Cancer, Infectious Disease, Diagnosing Chromosome Disorders, Identifying Genetic Influence on Disease, Studying the Genome to Understand the Sequence
	6a Use scientific evidence to evaluate the effects of pathology on the nervous system (e.g., Parkinson disease, Alzheimer disease, cerebral palsy, head trauma) and argue possible prevention and treatment options.	Personal Genome Analysis, Personalized Medicine, Identifying Genetic Influence on Disease, Pharmacogenomics
	6b Design a medication to treat a disorder associated with neurotransmission, including mode of entry into the body, form of medication, and desired effects.	Personalized Medicine, Identifying Genetic Influence on Disease, Pharmacogenomics
	9a Engage in argument from evidence describing how environmental (e.g., cigarette smoke, polluted air) and genetic factors may affect the respiratory system, possibly leading to pathological conditions (e.g., cystic fibrosis).	Cancer, Personal Genome Analysis, Identifying Genetic Influence on Disease, Personalized Medicine, Studying the Genome to Understand the Sequence

COURSE OF STUDY CONNECTED TO GUIDEBOOK TOPICS

Course	Objective and Applicable Subheading	Linking Scientific Concept
Environmental Science	**1** Investigate and analyze the use of nonrenewable energy source (e.g., fossil fuels, nuclear, natural gas) and renewable energy sources (e.g., solar, wind, hydroelectric, geothermal) and propose solutions for their impact on the environment.	Agriculture: Sequencing Plant Genomes for Food and Bioenergy Needs
	6 Obtain, evaluate and communicate information to describe how human activity may affect biodiversity and genetic variation of organisms, including threated and endangered species.	Comparative Genomics, Recombinant DNA and Genetic Engineering, Agriculture
AP Biology ***Big Idea 1*** **Evolution**	***The process of evolution drives the diversity and unity of life*** Enduring Understanding 1.A. Change in the genetic make-up of a population over time is evolution. Enduring Understanding 1.C Organisms are linked by lines of descent from common ancestry.	DNA Sequencing, RNA and Protein Analysis, Bioinformatics, Comparative Genomics
Big Idea 2 **Free Energy and Molecular Building Blocks**	***Biological systems utilize free energy and molecular building blocks to grow, to reproduce, and to maintain dynamic homeostasis*** Enduring Understanding 2C. Organisms use feedback mechanisms to regulate growth and reproduction, and to maintain dynamic homeostasis. Enduring Understanding 2.D. Growth and dynamic homeostasis of a biological system are influenced by changes in the system's environment. Enduring Understanding 2.E. Many biological processes involved in growth, reproduction and dynamic homeostasis include temporal regulation and coordination.	Cancer Identifying Genetic Influence on Disease, Studying the Genome to Understand the Sequence
Big Idea 3 **Information**	***Living systems store, retrieve, transmit and respond to information essential to life processes.*** Enduring Understanding 3.A. Heritable information provides for continuity of life. Enduring Understanding 3.B Expression of genetic information involves cellular and molecular mechanisms. Enduring Understanding 3.C. The processing of genetic information is imperfect and is a source of genetic variation. Enduring Understanding 3.D. Cells communicate by generating, transmitting and receiving chemical signals. Enduring Understanding 3.E. Transmission of information results in changes within and between biological systems.	DNA Sequencing, RNA and Protein Analysis, Pharmacogenomics, Noninvasive Prenatal Diagnosis, Genetic Information Nondiscrimination Act, Identifying Genetic Influence on Disease, Recombinant DNA and Genetic Engineering, Studying the Genome to Understand the Sequence, Copy Number Variation, Epigenetics, Stem Cells, Synthetic Biology, Therapeutic Approaches: RNAi
Big Idea 4 **Biological Systems**	***Biological systems interact and these systems and their interactions possess complex properties*** Enduring Understanding 4.A. Interactions within biological systems lead to complex properties.	Identifying the Genetic Influence on Disease

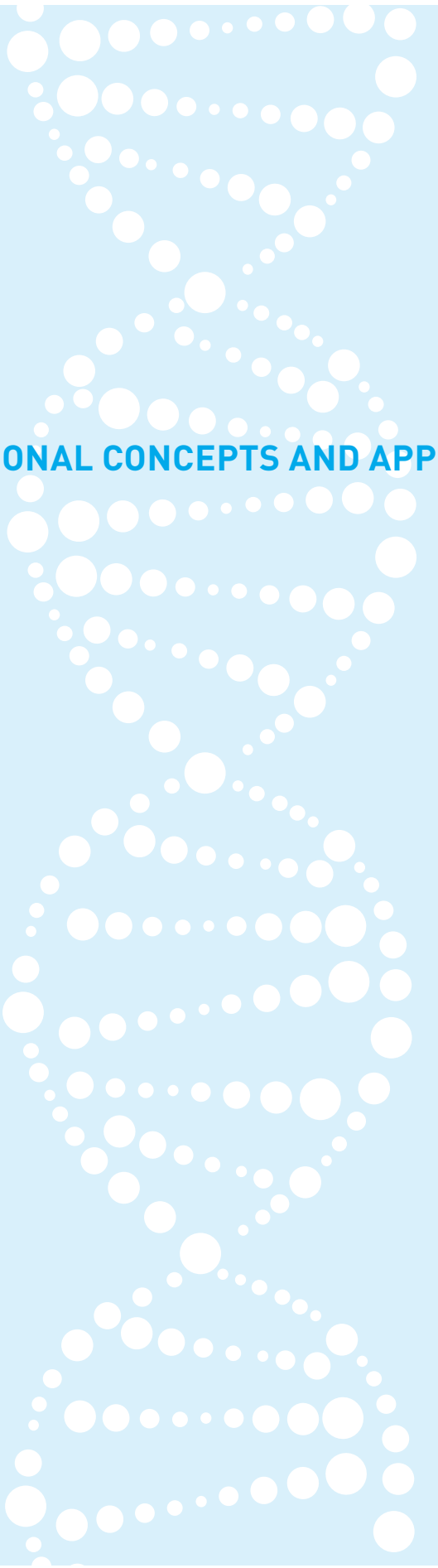
Course	Objective and Applicable Subheading	Linking Scientific Concept
Health	**5** Evaluate negative and positive impacts of technology on health.	Agricultural, Cancer, Identifying Genetic Influence on Disease, Noninvasive Prenatal Diagnosis, Personalized medicine, Pharmacogenomics, Recombinant DNA and Genetic Engineering, Stem Cells, Synthetic Biology
	6 Discuss valid and essential information for the safe use of consumer goods and health products.	Agricultural, 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
Technology Education	**26** Explain uses and advantages of databases.	Bioinformatics
	27 Apply appropriate techniques for producing databases.	Bioinformatics
Agriscience	**10** Determine characteristics and functions of plants. Explain how agricultural crops can be utilized as alternative fuel sources.	Agricultural applications
Forensic and Criminal Investigations	**7** Describe presumptive and confirmatory forensic tests. *Examples: blood type comparison, DNA testing*	Criminal Justice and Forensics
	8 Describe the importance of genetic information to forensics Using the process of gel electrophoresis for deoxyribonucleic acid (DNA) fingerprinting.	Bioinformatics, Criminal Justice and Forensics
Foundations of Health Sciences	**10** Recognize legal responsibilities, limitations, and implications within the health care delivery setting. *Examples: Patients' Bill of Rights, legal documentation requirements, Health Insurance Portability and Accountability Act (HIPAA)*	Genetic Information Nondiscrimination Act, Personal Genome Analysis
Health Informatics	**5** Describe legal and ethical regulations as they relate to health informatics. *Examples: Patients' Bill of Rights, legal documentation requirements, Health Insurance Portability and Accountability Act (HIPAA)*	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 Biotechnology Timeline (pg 5) Agricultural Applications, Bioinformatics, Criminal Justice and Forensics, Diagnosing Chromosome Disorders, DNA Sequencing, Pharmacogenomics, See also Biotechnology Timeline (pg 5)
	4 Correlate key cellular components to function.	See HudsonAlpha iCell® (pg 5)
	5 Describe the process of meiosis and the cell cycle, including the hereditary significance of each.	Cancer, Diagnosing Chromosome Disorders, Noninvasive Prenatal Diagnosis, Stem Cells
	8 Describe occurrences and effects of sex linkage, autosomal linkage, crossover, multiple alleles, and polygenes.	Cancer, Copy Number Variation, Genetics of Eye Color, Identifying Genetic Influence on Disease

continued on pg. 28

COURSE OF STUDY CONNECTED TO GUIDEBOOK TOPICS

Course	Objective and Applicable Subheading		Linking Scientific Concept
Introduction to Biotechnology	**9**	Describe the structure and function of deoxyribonucleic acid (DNA), including replication, translation, and transcription.	Recombinant DNA and Genetic Engineering, Studying the Genome to Understand the Sequence
		Applying the genetic code to predict amino acid sequence	Bioinformatics
		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.	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
	16	Explain the historical significance of plant biotechnology.	Agricultural Applications, Comparative Genomics; See also Biotechnology Timeline (pg 5)
	17	Describe methods of genetic engineering.	Agricultural Applications

FOUNDATIONAL CONCEPTS AND APPLICATIONS

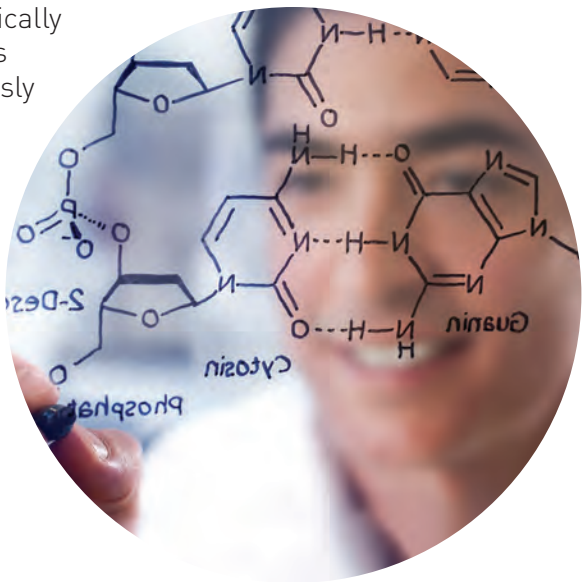


DNA Sequencing

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 something that is approaching $1,000.

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.



Second and third generation sequencing technologies should be briefly discussed in Biology courses as part of course of study (COS) standard 1 as the structure of DNA plays a role in modern sequencing techniques. DNA sequencing should be more thoroughly explored in COS standard 3 and related substandards (3a, 3b) particularly as it relates to how advancements in DNA technology have contributed to our understanding of the impacts of DNA change. This topic would also be appropriate for discussion in the Career/Tech Intro to Biotechnology course as part of objectives 1,13 and 15 and in AP Biology during investigation of Big Idea 3: Information.

The HudsonAlpha-developed, high school lab activity Genes & ConSEQUENCES®, connects information produced by a type of DNA sequencing system to genes, mutations and human 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 the "Bioinformatics" article for more details). The activity has been incorporated into the *Laying the Foundation* curriculum for A+ College Ready and is available through the AMSTI/ASIM high school program across the state. The kit is available for purchase through a partnership with Carolina Biological Supply. More information can be found at **www.hudsonalpha.org/available-educational-kits**.

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.

RNA and Protein Analyses

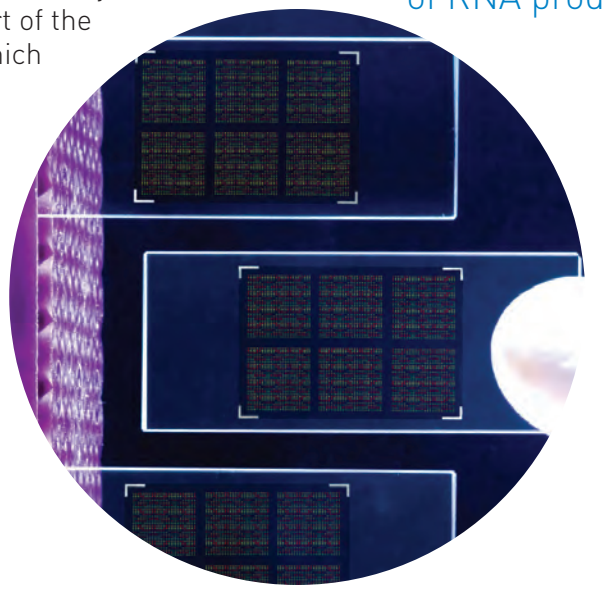
As sequencing techniques identify the genetic recipes of an organism, understanding the function of those genes becomes increasingly important. Many of the fundamental measurements used by molecular biologists seek to determine the presence, absence or relative amounts of RNA produced by a gene. Initially,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

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

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

RNA- and protein-based technologies should be noted in Biology relating to standards 1 and 3 as students strive to identify the relationship between structure and function of proteins and nucleic acids in cellular activities. These technologies can be examined in greater detail in an AP biology course as students investigate the occurrence and effects of genetic variability on populations, and methods used to regulate gene expression (Enduring Understanding 1.A and 3.B). These are also useful technologies to cover in the Career/Tech Intro to Biotechnology course, linking to COS objectives 9 and 14.

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

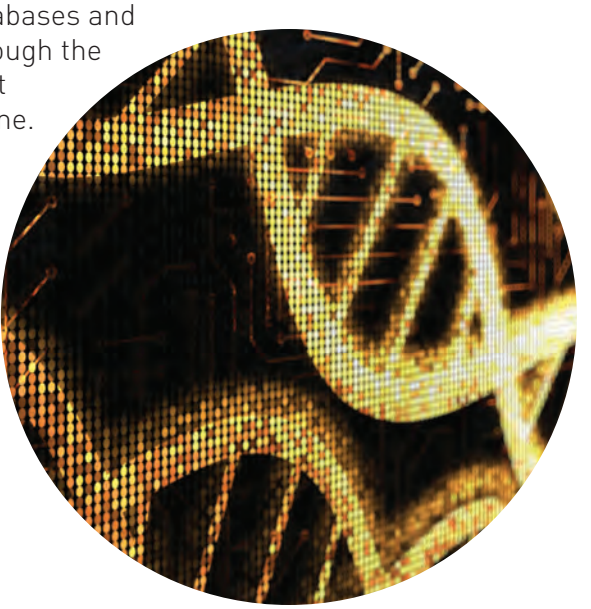


Bioinformatics

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

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

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



The concept of bioinformatics is a critical component to understanding modern genomic discoveries. It provides software 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 such ENCODE and the 1000 Genome Project (Biology Standard 3b). This topic could be incorporated in an AP Biology class under Enduring Understanding 1.A and to support the Science Practices of data analysis and mathematical modeling. Lastly, the creation, management and utilization of bioinformatics databases can be incorporated into the Technology Education course (COS objectives 26 and 27).

The HudsonAlpha-developed high school lab activity *Genes & ConSEQUENCES*® connects information produced by a DNA sequencing system to genes, mutations and human disease. The activity incorporates 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 student entered sequence data to known sequences from a number of organisms, including humans, to identify genetic matches. The dataset allows students to determine chromosomal location of key genes, investigate their role in disease, and compare sequence in healthy individuals to patients experiencing symptoms. The activity has been incorporated into the *Laying the Foundation* curriculum for A+ College Ready and is available through the AMSTI/ASIM high school program across the state. The kit is available for purchase through a partnership with Carolina Biological Supply. More information can be found at **www.hudsonalpha.org/available-educational-kits**.

APPLICATION

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 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% 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.

This application of genetic information and genetically modified organisms to increase agricultural yields, improve nutritional content, craft insect resistance or increase bioenergy yields provides topics of discussion for AP Biology as part of Big Ideas: (1) Evolution and (3) Information. In Biology, this topic can provide additional evidence as students develop an explanation of the impact of alterations of DNA to traits (COS standard 3). Environmental science classes may investigate the role of genome sequencing in selecting plants for sources of biofuels (COS standard 1). It also has a direct connection to Career/Tech courses in Agriscience (COS Objective 1 and 13) and Plant biotechnology (COS objectives 1, 14, 16, and 17).

Analyzing genomic data for agriculturally important plants is a rich field of study. HudsonAlpha has created Aluminum Tolerant Corn™ an activity designed to introduce students to genome analysis and copy number variation. Students analyze a section of the corn genome, using the bioinformatics website DNA Subway (developed by the DNA Learning Center). Students use NCBI's BLAST program to identify and explore the genes that are located within this region. Throughout the process, students link an aluminum-tolerant growth phenotype to structural variants that alter the copy number of key genes. This kit is currently undergoing pilot testing. For more information, contact **edoutreach@hudsonalpha.org**.

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 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.

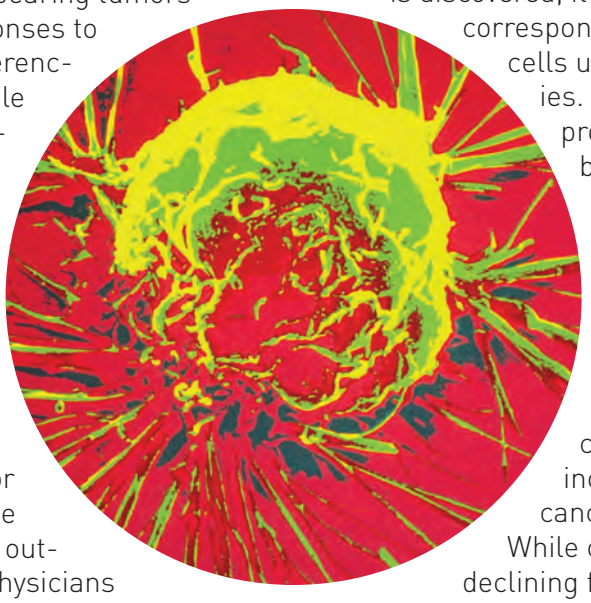
Researchers are also developing biofortified food plants to boost the levels of nutrients, vitamins and minerals in foods such as rice, cassava, carrots and tomatoes. It is hoped that these fortified foods will reduce the incidence of global hunger and micronutrient malnutrition (taking in adequate calories,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

Cancer

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

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

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



The idea that all cancers are genetic in nature and occur as a stepwise accumulation of additional of mutations, many of which are initiated by environmental factors, is a natural addition to investigation of the cell cycle control mechanisms (Biology COS standard 4, AP Biology Enduring Understanding 2.C). In Anatomy and Physiology classes, cancers of various body systems are examined (COS standards 3a and 9a). 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.

Hereditary Nonpolyposis Colorectal Cancer (HNPCC)® is a high school lab, developed at HudsonAlpha, which focuses on inherited cancer risk and detection. Students complete a family pedigree and interpret the pedigree to determine individual family member's risk for developing HNPCC. Students then complete a gel electrophoresis-based DNA analysis to diagnose family members with the HNPCC linked mutation. The lab introduces students to a genetic counselor and laboratory technician for career exploration. The HNPCC lab has been incorporated in the AMSTI *Science in Motion* program for high school life science classes across Alabama. The kit is available for purchase through a partnership with Carolina Biological Supply. More information can be found at **www.hudsonalpha.org/available-educational-kits**.

Microarray experiments are currently too cumbersome to perform in a clinic, so they are not likely to be used routinely to diagnosis patients.

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.

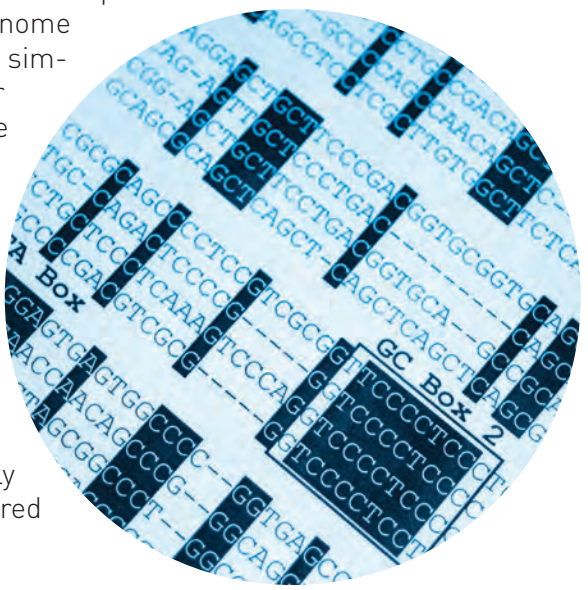
In the 2008 Annual Report to the Nation, the National Cancer Institute noted that both the incidence and death rate for all cancers combined is decreasing. While cancer death rates have been declining for several years, this marks the first decline in cancer incidence, the rate at which new cancers are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

Comparative Genomics

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

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

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



Comparative genomics provides evidence for the molecular process that underlies evolutionary theory and explains the nature and diversity of organisms, as outlined in Biology COS Standard 15. Biology Standard 16 specifically tasks students with analyzing DNA evidence to support the hypotheses of common ancestry. Comparative genomics and its relationship to evolution intersect AP Biology, in the Big Idea: Evolution and Enduring Understandings 1.A and 1B. Career/ Tech courses will also benefit from a discussion of comparative genomics in Veterinary Science (COS objective 3) and Intro to Biotechnology (COS objectives 9,11, and 14).

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

Copy Number Variation

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

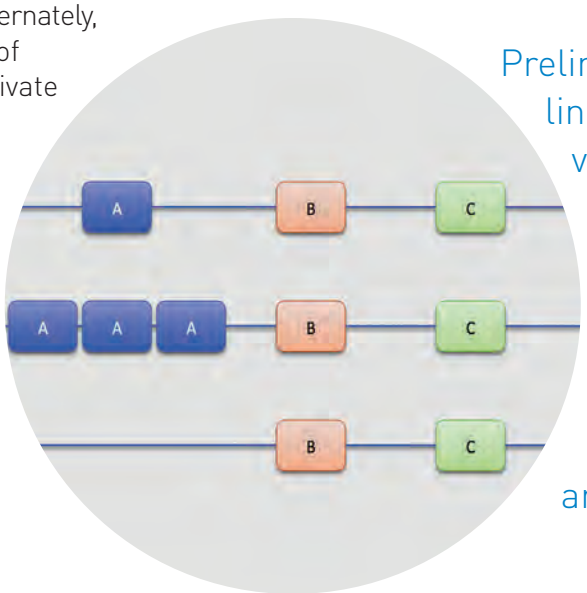
It came as a surprise to many scientists just how much DNA variation is due to copy number changes. Previous studies based primarily on SNPs suggested that any two randomly selected human genomes would differ by 0.1%. CNVs revise that estimate: the two genomes differ by at least 1.0%. While this may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

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

Relating genetic variation to human disease and inheritance is identified in the Biology COS in standard 3 and 12a. Genetic variation is also highlighted under standard 3b, which explores the ongoing impacts from the Human Genome Project and subsequent large-scale research projects. The impact of copy number variation intersects AP Biology in Enduring Understanding 3.B with discussions of gene regulation. Career/Tech Intro to Biotechnology should include discussion of copy number variation (COS objective 8).

Scientists are just beginning to understand the impact of structural genomic variation on plant, animal and human phenotypes. HudsonAlpha has created *Aluminum Tolerant Corn*™ an activity designed to introduce students to genome analysis and copy number variation. Students analyze a section of the corn genome, using the bioinformatics website DNA Subway (developed by the DNA Learning Center). Students use NCBI's BLAST program to identify and explore the genes that are located within this region. Throughout the process, students link an aluminum-tolerant growth phenotype to structural variants that alter the copy number of key genes. This kit is currently undergoing pilot testing. For more information, contact **edoutreach@hudsonalpha.org**.

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



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 disease.

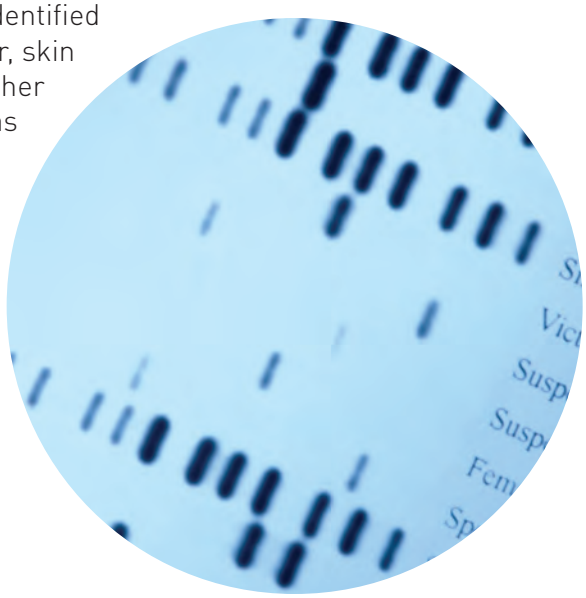
Criminal Justice and Forensics

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 profiling is a critical component of the Career/Tech course Forensic and Criminal Investigation (COS objectives 7 and 8) and Intro to Biotech (COS objectives 1,13 and 14). It can also be explored in AP Biology as part of the Big Idea 3: Information. 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 be incorporated into these discussions.

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.



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 (Standards 4, 11, 12 and related sub-standards). Karyotyping to diagnose chromosomal disorders is examined 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 may be outside the scope of the high school classroom. 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 team has crafted Disorder Detectives®, a kit that allows students to take on the role of a cytogeneticist working in a hospital or clinic. Students are given a case study and set of human chromosome clings and must arrange the chromosomes into a completed karyotype, analyze the karyotype and diagnose their patient. Many types of typical and atypical karyotypes are presented. Students explore technologies such as FISH, aCGH and sequencing to learn how laboratories can diagnose even the smallest genetic structural changes. Geneticists, genetic counselors and laboratory technicians are highlighted as careers that utilize these types of technologies. The activity is available from AMSTI/ASIM and can be purchased from Carolina Biological Supply. More information can be found at **www.hudsonalpha.org/available-educational-kits**.

ChromoSock® kits, developed at HudsonAlpha, use custom-made socks as models for chromosomes to examine the movement of chromosomes during cell division. Providing a physical manipulative allows students to investigate the impacts of errors in cell division. ChromoSock Meiosis™ and Modeling Mendel's Laws® are available through AMSTI/ASIM and for purchase through Carolina Biological Supply. More information can be found at **www.hudsonalpha.org/available-educational-kits**.

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 twins' 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 examining exceptions to standard Mendelian inheritance, for Biology COS Standard 11 (and it's sub-standards) as well as Intro to Biotech COS objective 9. The relationship between methyl modifications on the DNA and the silencing of genes links epigenetics to AP Biology Enduring Understanding: 3B, during discussions of gene 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 (GINA)

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 pre-disposed. 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.

Genetic discrimination has historical, legal and social implications and is often discussed in life science classrooms. In biology classes, this discussion often coincides with discussions of the impacts of the Human Genome Project and subsequent large-scale genomic research initiatives (COS standard 3b). Discussion of current legislation related to genetic discrimination should be included Career/ Tech courses Foundations of Health Science (COS objective 10), Health Informatics (COS Objective 5) and Intro to Biotechnology (COS objective 14). Exploration of genetic discrimination occurs in AP Biology courses as part of Enduring Understanding 3.A.

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.

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.

The multifactorial genetics of eye color should be discussed in Biology courses as part of COS standard 11c, especially since most text books still explain this trait in terms of a single gene effect. It could also be explored in AP Biology courses Enduring Understanding 3B alongside discussion of cellular and molecular processes involved in translating genetic information into phenotype. In the Career/Tech Intro to Biotechnology courses, eye color genetics could be explored under COS objectives 8 and 11.

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% 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.

Identifying Genetic Influence on Disease

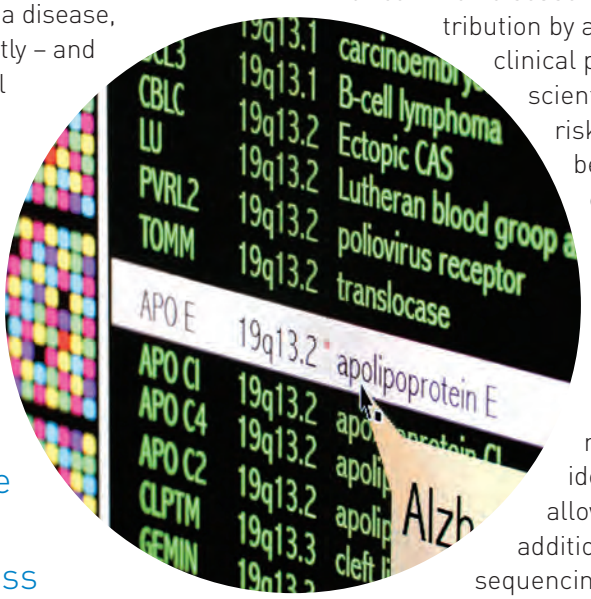
Much progress has been made in identifying the genetic causes of single gene diseases such as cystic fibrosis, phenylketonuria and Huntington disease. This has led to more accurate risk analysis, better testing approaches and, in some instances, more effective methods of treatment. Even though there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

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

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

The premise behind GWA studies: 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. Basically, there is an association between the specific allele and the incidence of disease.

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% of the population.



Relating genetic variation to human disease and inheritance is identified in Biology COS standard 11c during investigation of disease risks from both genetic and environmental factors. This would also be appropriate discussion in AP Biology (Enduring Understandings 2e and d, 3a, b, c and 4a), Health (COS objectives 5 and 10) and the Career/Tech Intro to Biotechnology course (COS objective 14).

Touching Triton® is a serious game that has been developed by HudsonAlpha through a National Institutes of Health Science Education Partnership Award, with additional support from Lockheed Martin. This free web-based game challenges students to analyze and interpret data related to the risks for developing common complex disease. Through the storyline of long-term space flight, students learn about the complexity of risk for common disease such as diabetes, colon cancer and Parkinson disease. Students analyze data from crew members' medical record, family history and genomic report to make medical packing decisions for a 20-year space mission. More information can be found at **www.triton.hudsonalpha.org**.

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

Infectious Disease

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

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

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

Discussions of pathogens occur in biology classes as part of standard 13a (viral classification), and as examples of rapidly evolving organisms (antibiotic resistance) and during discussions about co-evolution among host and pathogen as part of standard 14. In Career/Tech Intro to Biotechnology courses, infectious disease could be explored under COS objectives 11,13,14 and 15.

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 tests, treatments and preventatives. Although the genomes of pathogens have the capability to change rapidly, the genomes are small and often change in semi-predictable ways. Scientists may never be able to cure the flu or common cold, but through genetics and biotechnology, more accurate and faster diagnostics can be made.

Non-Invasive Prenatal Testing

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 pg. 36) – or more narrow in scope, such as testing a single gene. Both amniocentesis and CVS carry a small but significant risk of miscarriage.

Scientists have recently developed a novel, non-invasive testing method. In the 1990s, it was discovered that fetal DNA crosses the placenta into the maternal bloodstream. Today, relatively straightforward techniques can 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 testing. In 2012, three companies introduced this form of non-invasive prenatal testing into the clinic.

Non-invasive prenatal testing is currently classified as a screening, rather than a diagnostic test. It signals whether further, often more invasive forms of testing should be considered.

Prenatal diagnosis is a standard part of discussion around egg and sperm formation and abnormalities that can occur during meiosis. The advent of non-invasive techniques is an exciting addition for Biology (COS standards 12, 12a, 11, 3c). The application of this new technology to health and society links to classroom conversation in AP Biology (Big Idea: Information) 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.

Whether this will ultimately replace CVS and amniocentesis as a diagnostic test will depend upon improvements in 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?

Personal Genome Analysis

As sequencing costs drop, it has become feasible to analyze large portions of a human genome relatively quickly and comparatively inexpensively. This has most often been performed in either a research setting to better understand the functional impact of genetic variation or in the clinic to identify the molecular cause of a suspected genetic disease. However, there is a growing market for providing genomic information to what are sometimes termed "ostensibly healthy participants" – individuals without visible disease or health complications but who want to know their genomic information and understand how it informs their ancestry, personal traits and potential future risks for developing certain diseases.

An initial step towards personal genome analysis has been direct to consumer genotyping – a targeted analysis of between 500,000 and 1,000,000 variable regions from across the genome. A small but increasing proportion of these variants is connected to ancestry, physical traits or disease risk, although the predictive value of medical decisions of these risks is often unclear. The FDA ordered the health-related versions of these tests halted in 2013, although it has recently allowed a limited number of direct to consumer genetic tests back onto the market. Consumer genotyping is also available for individual genes such as the ACTN3 genetic variant involved in muscle strength and spring ability. These genetic differences are poor predictors of athletic skill as well as musical or artistic talent and overall intelligence, as most of the genetic and environmental influences on these traits are still unknown.

Today, predispositional (or presymptomatic) genomic screening – PPGS – analyzing the exome or entire genome of an ostensibly healthy individual – is controversial. There is little data about the response of people who have received genomic information about their trait and disease risk factors. At the same time, there is a powerful and growing recognition among personal genomic stakeholders that such information may provide a positive benefit on an individual's life and actions, even if the direct health benefit is uncertain or marginal. A number of research projects have been initiated to inform our understanding of these impacts, collectively involving more than 1,000 individuals. Common motivations for participating in PPGS initiatives include the desire to learn health-related information, a sense of general curiosity about personal genomic information and the desire to contribute to research that may benefit others. In keeping with the early adopter status of these studies, current participants tend to be highly educated, technically savvy and from a high socio-economic status. There have been few published studies of the impact of PPGS on the participants and the short and long-term benefits and concerns are primarily speculative. A long-term analysis of this sort of information is being conducted by the PeopleSeq Consortium, a collaboration between multiple PPGS projects using a common set of questions and techniques.

Personal genome studies offered direct-to-consumer should be a component of students' efforts to obtain, evaluate and communicate information about advancements in genetic technology. This provides fertile ground for a discussion of the implications of genetic information. (Biology COS standard 3) These topics can also be incorporated into an AP Biology course in Big Idea 3: Information. Personal Genome Analysis provides modern content and context for students in Health (COS objective 6) and the Career/Tech course Foundation 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.

Determining the significance of a particular variant identified through genome sequencing is still in its infancy. Classifying a variant as benign or pathogenic requires multiple lines of supporting evidence, curated by a team of researchers and clinicians. For many variants, such supporting evidence either doesn't exist, or is contradictory, resulting in a classification of "variant of uncertain significance (VUS)." Currently in pilot testing, HudsonAlpha has developed a card-based kit that brings the process of variant analysis to high school classrooms. Making Sense of Uncertainty™ provides opportunities for students to classify variants and argue from evidence to justify their classification. A pilot version of this kit is available to Alabama high school life science educators who attend the two-day GREAT workshops scheduled across the state during the 2016-17 school year. More information can be found at **www.hudsonalpha.org/GREAT or p.56 of this Guidebook**.

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.

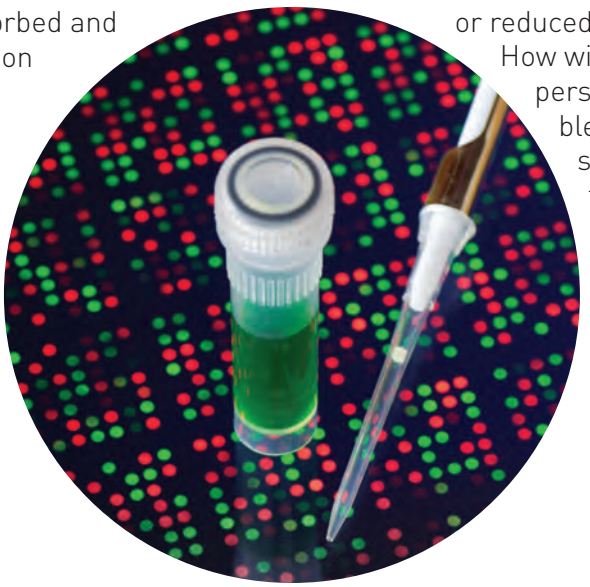
Drawing an analogy from the world of fashion, personalized medicine is the equivalent of a custom-made suit or outfit, designed with an individual's unique body measurements. This type of tailored approach provides a much better fit than purchasing something off the rack.

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 pg. 47).

This newfound knowledge is rapidly moving into the clinical setting. At the forefront are a series of drugs such as Gleevec®, 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. 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 genetic profile.

The implication of personalized medicine impacts biology-based science courses, Health Education and pre-healthcare options at the high school level. Biology COS 3b addresses modern understandings of the central dogma and the research projects that have enhanced those understandings. Biology COS standard 11c and 12a, as well as AP biology Big Idea: Information, involve the impacts of genetic variation on human disease. In a Health course, COS objective 5 asks students to evaluate negative and positive impacts of technology health. Personalized medicine is an excellent candidate for this discussion, as well as showing application to the Career/Tech courses Introduction to Pharmacy (COS objective 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 have moved 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. More than 200 pharmaceutical products 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 l evels, leading to increased cell growth and tumor formation. In these cases, the anti-cancer drug Herceptin® is added to the patient's treatment plan where it increases the efficacy of chemotherapy.

Molecular testing is needed because only 25% of breast cancer patients will see any benefit from Herceptin® – the rest should be given another treatment. In a similar manner, 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 3b addresses modern understandings of the central dogma and the research projects that have enhanced those understandings. Biology COS standard 11c and 12a, as well as AP Biology Big Idea: Information, involve the impacts of genetic variation on human disease. In a Health course, COS objective 5 asks students to evaluate negative and positive impacts of technology health. Pharmacogenomics, as part of personalized medicine, is an excellent candidate for this discussion, as well as showing application to the Career/Tech courses Introduction to Pharmacy (COS objective 9 and 11) and Intro to Biotechnology (COS objectives 11 and 14).

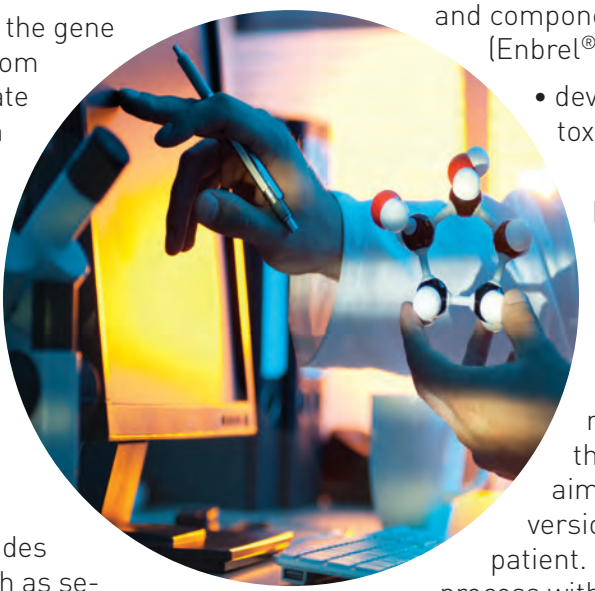
More than 200 pharmaceutical products either recommend genetic testing or point to the influence of genetic variability on the drug's response.

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.



Investigating recombinant DNA offer a way to re-emphasize central dogma (the information in DNA is transcribed into RNA and then translated into protein) which is a central feature of Biology COS standard 3 and AP Biology Enduring Understanding 3.A, and clearly links biological concepts to real-world application. 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 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 or soybean, adding pest or herbicide resistance, or increasing nutrient content (see Agricultural Applications, pg. 33)
- 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 pg. 50.

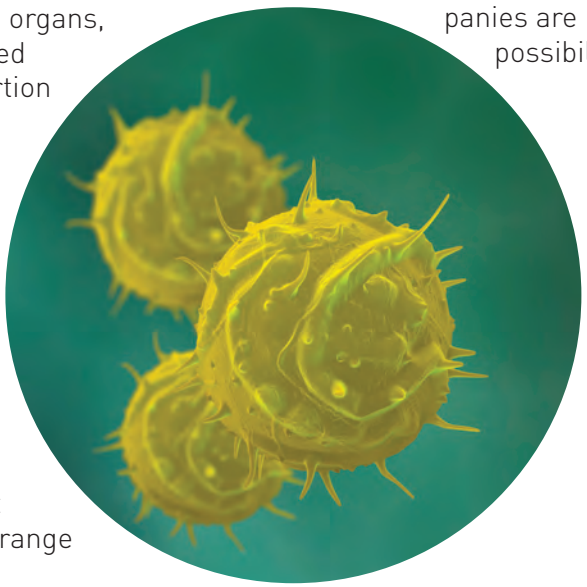
Stem Cells

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 a lung or brain cell.

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 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 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 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.



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.

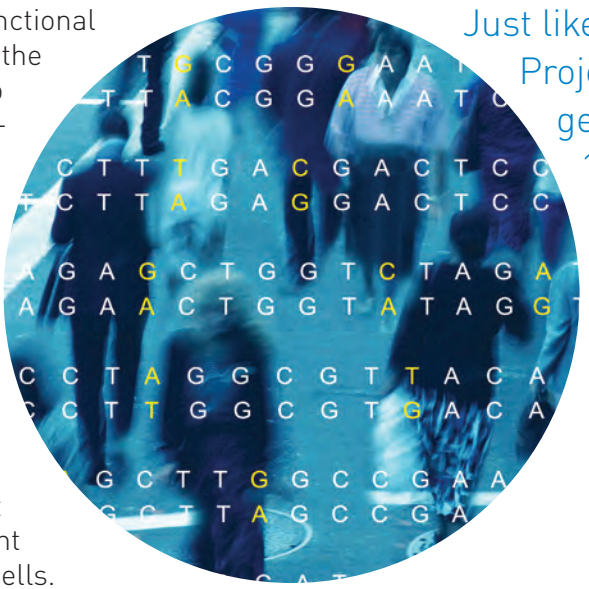
The concept of stem cells connects to several components of the standard Biology Course. It can be highlighted during modeling of the cell cycle (COS standard 4). In addition, exploring the similarities and differences between stem cells and differentiated cells would reinforce concepts about structure and function of cells and how specific functions are performed (COS standard 2). The role of biotechnology in the development of induced Pluripotent Stem cells connects to Introduction to Biotechnology (COS objective 5) and AP Biology (Enduring Understanding 3.B).

Studying the Genome to Understand the Sequence

In 2001 the completion of the Human Genome Project and the publication of the DNA sequence found inside every human cell were announced with much fanfare. Although it may have seemed like the end of an era, in reality it was only the beginning. Little was known about how cells used DNA information to function and interact. There was not a clear understanding of how genes maintain human health or predispose to disease. A representative genome had been sequenced, but how many differences in the genetic information were present from person to person? How did the sequence compare to other organisms? Sequencing the human genome raised far more questions than it answered.

Since that time, several large-scale projects have expanded our understanding of the human genome. The International HapMap Project identified common genetic variants and compared them across world populations. This was followed by the 1,000 Genomes Project, which sought to categorize rare genetic changes across an even larger number of global communities. Collectively, these two projects identified more than 88 million DNA changes.

ENCODE, the Encyclopedia of DNA Elements, was launched to determine the functional significance of every nucleotide in the genome. This project is working to detect and classify those sequences that stimulate or silence the transcriptional activity of all genes. Data published in 2012 suggest that as much as 80% of the genome is involved in some sort of "biochemical function." This includes those regions of the DNA that are bound by proteins necessary to regulate transcription or DNA folding, but also includes DNA sequences that correspond to evolutionarily ancient mechanisms not used by human cells. Additional analyses suggest that 8-20% of the genome is functionally important for human life.



Just like the Human Genome Project, information generated by HapMap, 1,000 Genomes and ENCODE are freely accessible to scientists and the public around the world.

The history of and findings of the Human Genome Project, ENCODE, 1000 Genome Project and other large scale genomic research projects that have shaped our understanding of how DNA influences traits are the specific focus of Biology COS Standard 3b. These projects have shed light on structure of eukaryotic chromosomes, the influence of genetic change on human diversity and the functionality of noncoding regions of DNA. The annual Biotechnology Guidebook provides a resource for information about these projects that students evaluate to meet the standard. These topics also have merit for discussion in the Career/Tech Veterinary Science (COS objective 3) and Intro to Biotechnology (COS 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 Hap Map, ENCODE or the 1,000 Genomes Project. 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.

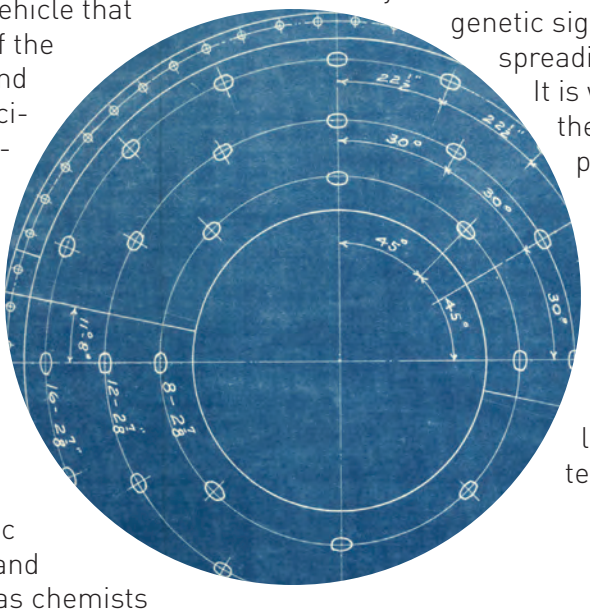
Following such a foundational activity, The Progress of Science™ timeline (**timeline.hudsonalpha.org**) is a resource that provides information about key discoveries in the development of genomic sciences.

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.

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.

The concepts behind synthetic biology link to Biology COS standard 3, particularly as it relates to understanding of the elements of DNA code necessary for biological function. Discussion of synthetic biology also connects to AP Biology (Enduring Understanding 3.A and 3.E) and the CTE Intro to Biotechnology (COS objective 13). This is a natural connection to synthetic biology, which uses recombinant DNA techniques as the cornerstone to creating artificial bioparts, systems and devices.

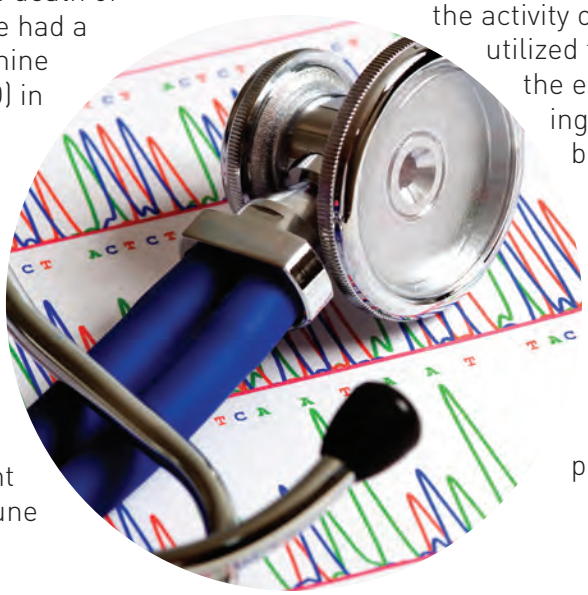
THERAPEUTIC APPROACHES

Gene Therapy

Gene therapy is defined as the correction of a non-functioning 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 or silencing protein synthesis could be included in a Biology course (COS standard 3) during the gathering of evidence to explain how DNA determines the structure of proteins and the relationships between DNA, RNA and protein. In an AP Biology course, these approaches provide additional insight into how living systems store and transmit essential information (Enduring Understanding 3.A and 3.B). The role of these approaches in treating human disease would be discussed in the Career/Tech Intro to Biology 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

Code of Alabama

Section 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.

REFERENCES AND IMAGE CREDITS

Science Snapshot (pg. 9)

DNA sequencing in space. Gebhardt, Chris. "NASA Prepares for First-ever In-space DNA Sequencing Experiment." NASASpaceFlight.com. N.pg., 22 July 2016. Web. 23 July 2016.

Migraine Loci identified. Gormley Pg. et al. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. Nature Genetics (2016) published online June 20, 2016 doi:10.1038/ng.3598.

Minimal Genome. Hutchison C.A. 3rd et al. Design and synthesis of a minimal bacterial genome. Science (2016) 351:1414- doi:10.1126/science.aad6253.

GDF6 and human foot shape. Indjeian V. et al. Evolving New Skeletal Traits by cis-Regulatory Changes in Bone Morphogenetic Proteins. Cell (2016) 164:45-56 doi:10.1016/j.cell.2015.12.007. www.hudsonalpha.org/researchers

26 hour record for WGS clinical analysis. Miller N. et al. A 26-hour system of highly sensitive whole genome sequencing for emergency management of genetic diseases. Genome Medicine (2015) 7:100 doi:10.1186/s13073-015-0221-8. We can now sequence a whole human genome in 26 hours. (September 30, 2015). Retrieved July 28, 2016 from http://www.popsci.com/scientists-can-now-sequence-whole-genome-in-26-hours

GM mosquitoes reduce dengue incidence. "Dengue Fever Cases Drop 91% in Neighbourhood of Piracicaba, Brazil, Where Oxitec's Friendly™ Aedes Were Released | Oxitec." Oxitec. N.pg., 14 July 2016. Web. 23 July 2016.

California sugar pine sequenced. Filmer, A. (2015, December 16). Genome sequencing may save California's legendary sugar pine. Retrieved July 24, 2016, from https://www.ucdavis.edu/news/genome-sequencing-may-save-california's-legendary-sugar-pine.

The BabySeq Project. (2015). Retrieved July 24, 2016, from http://www.genomes2people.org/babyseqproject/

Dun horse pigmentation. Imsland F. et al. Regulatory mutations in TBX3 disrupt asymmetric hair pigmentation that underlies Dun camouflage color in horses. Nature Genetics (2016) 486:152-8 doi:10.1038ng.3475. www.hudsonalpha.org/researchers

Hair related genes. Adhikari K. et al. A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features. Nature Communications (2016) 7:10815 doi:10.1038/ncomms10815.

page 9 – Sugar pine – By willmcw at en.wikipedia [Public Domain] via Wikimedia Commons https://commons.wikimedia.org/wiki/File:Sugar_Pine.jpg

page 10 – Mosquito – Centers for Disease Control – PHIL images #9258, James Gathany hppt://phil.cdc.gov/phil/home.asp

page 10 – Peppered moths (white and black) – By Olaf Leillinger via Wikimedia Commons https://commons.wikimedia.org/wiki/File:Biston.betularia.7200.jpg; https://commons.wikimedia.org/wiki/File:Biston.betularia.f.carbonaria.7209.jpg

page 11 – Photo: *C. Elegans* sperm - WormAtlas, Sam Ward, Paul Muhlrad, Craig LaMunyan, www.wormatlas.org

page 11 – Monocercomonoides – Tree of Life Media, Kevin Carpenter and Patrick Kelling http://tolweb.org/onlinecontributors/app;jsessionid=AD718C36220D1754EC-41142C568E1506?page=ViewImageData&service=external&sp=29835

page 11 – Zika virus – RCSB Molecule of the Month, Drawn by David Goodsell via Wikimedia Commons https://commons.wikimedia.org/wiki/File:197-Zika_Virus-Zika Virus.tif

page 12 – Population Architecture Using Genomics and Epidemiology - NHGRI Digital Media Database, Jonathan Bailey http://www.genome.gov

page 12 – Pediatric Genome Sequencing - NHGRI Digital Media Database, Ernesto del Aguila III http://www.genome.gov

page 13 – Retina Therapy - NHGRI Digital Media Database, Ernesto del Aguila III http://www.genome.gov

page 14 – Moon – By Nigel Howe via Wikimedia commons https://commons.wikimedia.org/wiki/File:Moon_24_03_2011_(5555607876).jpg

page 14 – The Cancer Genome Atlas (TCGA) - NHGRI Digital Media Database, Ernesto del Aguila III http://www.genome.gov

page 14 – Treatment-resistant breast cancer cells – National Cancer Institute\Dana-Farber Harvard Cancer Center at Massachusetts General Hospital, Sheheryar Kabraji, Sridhar Ramaswamy https://visuals.nci.nih.gov/details.cfm?imageid=10574

page 15 – African Elephant – Photographed by Bernard Dupont via Wikimedia Commons https://commons.wikimedia.org/wiki/File:African_Elephant_(Loxodonta_africana)_male_(16723147361).jpg

page 16 - CRISPR/Cas space filing model from Escherichia coli – RCSB Molecule of the Month, Drawn by David Goodsell via Wikimedia Commons https://commons.wikimedia.org/wiki/File:181-CascadeAndCRISPR_1vy8.tiff

page 17 – Centers of Excellence in Ethical, Legal and Social Implications Research (CEER) - NHGRI Digital Media Database, Darryl Leja http://www.genome.gov

page 17 – Teal and Pink Awareness Ribbons – By Messer Woland via Wikimedia Commons https://commons.wikimedia.org/wiki/File:Pink_ribbon.svg; https://commons.wikimedia.org/wiki/File:Teal_ribbon.svg

page 18 – Hereford Cow - NHGRI Digital Media Database, Michael MacNeil, USDA http://www.genome.gov

NOTES

Genetic Resources to Empower Alabama Teachers

GREAT Standards

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 third round of GREAT workshops will dig into **the newly adopted Alabama Science Course of Study** by working through the following:

- Using models to help students predict cell membrane behavior
- Using genetic variant analysis to argue from evidence
- Addressing large scale genetics projects such as ENCODE and the 1000 Genomes Project
- Analyzing data from environmental and genetic factors to inform lifetime risk for complex disease
- Evaluating and modifying classroom resources in light of the new course of study

In two full days of small group concurrent sessions and talks by dynamic speakers, teachers will learn about recent findings in genetics and genomics and methodologies to address this content with their students.

Teachers who complete both days of the workshop will return to the classroom with lesson plans and hands-on materials that are student-tested and informative and that link to state course of study objectives.

For more information, visit **www.hudsonalpha.org/GREAT**

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.

HudsonAlpha
INSTITUTE FOR BIOTECHNOLOGY

601 Genome Way Huntsville, AL 35806 | 256.327.0462 | www.hudsonalpha.org