

An abstract graphic featuring a thick green helical ribbon that winds around a large black sphere. The ribbon continues to wind around a smaller black sphere below it, and then forms a series of smaller, stacked black spheres at the bottom right. A smaller green helix is visible in the bottom left corner.

# Biotechnology Discoveries and Applications

Extensions to high school science curriculum

The 2017-2018 guidebook



## 2017-2018 GUIDEBOOK CONTENTS

<b>About HudsonAlpha Institute for Biotechnology</b>	2
<b>Educational Opportunities at HudsonAlpha</b>	4
<b>Executive Summary</b>	8
<b>Science Snapshots</b>	9
<b>New Findings</b>	10
<b>Types of Genetic Testing</b>	20
<b>The Course of Study Connected to Guidebook Topics</b>	22
<b>Foundational Concepts and Applications</b>	27
<b>Key Technologies</b> (DNA Sequencing, RNA and Protein Analyses, Bioinformatics)	28
<b>Applications</b>	31
1. Agriculture	31
2. Cancer	32
3. Comparative Genomics	33
4. Copy Number Variation	34
5. Criminal Justice and Forensics	35
6. Diagnosing Chromosome Disorders	36
7. Epigenetics	37
8. Genetic Information Nondiscrimination Act (GINA)	38
9. Genetics of Eye Color	39
10. Genome Editing	40
11. Identifying Genetic Influence on Disease	41
12. Infectious Disease	42
13. Noninvasive Prenatal Testing	43
14. Personal Genome Analysis	44
15. Personalized Medicine	45
16. Pharmacogenomics	46
17. Recombinant DNA and Genetic Engineering	47
18. Stem Cells	48
19. Studying the Genome to Understand the Sequence	49
20. Synthetic Biology	50
21. Therapeutic Approaches (Gene Therapy, RNAi)	51
<b>Statutes and Session Law</b>	52



## 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.

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

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

## Genomic Research

Current research focus areas are:

### Basic Research

Foundational research aimed at improving scientific theories and understanding

### Pediatrics

Undiagnosed childhood genetic disorders

### Neurological and Psychiatric Disorders

including Alzheimer disease, Parkinson disease, ALS, Huntington disease, bipolar disorder, schizophrenia, autism and epilepsy

### Cancer

Multiple forms of cancer, including breast, ovarian, prostate, kidney, brain, colon and pancreatic

### Genomic Medicine

Leveraging the power of the human genome to diagnose, predict and prevent disease

### Agriscience

Applying genomic knowledge to agriculture and bioenergy to create a more sustainable world

### Immunogenomics

Application of genomic technology to understand the immune system's role in health and disease

### Computational Biology and Bioinformatics

Deep computational analysis and interpretation of vast amounts of data, critical to the science of genomics

## Biotech Enterprises



HudsonAlpha strengthens and diversifies Alabama's economy by fostering success in life sciences companies of all stages and sizes. Its 152-acre biotech campus within Cummings Research Park supports more than 30 tenant companies, from startups to global leaders, with space for more. HudsonAlpha offers turnkey and build-to-suit laboratory and office space for lease in an energizing environment with superior shared amenities. Bioscience enterprises on campus benefit from access to HudsonAlpha researchers as well as strategic support through investor forums, workforce and business assistance, marketing resources, and bioscience networking events.







## HudsonAlpha's Educational Programs

HudsonAlpha's Educational Outreach team inspires the next generation of life sciences researchers and workforce while building a more biotech-literate society. The dynamic educators at HudsonAlpha reach students, educators, medical providers, patients and the community through hands-on classroom modules, in-depth school and workshop experiences, and digital learning opportunities. HudsonAlpha also provides educational opportunities for healthcare providers and learning tools for patients who are making medical decisions using their personal genomic information. Additionally, the team builds genomics awareness through community outreach classes and events. More than 1.5 million individuals were impacted through HudsonAlpha Educational Outreach during the 2016-2017 academic year.



## Teacher Professional Development

In addition to this guidebook, HudsonAlpha has several opportunities for teacher professional development, ranging from single-day workshops to ongoing classroom support. These increase an educator's comfort in discussing genetic concepts and terminology and the associated ethical, social and legal issues.



## 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.

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.



See pg 6  
for more  
details about  
HudsonAlpha  
Digital  
Resources.

## Digital Resources

HudsonAlpha has crafted a suite of free digital activities available to students, educators and anyone who uses the Internet on a computer or mobile device. iCell® is an interactive simulation that allows users to explore and understand the inner workings of a typical animal, plant or bacterial cell with 3-D animation. Touching Triton is a web-based activity that builds understanding of common complex disease risk, influenced by factors from family history, environment and genomic data. The Progress of Science Timeline showcases the history of genetics and biotechnology discoveries, and GenomeCache® explores the information contained in the human genome. Finally, Genome Gateway® provides educational resources for patients and physicians as they apply genomic information to the practice of medicine.

# EDUCATION PROGRAMS

The HudsonAlpha Educational Outreach team inspires and prepares society to embrace and use genomic information. The team nurtures the workforce for tomorrow's life science laboratories and companies.

HudsonAlpha provides educational programs for students, teachers, health professionals and the community at large. This graphic highlights the various outreach opportunities.

**BioTrain®:** summer student internship  
**Genetic Counseling** mini rotation  
**PORTALS:** biotech skills course  
**CODE:** DNA variant modeling  
**Nursing student training**  
**Digital applications\***

**Code of Life, I Want to Work in a Lab Coat,**  
**UDSO Challenge:** summer camp series  
**Digital applications\***

Lifelong Learners

Medical Professionals

High School Educators

Middle School Educators

Undergraduate and Graduate Students

High School Students

Middle School Students


Elementary Students




**Biotech 101 & 201: public seminar series**  
**DNA Day GeneChat: social media event**  
**Digital applications\***




**Genomics & Your Practice: healthcare provider continuing education**  
**Genome Gateway®: patient communications and education**



**Alabama Science in Motion: classroom kits**  
**Annual Guidebook: recent biotechnology discoveries and applications**  
**Compendium: a Biology teacher's fieldguide**  
**GREAT: genetic content updates for Alabama educators**  
**GTAC®: summer educator professional learning academies**  
**Bicentennial Barcoding: cataloging Alabama's native plants**  
**Field trips to biotech campus**



**Alabama Math, Science and Technology Initiative:**  
**genetics and biotechnology classroom module**  
**Middle School GPS: genetics content updates for Alabama educators**  
**Fieldtrips to biotech campus**



**APPLE: scientists on site with AP Biology classes**  
**Biotech Academy: summer intensive experience**  
**Summer Short Course: themed science camp**  
**LABS: after-school lab skills program**  
**Digital applications \***



**Community events and activities**

**\* Digital applications** *(see pg. 6 for application details)*

Visit **[www.hudsonalpha.org/education](http://www.hudsonalpha.org/education)** to find out more about the programs offered at HudsonAlpha.

**\* Digital applications: iCell®, GenomeCache®, The Progress of Science Timeline, Touching Triton®, the HudsonAlpha Vimeo® channel: [vimeo.com/hudsonalpha](http://vimeo.com/hudsonalpha)**

# 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 3 million times by students and educators around the world.

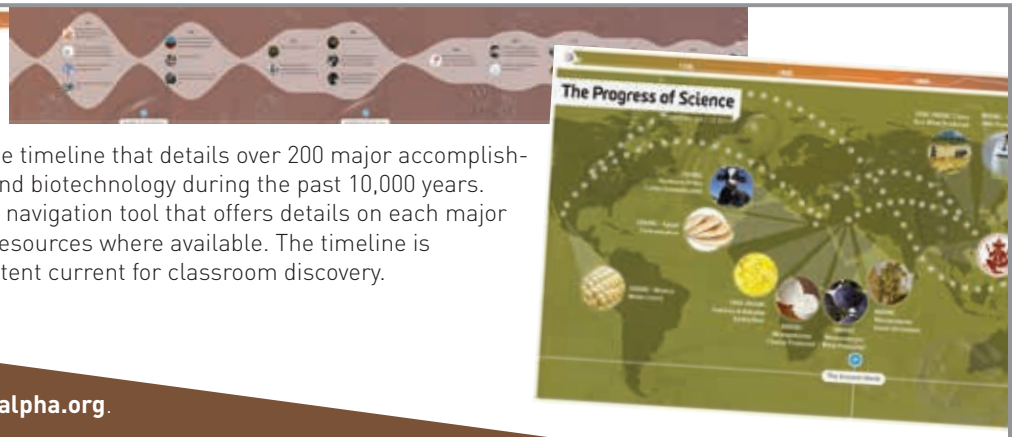


**iCell is available** on Apple® and Android® devices, Windows 8+ tablets, as a downloadable program for Mac® and Windows®, and at [icell.hudsonalpha.org](http://icell.hudsonalpha.org).

## 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](http://timeline.hudsonalpha.org).



**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.

## FREE Digital Activity



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

Made possible by: Grant Number BR25 OD010981-02

LOCKHEED MARTIN

SEPA SCIENCE EDUCATION PARTNERSHIP AWARD  
Supported by National Institutes of Health

HUDSONALPHA  
INSTITUTE FOR BIOTECHNOLOGY

[triton.hudsonalpha.org](http://triton.hudsonalpha.org)



# genome cache®

Build your own genome, or walk ours. **Genome-Cache** 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. Genome-Cache 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](http://genomecache.hudsonalpha.org).



# Biotechnology Discoveries and Applications 2017-2018

---

## HOW THIS GUIDE IS ARRANGED

Recent research findings are grouped on pages 9 through 19 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

Ten years ago, HudsonAlpha printed the first “Biotechnology Discoveries and Applications” The 36 page pamphlet featured a primer about DNA and a summary of 17 applications of genetics and biotechnology. Since then, the guidebook has grown significantly – this year’s volume clocks in at 56 pages, summarizes 24 foundational concepts and highlights 40 research discoveries published from August 2016 to July 2017.

Over the last decade, the field of genetics has undergone a head-spinning transformation. Sequencing an organism’s entire genome has become almost commonplace, with scientists routinely analyzing thousands of genomes for a single study. Our understanding of the functional significance of genes and their regulatory sequences continues to expand. The field of genome editing has burst upon the scene, potentially reshaping human health, agriculture and our battle with infectious disease.

Throughout it all, our goal for this guidebook has remained constant: **to showcase recent advances in genetics and biotechnology so educators can easily share the findings with their students.**

The articles highlighted in this year’s guidebook include:

- preliminary genetic findings from the NASA Twin Study (pg 9)
- a new form of imaging that upends traditional views about leading and lagging strand DNA replication (pg 10)
- the potential impact of gut microbiomes on the “yo-yo effect” for dieting (pg 13)
- a cautionary tale about genetic testing (pg 14)
- the surprising set of genes that are transcriptionally activated after death (pg 16)
- good news for those who bemoan the bland taste of grocery store tomatoes (pg 19)



**Neil Lamb, PhD**

*Vice President for Educational Outreach  
HudsonAlpha Institute for Biotechnology*

Each finding links to one or more of the foundational stories, covered in detail beginning on page 27. As in the past, the foundational topics are tied to Alabama course of study objectives for health, science and relevant career technical education classes (pages 22–26). Educators from other states will find that these foundational topics align with their own state’s objectives in a similar fashion.

A special word of thanks to Madelene Loftin, Vanessa Wamsley and Cathleen Shaw. These three amazing individuals were an integral part of article selection/development and the overall layout/design. I am grateful to count them as HudsonAlpha colleagues.

Happy Reading!

A handwritten signature in dark ink, which appears to read "Neil E. Lamb". The signature is stylized with a large, looping "N" and a cursive "Lamb".

**Neil E. Lamb, Ph.D.**

Vice President for  
Educational Outreach

*HudsonAlpha Institute  
for Biotechnology*

[nlamb@hudsonalpha.org](mailto:nlamb@hudsonalpha.org)

## SCIENCE SNAPSHOTS

### a quick rundown of 10 genetics and biotech stories

**1.** Astronauts Mark and Scott Kelly were the subjects of the “NASA Twin Study” – Scott recently spent 340 days at the International Space Station, while his identical twin sibling remained on Earth. A number of biological tests were performed on both twins before, during and following Scott’s time in space and preliminary results were announced in early 2017. While in flight, the telomeres on Scott’s white blood cells lengthened, genome-wide levels of DNA methylation dropped, and the composition of his gut microbiome underwent a dramatic shift. All these readings returned to pre-flight levels after landing, shedding light on how the body is impacted by extended time in space.

**2.** Researchers have identified the molecular pathway that leads to itching when exposed to urushiol, the oily sap found in poison ivy, poison sumac and poison oak. The Interleukin-33 (*IL-33*) gene is upregulated following contact with urushiol. The *IL-33* protein excites nerve fibers within the skin, sending the itch message to the brain. When the researchers blocked *IL-33* activity with neutralizing antibodies, urushiol-exposed mice reduced their scratching. This suggests a potential treatment approach – an antibody for human *IL-33* is currently being evaluated in early stage clinical trials.

**3.** Lifting the moratorium it imposed in 2013, the US Food and Drug Administration now allows 23andMe to sell genetic tests that provide disease risk information directly to consumers. The FDA has given approval for tests that provide information about genetic predisposition for 10 diseases and/or conditions. The FDA has also signed off on a process for expedited approval of future genetic health risk tests opening the door for broader consumer access to genetic information.

**4.** In an ambitious search for genes associated with human height, over 700,000 individuals were tested for nearly 200,000 rare DNA variants known to alter the function of protein-coding genes. Eighty-three variants were identified, some that altered height by more than 2cm (~ 8/10 of an inch) – ten times the effect of previously identified alleles. To date, over 800 genetic variants have been linked to human height, although the authors estimate more than 70% of the heritability of height is still unidentified.

**5.** The US Food and Drug Administration recently approved Nusinersen (trade name Spinraza) a new drug that alters mRNA splicing patterns to treat spinal muscular atrophy (SMA). Caused by mutations in the *SMN1* gene, the disease results from the degeneration of motor neurons in the spinal cord and lower brain stem. Nusinersen is an antisense oligonucleotide – a short synthetic strand of nucleotides that selectively binds mRNA transcribed from the *SMN2* gene, altering its splicing to produce a protein that functionally substitutes for the mutated *SMN1*. In clinical trials, treatment prolonged overall survival of infants with SMA and increased their ability to achieve motor milestones such as head control, rolling over and sitting independently. The biggest benefit was seen among babies not yet showing symptoms, suggesting the importance of early treatment.


See article references on pg.53

**6.** Octopuses, squid, and cuttlefish frequently turn to RNA editing to adjust or tweak the proteins produced from their genes. RNA editing involves swapping out adenosine bases with inosine in messenger RNAs, using enzymes that deaminate adenosine. Inosine is recognized as guanine during the process of translation, resulting in proteins different from what is encoded in the genomic sequence.

More than half of the transcripts in these animals show signs of editing, compared to just a handful of RNA edits in humans. Researchers speculate RNA editing may provide a more dynamic way to respond to altered environmental conditions or stimuli. However this RNA flexibility comes at a price – the regions where RNA editing occurs acquire DNA changes only rarely, resulting in slower evolution.

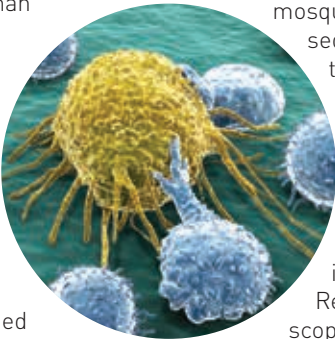
**7.** Asparagusic acid, a compound found in asparagus, is broken down into ammonia and various sulfur-containing products that produce a pungent odor in the urine of individuals who have recently eaten the vegetable. However, not everyone is able to detect this odor. Confirming and extending prior studies, a recent genome-wide analysis of nearly 7,000 individuals (40% who noted they could smell the odor in their urine) pinpointed a region on human chromosome 1. This area contains several members of the olfactory receptor 2 gene family, thought to be involved in the binding of an odor molecule and the cell signaling that notifies the brain of an aroma’s presence.

**8.** Researchers have identified a key regulator of stripes and pigment patterns in mammals. Working with the African striped mouse (*Rhabdomys pumiliom*), the scientists discovered the *Alx3* gene is expressed in cells where lighter stripes form, inhibiting a protein that would otherwise activate pigment production. Comparative genomics suggests this gene regulates pigmentation across most mammals, from chipmunks to zebras. Intriguingly, genes that coordinate skin and hair pigmentation also regulate the development of certain bone and nerve tissues. Mutations in *Alx3* can lead to neural-tube defects in mice and facial malformations in humans.

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

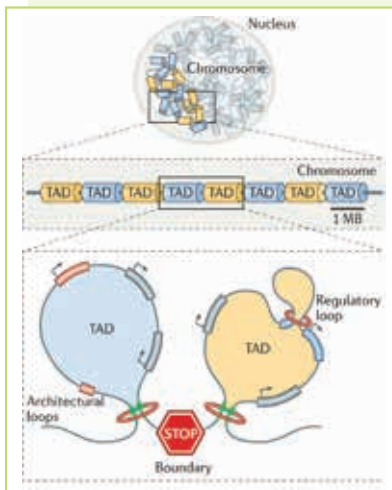
**9.** *Anopheles arabiensis* is the primary species of mosquito that carries malaria in East Africa. Some of these mosquitoes prefer feeding on humans while others gravitate towards cattle. To test whether genes influence host choice, cattle-fed and human-fed mosquitoes were gathered from across Tanzania. Genome sequencing showed mosquitoes that prefer cattle tend to carry a paracentromeric inversion on the R arm of chromosome 3. A greater understanding of the biology behind this preference may lead to approaches that reduce the incidence of malaria in humans.

**10.** Determining optimal treatments for cancer often requires molecular diagnostics to help scientists identify the genetic mutations that cause the tumor. Researchers have designed a smartphone-based microscope that can image and analyze specific DNA sequences and genetic mutations in tumor cells and tissue samples – without having to first extract DNA from them. The mobile microscope can also identify small numbers of cancer cells within a larger population of normal cells. The authors believe the device could be mass-produced for less than \$500, twenty times cheaper than standard multi-imaging microscopes. This tool could allow molecular testing to be offered anywhere, not just places with well-equipped medical laboratories.



## Chromosomal neighborhoods

Inside a cell's nucleus, each chromosome occupies a distinct "territory". Within that space, the DNA is folded into a number of large chromosomal loops (generally 800,000 to 1 million nucleotides) called topologically associated domains (TADs).



TADs are like neighborhoods – interactions between genes, enhancers and other regulatory regions take place within their borders. The edges of a TAD often include insulator proteins that block the regulatory elements in one TAD from inappropriately activating non-target genes in another.

This organizational system can be disturbed when DNA segments are duplicated, a fairly common feature of the human genome. For example, the *SOX9* gene

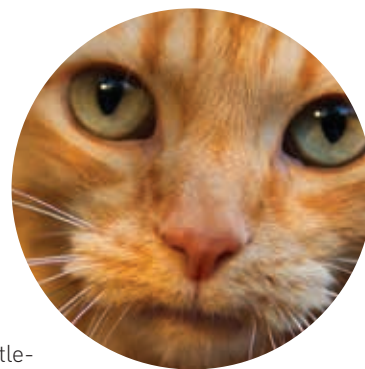
is a transcription factor located on chromosome 17. Duplications in this region of the genome can result in disorders with very different clinical profiles. Small duplications of the *SOX9* regulatory elements leads to female to male sex reversal (two X chromosomes, but anatomically male). Larger duplications don't impact gender, but instead cause an abnormality of the hands and feet. Intriguingly, duplications in-between these two lengths have no symptoms at all.

Scientists have learned that this region has two large TADs. One includes the *SOX9* gene and its regulatory elements, while the other contains two potassium channel genes (*KCNJ2* and *KCNJ16*) and their regulatory elements. The smallest duplication remains fully within the *SOX9* TAD, limiting its impact to the genes and regulators within this region. A duplicated enhancer increases *SOX9* expression, inappropriately activating male development pathways in a chromosomally female individual.

In contrast, the largest duplication extended into the second TAD, copying *SOX9* regulatory sequences, the *KCNJ2* gene and the boundary sequence between the TADs. A brand new TAD was formed where the *SOX9* regulatory sequences activate *KCNJ2* at the wrong time and in the wrong tissues, leading to the hand and foot abnormalities. A new TAD was also created from the mid-length duplication, but it only contained regulatory sequences. Because the adjacent *SOX9* and *KCNJ2* genes were shielded by the boundaries of the new TAD, there were no clinical consequences of the duplication.

**REFERENCE:** Franke M. et al. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* (2016) 538:265-9 doi: 10.1038/nature19800.

## Cat (non)domestication



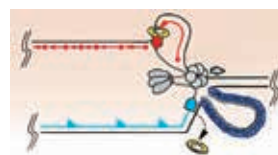
Using samples from across the last 9,000 years, a comparison of more than 200 cat genomes reveals information about the path to cat domestication. Archeological evidence suggests that cats began living alongside human settlements in the Fertile Crescent as early as 4400 B.C., likely attracted by rodents among human crops. This suggests that cats domesticated themselves, choosing to live near rodent-rich farms.

A second lineage of small African cats was first popularized in Egypt before ultimately being introduced across the Mediterranean. These cats may have had behaviors such as sociability that made them likeable human companions. It appears that cats were introduced widely across Asia and Europe as ancient humans brought them along land and sea trade routes – likely to control rodents during the journey.

In sharp contrast to dog genomes – which bear the signatures of extensive selective breeding, cat genomes remained relatively unchanged even after living in close quarters with humans for long periods of time. The tabby coat trait is one of the few traits not found in the wild cat relatives. It dates back to the Ottoman empire and was more widely introduced in domesticated cats during the Middle Ages.

**REFERENCE:** Ottoni C. et al. The paleogenetics of cat dispersal in the ancient world. *Nature Ecology & Evolution* (2017) 1, 0139 doi: 10.1038/s41559-017-0139.

## Replication re-imagined



The basic mechanism of DNA replication in *E. coli* is well studied and requires that helicase unwind the two strands as DNA polymerase copies the leading strand in a single continuous piece, while the lagging strand is copied in a series of fragments that must be further processed. Replication is thought of as an elegant dance of enzymatic action, with the construction of the two strands proceeding apace, even though the mechanisms differ. The presence of a mechanism that manages the speed of synthesis for each strand has been long hypothesized. These assumptions are evident in textbook descriptions and documentary animations.

For the first time, scientists have recently been able to watch the replication of a single DNA molecule, and surprisingly their findings do not match the old assumptions. Using sophisticated imaging technology and a great deal of patience, the researchers were able to watch DNA from *E. Coli* bacteria as it replicated and measure how fast enzyme machinery worked on the different strands. As scientists started watching individual DNA strands, they noticed something unexpected. Replication stops unpredictably and when it starts up again can change speed. The speed can vary about ten-fold. Sometimes the lagging strand synthesis stops, but the leading strand continues to grow. There doesn't appear to be any coordination between the strands.

The two strands appear to be replicated independently. Researchers describe the process as "like heavy traffic on the freeway" with lanes slowing or stopping and other lanes flowing rapidly. Despite these intermittent variations in speed, replication of the two strands of the template DNA molecule completes at about the same time.

**REFERENCE:** Graham J.E. et al. Independent and Stochastic Action of DNA Polymerases in the Replisome. *Cell* (2017) 169:1201-13 doi: 10.1016/j.cell.2017.05.041.



## Populating the planet

After evolving in Africa about 200,000 years ago, how have humans spread to populate the entire globe? Three independent research projects, all using genomic sequencing, have concluded that today's non-Africans descended from one migration out of Africa between 50,000 and 80,000 years ago. The teams sequenced 787 genomes from indigenous populations around the globe. This model of human migration does not mean that earlier voyages out of Africa did not occur – just that those first travelers made little to no contribution to the modern non-African genome. In fact, one group found that about 2 % Papua New Guineans carry DNA from ancestors who left Africa before other Eurasians.



**REFERENCES:** Malaspinas A.S. et al. A genomic history of Aboriginal Australia. *Nature* (2016) 538:207-14 doi: 10.1038/nature18299.

Mallick S. et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* (2016) 538:201-6 doi: 10.1038/nature18964.

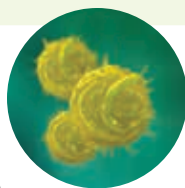
Pagani L. et al. Genomic analyses inform on migration events during the peopling of Eurasia. *Nature* (2016) 538:238-42 doi:10.1038/nature19792.

## Pinning down DNA

X-inactivation silences transcription on one of the two X chromosomes inside the cells of female mammals. As a result, protein levels encoded by most genes on the X chromosome are equivalent between females and males (who only have 1 X chromosome). This process of dosage compensation alters the three-dimensional structure of the X chromosome, leading to a tightly packed structure known as a Barr body. X inactivation requires a > 17,000 nucleotide-long noncoding RNA (lncRNA) molecule called Xist. lncRNA molecules are not translated into protein, but carry out their function by directly binding to DNA, altering the structure and activity of nearby genes. Xist coats the soon-to-be inactive X chromosome, setting in motion a chain of events that leads to silencing.

Recently, researchers working with mouse embryonic stem cells have identified part of the mechanism that allows Xist binding to spread across the X chromosome. In addition to binding the X chromosome, a subset of Xist molecules attach to the lamin B receptor (LBR) which is integrated into the nuclear membrane. The interaction between Xist and LBR “tethers” the X chromosome to the nuclear membrane, holding it in place. This brings other parts of the X chromosome into proximity, allowing Xist to bind more distant DNA regions. Other proteins work alongside Xist to increase DNA methylation, remove acetyl groups from DNA histones and transform the X chromosome into the small and generally silent Barr body.

**REFERENCE:** Chen C.K. et al. Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* (2016) 354:468-72 doi: 10.1126/science.aae0047.



## Sonic hedgehog and snake legs

The fossil record suggests that snakes arose from lizards over 100 million years ago. Along the way, snakes lost their legs – at least most of them did. Modern day pythons and boas have tiny remnants of back legs, buried in the muscles towards their tails. Most other snakes, such as garters, vipers and cobras, are completely limbless. Analysis of a DNA enhancer has shed new light on the formation of snake limbs.



Enhancers are regulatory regions of the genome that – when bound by certain proteins – increase the transcription of specific genes. In vertebrates, the Zone of Polarizing Activity Regulatory Sequence (ZRS) is an enhancer for the *Sonic Hedgehog* (*Shh*) gene (named after the SEGA® character Sonic the Hedgehog because mutations in this gene result in embryos covered in small pointy spike-like projections). *Shh* is a master on/off switch for forming multiple organs during embryonic growth. Separate enhancers activate *Shh* to oversee different body parts – ZRS coordinates *Shh* activity associated with arm and leg development. In humans, mutations in the ZRS enhancer can cause malformations like shortened limbs or extra fingers/toes. The ZRS enhancer lies over 1 million nucleotides upstream of *Shh*, within an intron of an unrelated gene.

Researchers compared the sequence of ZRS across multiple reptile and mammalian species. In lizards and mammals, ZRS keeps *Shh* active throughout the process of leg development. However, a number of mutations and deletions are present in snakes that minimize the ability of ZRS to stimulate *Shh*.

In most snakes, the ZRS enhancer is so degraded that *Shh* remains silent in the cells that would otherwise give rise to limbs. Intriguingly, python ZRS contains fewer mutations, allowing it to briefly activate *Shh*. As a result, python embryos form hind limb buds and even build cartilage-based legs, feet and toes. However, without constant *Shh* activity, most skeletal structures don't develop so that by the time baby pythons hatch, they only contain rudimentary legs.

**REFERENCES:** Kvon E.Z. et al Progressive Loss of Function in a Limb Enhancer during Snake Evolution. *Cell* (2016) 167:633-42 doi: 10.1016/j.cell.2016.09.028.

Leal F. and Cohn M.J. Loss and Re-emergence of Legs in Snakes by Modular Evolution of Sonic hedgehog and HOXD Enhancers. *Current Biology* (2016) 26:2966-73 doi: 10.1016/j.cub.2016.09.020.

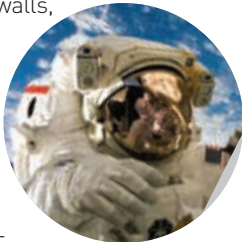


UF Health researchers uncover how snakes lost their legs  
[youtu.be/OMKRNhl8s0c](https://youtu.be/OMKRNhl8s0c)

## NEW FINDINGS — BACTERIA & VIRUSES

### Bacterial growth in space

Bacteria behave differently in space, exhibiting decreased lag phase, higher final cell densities, thickened cell walls, and increased biofilm formation. These characteristics appear to increase virulence and antibiotic resistance, with important implications for the long-term human presence in space. A recent comparison of gene expression between ground-based and space-based liquid cultures of *E. coli* has provided a molecular explanation for these differences.



On earth, the exchange of nutrients and metabolic waste across the bacterial membrane relies on convection, which is influenced by gravity. In space, a lack of gravity limits molecular transport to diffusion only. Over time, the microclimate around each bacterial cell becomes depleted of glucose and overrun with acidic waste products. Essentially in starvation mode, the bacteria respond by transcribing genes involved in the metabolism of other nutrients. They also upregulate genes associated with acid resistance, including factors known to increase virulence.

The confirmation of the limited transport hypothesis provides scientists with new approaches for maintaining microbes in microgravity environments associated with vaccine and antibiotic production. It also sheds light on growing non-bacterial cells, which are likely subject to similar transport conditions.



*The lab of HudsonAlpha researcher Shawn Levy, PhD, contributed to this research.*

**REFERENCE:** Zea L. et al. A molecular genetic basis explaining altered bacterial behavior in space. *PLoS ONE* (2016) 11:e0164359 doi: 10.1371/journal.pone.0164359.



connected to the 1545 *cocoliztli*. The researchers recovered and sequenced bacterial DNA found in the teeth of several people in that group that matched a *Salmonella enterica* strain called Paratyphi C. This strain of salmonella causes enteric fever, more commonly known as typhoid. A second bioRxiv preprint presents evidence that Paratyphi C was circulating in Europe more than 300 years before Europeans reached Mexico. Together, these studies provide what is possibly the first genetic evidence for a European pathogen that caused a massive population decline in indigenous populations of Mexico.

**REFERENCES:** Vågene A.J. et al. *Salmonella enterica* genomes recovered from victims of a major 16th century epidemic in Mexico. (2017) bioRxiv doi: 10.1101/106740.

Zhou Z. et al. Millennia of genomic stability within the invasive Para C Lineage of *Salmonella enterica*. (2017) bioRxiv doi: 10.1101/105759.

### Gut microbes and anxiety

Irritable bowel syndrome (IBS) affects the large intestine, causing abdominal cramping and pain, bloating and a change in bowel habits. Some individuals with IBS also report symptoms of generalizable anxiety. A recent study in mice suggests the gut microbiome contributes to both bowel and behavioral symptoms. Germ-free mice received fecal microbiota from healthy human donors, individuals with IBS and anxiety, or individuals with IBS alone. The microbiome from either category of IBS patients impacted gut function, leading to faster content movement and higher gut permeability. Additionally, mice that received microbiota from IBS patients with anxiety showed increased measures of anxiety themselves.



**REFERENCE:** De Palma G. et al. Transplantation of fecal microbiota from patients with irritable bowel syndrome alters gut function and behavior in recipient mice. *Science Translational Medicine* (2017) 9:eaaf6397 doi: 10.1126/scitranslmed.aaf6397.

### Did salmonella decimate the Aztecs?

In the 16th century, infectious diseases devastated the native population of modern Mexico, taking their numbers from 24 million to 1 million in a single century. Many of those disease were introduced by Europeans, but little evidence remains to verify the actual pathogens that caused the outbreaks. However, a preprint published in bioRxiv presents a study that may identify salmonella as the cause of a major epidemic – known locally as a *cocoliztli* – that swept through Mexico from 1545-1550. For the study, researchers extracted DNA from the teeth of 29 people buried in southern Mexico. All but five were



### Atopic dermatitis

People with eczema have dry, flaky skin that is prone to infection by the harmful bacteria *Staphylococcus aureus*. Researchers have identified two other strains of *Staphylococcus* found on healthy human skin that inhibit and even kill *S. aureus*. These bacteria secrete peptides that work with the human immune system to reduce the growth of pathogenic bacteria. Unlike antibiotic ointments, the peptides do not interfere with the growth of other skin microbes. In a clinical trial, patients with eczema applied a lotion loaded with the helpful bacteria. After 24 hours, their skin showed dramatically reduced levels of *S. aureus*. These promising results may lead to new ways to treat eczema.

**REFERENCE:** Nakatsuji T. et al. Antimicrobials from human skin commensal bacteria protect against *Staphylococcus aureus* and are deficient in atopic dermatitis. *Science Translational Medicine* (2017) 9:eaah4680 doi: 10.1126/scitranslmed.aah4680.



## A deadlier form of Ebola

The 2013-16 Ebola epidemic was unlike any that had occurred before, affecting over 28,000 West Africans. The virus appeared in heavily populated cities, and poor public health infrastructure made it difficult to track, treat and quarantine. Recent findings suggest genetics may have also driven its rapid spread.

Two research teams independently sequenced over a thousand viral genomes isolated from patients throughout the epidemic. Both groups identified a mutation called GP-A82V that first appeared in patients about 3 months into the outbreak. Viruses carrying this mutation rapidly spread while the original version of the virus disappeared within a few months.

The mutation alters a glycoprotein – a docking station on the viral surface that makes contact with a host cell and opens a passageway for entry. This mutation allows the virus to infect two to four times as many human cells. Simultaneously, it confers greater lethality – individuals were twice as likely to die if they had been infected with virus carrying GP-A82V.

**REFERENCES:** Urbanowicz R.A. et al. Human Adaptation of Ebola Virus during the West African Outbreak. *Cell* (2016) 167:1079-87 doi: 10.1016/j.cell.2016.10.013.

Diehl W.E. et al. Ebola Virus Glycoprotein with Increased Infectivity Dominated the 2013–2016 Epidemic. *Cell* (2016) 167:1088-98 doi: 10.1016/j.cell.2016.10.014.



## Genetic risks for catching cold

Although the range varies widely, the CDC notes that the average healthy adult has 2 to 3 colds per year. Scientists recently published the case of a 5-year-old girl affected with over 40 viral respiratory infections since birth, many life-threatening. Suspicious that she might have a primary immune deficiency, whole exome sequencing was performed. A homozygous missense mutation was identified in the *IFIH1* gene, which encodes a protein known as MDA5. The dysfunctional MDA5 prevented the girl's immune system from recognizing and responding to human rhinoviruses – the main cause of the common cold.

**REFERENCE:** Lamborn, I.T. et al. Recurrent rhinovirus infections in a child with inherited MDA5 deficiency. *Journal of Experimental Medicine* (2017) 214:1949-72 doi: 10.1084/jem.20161759.

## Bacteria and the diet rebound

Over one-third of US adults struggle with obesity. While multiple approaches result in weight loss, more than 80% of dieters are unable to maintain their losses. When the pounds are regained, people often find they are at an even higher weight than before. The reasons behind this “yo-yo effect” are poorly understood, although new research suggests some of the blame may lie with the microbes inhabiting our gut.



To replicate the yo-yo effect, scientists fed mice on cycles of high-fat or normal mouse chow. Similar to humans, mice gained weight on the high-fat diet, lost it when eating normal chow and then regained the lost weight plus more after returning to the high-fat food. The initial round of dieting somehow predisposed the mice to more rapid weight gain post-diet.

Looking for explanations, researchers compared a host of biological features between cycling mice and those who remained on normal chow. While a number of traits (body fat content, serum cholesterol and energy expenditure) were altered during weight gain, these reverted to baseline once the mice returned to normal weight. In contrast, the gut microbiomes of the cycling mice maintained many of their obesity-induced characteristics weeks after weight had normalized. Scientists transplanted these gut microbes into mice that previously lacked a microbiome. Feeding these mice high-fat chow led to higher than expected weight gain. Scientists developed an algorithm to predict the degree of weight regain based on the frequency of 189 different gut microbes.

The gut microbiome assists our digestive system in extracting energy from the food we consume. The composition of specific bacterial communities within our gut microbiome shifts as our diet changes, responding to the types of nutrients available. It took 21 weeks for mice microbiomes to return to pre-experiment form – five times the length of the initial high fat or dieting cycles. The researchers hypothesize this microbial lag may help drive the yo-yo effect. If mice resume eating high-fat chow before their microbiome has fully shifted to a post-diet composition, the effect on weight regain is exaggerated. Until additional studies are performed, researchers caution against generalizing the results to humans.



**REFERENCE:** Thaiss C.A. et al. Persistent microbiome alterations modulate the rate of post-dieting weight regain. *Nature* (2017) 540:544-51 doi: 10.1038/nature20796.

### When genetic testing goes wrong

A previously healthy 13-year-old Hispanic boy unexpectedly died during his sleep. An autopsy revealed some unexplained cardiac abnormalities but nothing conclusive. The boy's family members were examined for signs of heart irregularities and a faster than normal heartbeat was briefly detected in a brother, leading to a surgically implanted defibrillator. A genetic test on this brother also identified a variant in *KCNQ1*, a gene associated with a cardiac arrhythmia called long QT syndrome (LQT). The genetic testing company categorized the DNA change as a "probable deleterious mutation." Believing this was the likely cause of the original sudden death, the extended family underwent genetic testing, resulting in over two dozen LQT diagnoses.

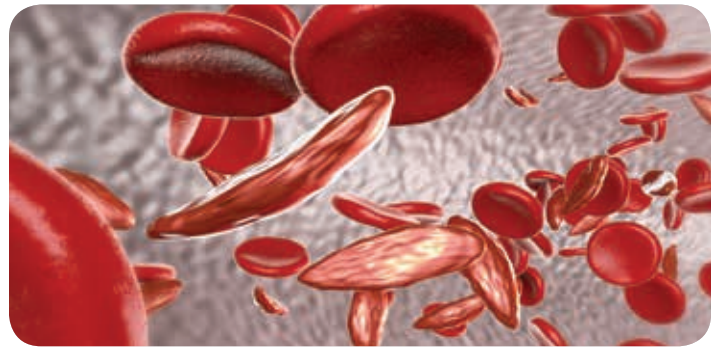
Because none of the affected individuals had history of cardiac problems or showed symptoms of arrhythmia, the family sought a second opinion. Although DNA samples existed from the boy who died, molecular analysis had never been performed. Genetic testing showed that he did not inherit the *KCNQ1* variant found in his brother. Rather, he had a spontaneous mutation in a different gene known to cause heart failure and early death. On further review, the *KCNQ1* variant present in this family was reclassified as completely benign, meaning none of the individuals were at risk of LQT. It was a genetic red herring.

This cautionary tale has several important points. Although precision medicine has great promise, without proper understanding of whom to test and how to critically interpret the results, molecular data may point in the wrong direction.

Instead of "surrogate testing" the living brother, the physician should have tested the DNA samples available from the autopsy. Additionally, the genotype and phenotype didn't match — none of the supposedly "affected" individuals actually showed symptoms of LQT. This highlights a flaw in databases used to assess whether DNA variants are harmful: in the past, if an amino-acid altering variant was rare, that was often sufficient to merit a "disease-causing" tag. Although today's criteria are much more stringent, many prior categorizations are likely incorrect, and it's estimated that 10% of "disease causing" long QT variants are misclassified. This points to a critical need to re-evaluate variants present in existing mutation databases.

#### REFERENCE:

Ackerman J.P. et al. The Promise and Peril of Precision Medicine: Phenotyping Still Matters Most. *Mayo Clinic Proceedings* (2016) 91:1606-16 doi: 10.1016/j.mayocp.2016.08.008.



### New treatments for sickle cell anemia

Sickle Cell Anemia (SCA) affects an estimated 25 million people worldwide. It is caused by a mutation in the beta globin gene that codes for one of the proteins that makes up the hemoglobin molecule. Under low oxygen conditions, the mutation causes hemoglobin to form long chains that distort the red blood cells into a sickle shape. Sickled cells have a much shorter lifespan and clog small capillaries and vessels, leading to pain and organ damage.

Although the mutation that causes SCA was identified over 65 years ago, stem cell transplants offer the only cure. Recently, several labs have developed therapies based on gene editing and/or gene therapy.

Two groups have directly modified the beta globin gene using CRISPR gene editing. Human stem cells were isolated from patients with SCA, edited to correct the mutation and injected into mice. The modified stem cells traveled to the bone marrow where they multiplied, differentiating into red blood cells that produced normal beta globin.

Other scientists are boosting the transcription of gamma globin genes, which encode a fetal form of hemoglobin that also transports oxygen. Normally silenced shortly after birth, re-activating gamma globin genes can significantly minimize the symptoms of SCA. One group is using CRISPR to delete the "off" switch that controls the fetal to adult hemoglobin switch. A second group is using short-hairpin RNA fragments to silence a gene which itself normally reduces gamma globin gene expression. Both groups have increased fetal hemoglobin in cell cultures and/or lab mice to levels that eliminate sickling.

Lastly, French scientists used gene therapy to treat a patient with SCA, inserting a modified version of the beta globin gene. The modification created a protein that inhibited the patient's own mutant beta globin from polymerizing into the long chains. Fifteen months post-treatment the patient had experienced no episodes of sickling and was symptom free.

**REFERENCES:** DeWitt M.A. et al. Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Science Translational Medicine* (2016) 8:360ra134 doi: 10.1126/scitranslmed.aaf9336.

Dever D.P. et al. CRISPR/Cas9  $\beta$ -globin gene targeting in human haematopoietic stem cells. *Nature* (2016) 539:384-9 doi: 10.1038/nature20134.

Traxler E.A. et al. A genome-editing strategy to treat  $\beta$ -hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nature Medicine* (2016) 22:987-90 doi: 10.1038/nm.4170.

Brendel C. Lineage-specific BCL11A knockdown circumvents toxicities and reverses sickle phenotype. *Journal of Clinical Investigation* (2016) 126:3868-78 doi: 10.1172/JCI87885.

Ribeil J.A. et al. Gene Therapy in a Patient with Sickle Cell Disease. *New England Journal of Medicine* (2017) 376:848-55 doi: 10.1056/NEJMoa1609677.



## Sequencing babies

BABYSEQ, an NIH funded study designed “to probe the risks and benefits of sequencing newborns’ DNA,” is finding few parents willing to enroll their newborns. Early surveys indicated a strong parental interest in having whole genome sequencing information to guide health care for newborns. However, that initial interest has not translated into participation, with only about 7% of the new parents who are approached ultimately deciding to enroll in the study. Even among parents whose newborns were in the neonatal intensive care unit, fewer than 8% of parents offered DNA sequencing for their baby consented. Parents cited fears of finding negative/unclear results or concerns regarding potential insurance discrimination as reasons for declining to have their newborn sequenced.

**REFERENCE:** American Society of Human Genetics. Researchers Assess Impact of Exome Sequencing on Newborns and Their Families. N.p., 19 Oct. 2016. Web. 2 Aug. 2017 <http://www.ashg.org/press/201610-BabySeq.html>.

## Using “good” mutations

More than twenty years ago, Anna Feurer participated in a health fair sponsored by her employer, where she gave a blood sample to measure her cholesterol. Little did she suspect that would begin a decades-long process leading to new medications for preventing heart disease.

It turns out that Ms. Feurer, along with other members of her family, have incredibly low triglyceride levels. Genetic analysis ultimately determined she carries inactivating mutations in both copies of the *ANGPTL3* gene, which typically regulates triglyceride metabolism. Recently, the coronary arteries of Ms. Feurer and two of her siblings were scanned for evidence of atherosclerosis. All three individuals had no evidence of plaque – even though one sibling had several risk factors for heart disease.

A study of more than 180,000 participants identified potentially inactivating DNA changes in *ANGPTL3*, many of which were functionally analyzed in a mouse model. The data suggest approximately 1 in 300 individuals carried a single *ANGPTL3* loss of function variant, which resulted in 11% lower total cholesterol, 12% lower LDL and 17% lower triglycerides. In addition, *ANGPTL3* mutations reduce heart attack risks by one-third. Other large-scale studies have identified similar results.

Since multiple lines of evidence show *ANGPTL3* deficiency protects against heart disease, several pharmaceutical companies have begun drug discovery efforts in an attempt to mimic the impact of Ms. Feurer’s genetic mutation. Pre-clinical studies in mice slowed the progression of atherosclerosis and early clinical trials in humans have led to reductions in levels of HDL, LDL and triglycerides. Ultimately, many experts believe these medications will become the next class of blockbuster drugs, used to reduce the risk of heart disease.

**REFERENCES:** Stitzel, N.O. et al. *ANGPTL3* Deficiency and Protection Against Coronary Artery Disease. *Journal of the American College of Cardiology* (2017) 69:2054–63 doi: 10.1016/j.jacc.2017.02.030.

Dewey, F.E. et al. Genetic and Pharmacologic Inactivation of *ANGPTL3* and Cardiovascular Disease. *New England Journal of Medicine* (2017) 377:211–21 doi: 10.1056/NEJMoa1612790.

Graham, M.J. et al. Cardiovascular and Metabolic Effects of *ANGPTL3* Antisense Oligonucleotides. *New England Journal of Medicine* (2017) 377:222–32 doi: 10.1056/NEJMoa1701329.

Kolata, G. “Rare Gene Mutations Inspire New Heart Drugs.” *New York Times*, 24 May 2017, [www.nytimes.com/2017/05/24/health/heart-drugs-gene-mutations.html?\\_r=0](http://www.nytimes.com/2017/05/24/health/heart-drugs-gene-mutations.html?_r=0). Accessed 19 June 2017.



## First “3 parent baby” born

A baby boy has been born through mitochondrial replacement IVF therapy – a controversial technique that uses the genetic information from three people. Profiled in the *2015 Guidebook*, the process seeks to prevent the maternal transmission of mitochondrial disease by combining the nucleus from a mother’s egg with the cytoplasm and mitochondria of a donor (the “third parent”), followed by fertilization with the father’s sperm. Currently banned in the US, the procedure was carried out in Mexico by US-based scientists. A relatively small amount of the mother’s mutated mitochondria transferred into the donor egg. Whether this will affect the child’s health remains unknown.

**REFERENCE:** Zhang, J. et al. Live birth derived from oocyte spindle transfer to prevent mitochondrial disease. *Reproductive Biomedicine Online* (2017) 34:361–8 doi: 10.1016/j.rbmo.2017.01.013.



## Providing patients answers

Exemplifying the power of large-scale sequencing to solve diagnostic odysseys, researchers recently described a project that provided a genetic diagnosis for 100 children with intellectual disability or developmental delay. The group sequenced the exome or genome of 371 individuals with symptoms such as impaired cognition, failure to meet developmental milestones, facial and skeletal abnormalities, autism and seizures. In 309 of those cases, the team also sequenced at least one parent. Comparing the child’s genome to the parent’s genome helped filter the data down to the DNA differences that were new in the child. They were able to provide an answer for 27% of the affected children who enrolled in the study. Variants of uncertain significance were identified in an additional 11%. According to the scientists, their research supports the value of whole genome sequencing as a diagnostic tool for clinical and research applications in pediatric neurological disease.

**REFERENCE:** Bowling K.M. et al. Genomic diagnosis for children with intellectual disability and/or developmental delay. *Genome Medicine* (2017) 9 doi: 10.1186/s13073-017-0433-1.



The labs of HudsonAlpha researchers Greg Cooper, PhD, Greg Barsh, MD, PhD, Neil Lamb, PhD, and Richard Myers, PhD contributed to this research.

### Childhood cancer and lifetime risk

Twelve percent of childhood cancer survivors have increased genetic risk of developing a secondary cancer later in life, according to a study presented at the 2017 annual meeting of the American Association for Cancer Research. For the study, a group led by researchers at St. Jude Children's Research Hospital sequenced and analyzed whole genomes from more than 3,000 long-term childhood cancer survivors. The team looked at germline mutations – changes to genes that can be inherited – in 156 genes that are associated with an elevated cancer risk. Extending this finding to the 400,000 US childhood cancer survivors suggests more than 30,000 carry predisposing genetic mutations that put them at greater risk for a secondary cancer. Since these mutations can be inherited, the children of these cancer survivors are also at increased risk for developing cancer. The project involved long-term cancer survivors who received treatment at St. Jude and were enrolled in the St. Jude Lifetime Cohort (St. Jude LIFE) study. Participants return to St. Jude periodically for several days of clinical and functional assessments. The St. Jude LIFE project represents the first time whole genome sequencing has been used in a large group of cancer survivors to study the impact of genetic factors on lifetime cancer risk. As a result of the study, the researchers recommend expanding genetic cancer screening and counseling to include childhood cancer survivors who have been diagnosed with second cancers. In addition, these findings could help doctors design personalized therapeutic approaches for children who have been newly diagnosed with cancer based on individual genetic profiles.

**REFERENCE:** Wang Z. et al. Germline mutations in cancer predisposition genes and risk for subsequent neoplasms among long-term survivors of childhood cancer in the St. Jude Lifetime Cohort. In: Proceedings: AACR Annual Meeting 2017.; April 1-5, 2017; Washington DC. Abstract 3001. doi: 10.1158/1538-7445.AM2017-3001.

 HudsonAlpha researcher Shawn Levy, PhD, contributed to this research.



### Gene activity after death

What happens inside cells when an organism dies? Does gene expression flip off like a switch or gradually fizzle out as raw materials and available energy are consumed? Surprisingly, new research reveals that neither option occurs, with some genes actually increasing transcription in the period after death. In both zebra fish and mice, over 1000 genes are actively transcribed up to 96 hours postmortem, including genes associated with stress, immunity, apoptosis, transport, embryonic development and cancer.

In some cases, the transcriptional increase makes sense. More transcription in genes related to stress, inflammation and the immune response suggest that individual cells react to changing conditions (loss of homeostasis, lack of blood flow, increased CO2 levels and cellular decomposition) much as they would to an injury sustained in life. Genes associated with apoptosis likely are responding to growing cellular damage, while up-regulating transport genes reflects an attempt to re-establish homeostasis. The increased expression of genes associated with embryonic development is unexpected and harder to explain. In adult cells, these genes are generally located within tightly packed chromatin to prevent transcription. After death, DNA may unravel, providing RNA polymerase with transcriptional access to these genes.

While additional research is needed, these data suggest a step-wise shutdown occurs in organismal death. At some point, measures of transcript abundance could more precisely calculate time of death for forensic purposes.



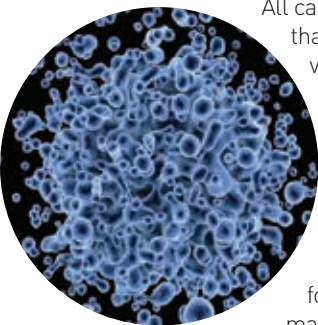
**REFERENCE:** Pozhitkov A.E. et al. Tracing the dynamics of gene transcripts after organismal death. Open Biology (2016) 7:160267 doi: 10.1098/rsob.160267.

### Somatic mutations and heart disease

Cancer is most frequently caused by the accumulation of mutations in the genome of somatic cells. It is believed that cells progress to pre-malignant and ultimately malignant stages as mutations are acquired in genes important for cell growth and regulation. For example, scientists have identified a set of genes that are commonly mutated in the hematopoietic stem cells. Found in bone marrow, these stem cells divide and mature into red blood cells, white blood cells and platelets. This set of DNA mutations is rarely observed in individuals under 40 but found in more than 10% of individuals older than 70. People who have acquired these mutations are at a ten-fold higher risk of cancers of the blood.

Recently, these age-related mutations have also been associated with atherosclerosis – the buildup in the arteries of fatty deposits known as plaques. When atherosclerosis occurs in vessels that supply blood to the heart, the disorder is known as coronary heart disease. This connection to atherosclerosis is surprising, as it marks the first time that acquired mutations have been associated with a disorder other than cancer. An analysis of data from over 8,000 individuals enrolled in cardiovascular case/control studies showed that having one of these mutations nearly doubled the risk for coronary heart disease.

## Recurrence and patient treatment



All cancers are not the same. Tumors, even those that form in the same part of the body, may have very different behaviors in terms of growth rate, likelihood of metastasizing, response to various therapies and risk of recurrence. Modern screening and detection methods are adept at finding cancers early but have done little to classify these behaviors. Accordingly, clinicians have historically treated all tumors aggressively, even if this meant overtreatment for certain patients. New molecular technologies may be changing this approach.

An FDA-approved diagnostic test called MammaPrint, measures the activity of 70 genes shown to be predictive of tumor recurrence and classified them into low or high risk categories. In 2016, researchers found that early-stage breast cancer patients with low risk had greater than 94% metastasis-free survival rate at five years – whether they received chemotherapy or not. In a follow-on study, the researchers analyzed a different population of women who had been diagnosed up to 20 years ago. A subset of tumors (15%) could be categorized as ultra-low risk, with less than 3% risk of recurrence even two decades after diagnosis. It's likely these women could be effectively treated by a straightforward lumpectomy rather than more radical surgeries, chemotherapies or long-term hormone treatments. Assuming these findings can be replicated, oncologists will have a powerful tool to tailor their patient's treatment.

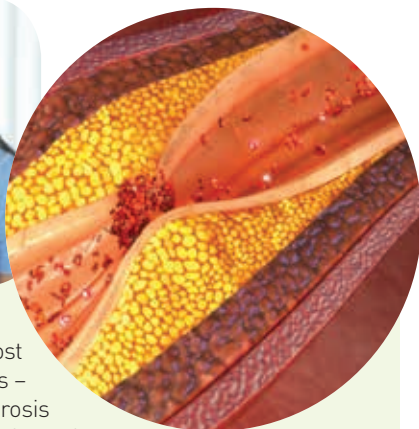
**REFERENCE:** Esserman L.J. et al. Use of Molecular Tools to Identify Patients with Indolent Breast Cancers with Ultralow Risk Over 2 Decades. *JAMA Oncology* (2017) published online doi: 10.1001/jamaoncol.2017.1261.



Follow-on studies in mice found that when one of the most commonly identified mutations – *Tet2* – was present, atherosclerosis was accelerated and the cells showed increased levels of inflammation, an early contributor to plaque formation.

Although several risks for atherosclerosis (high cholesterol, Type 2 diabetes, high blood pressure and smoking) have been well studied, many affected individuals have none of these risk factors. Acquired mutations in hematopoietic cells may represent an additional pathway for disease formation. Additional work is needed to determine whether the effect of these acquired somatic mutations can be blunted through cholesterol-lowering medications or anti-inflammatory treatments.

**REFERENCE:** Jaiswal, P. et al. Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *New England Journal of Medicine* (2017) 377:11-21 doi: 10.1056/NEJMoa1701719.



## Finding protective DNA variants

As more and more genomes become available, analysis of these huge data sets reveals important clues regarding risks and protectors for human disease. Mutations in the *BRCA1/2* genes are one of the most thoroughly studied genetic risks for breast cancer. Not all carriers of widely publicized *BRCA1/2* mutations develop cancer, and the reasons why have not been clear. A recent study screened a database of over 3,000 exomes, identifying 15 women with predicted protein-disrupting mutations in *BRCA1/2* who had not developed early-onset breast cancer. The exomes of these women were further analyzed for DNA changes that might protect against cancer. A significant number carried an additional mutation in the *COMT* gene, which encodes an enzyme that helps regulate hormones such as estrogen. This preliminary finding requires further confirmation, but highlights a previously unknown correlation between *BRCA1/2* and *COMT*. Additionally, it highlights how the growing availability of large-scale genomic information can be utilized to uncover biologically relevant DNA variants.



**REFERENCE:** Movassagh M. et al. Co-Occurrence of *COMT* and *BRCA1/2* Variants in a Population. *New England Journal of Medicine* (2017) 376:2090-1 doi: 10.1056/NEJMc1701592.

## Stem cells reverse aging in mice

There is a growing consensus that aging is controlled by epigenetic alterations such as DNA methylation, post-translational modifications of histones and chromatin packaging. Over an organism's life, epigenetic marks added to or removed from DNA alter the activity of certain genes, potentially reducing the ability to maintain and repair tissues. In a petri dish, adult cells have been reprogrammed to behave as embryonic stem cells by activating four transcription factor that reset epigenetic marks genome-wide. For the first time, these same transcription factors have been activated in a living organism – mice affected with a premature aging disorder. Treatment led to a 30% increase in lifespan and improvements in overall appearance as well as digestive, cardiovascular, and muscular function. However, there was a fine line between reversing aging and doing harm: the transcription factors could be activated for only two days per week – higher levels reprogrammed too many adult cells into undifferentiated embryonic stem cells, leading to organ failure and tumor formation.

**REFERENCE:** Ocampo A. et al. In Vivo Amelioration of Age-Associated Hallmarks by Partial Reprogramming. *Cell* (2016) 167:1719-33 doi: 10.1016/j.cell.2016.11.052.







### Approved for planting

The Environmental Protection Agency has approved a genetic engineering tool for controlling a major agricultural pest. The new insecticide is encoded within the genome of a corn plant and uses RNA interference (RNAi) to target the corn rootworm. RNAi occurs naturally in plants, animals and humans to regulate the activity of genes. Scientists have used this process to generate biotechnology tools that assist farmers. Scientists inserted a gene into the corn plant that manufactures a specific sequence of RNA. When rootworm larvae eat corn roots, or adult beetles munch on the leaves and pollen, they ingest the RNA fragment. Inside rootworm cells, this RNA fragment interferes or "silences" a rootworm-specific gene critical for growth, killing the pest. In contrast to a pesticide sprayed across a field, RNAi provides specific insect control - targeting only the rootworm. It is harmless to other insects, plants or animals.

RNAi technology has been used to create other genetically altered crops, like apples or potatoes that don't brown because the browning genes have been silenced. However, this is the first approval for an RNAi-based insecticide. The modified corn seeds are expected to be commercially available by 2020, through a collaboration between Monsanto and Dow.

**REFERENCE:** United States Environmental Protection Agency. Chemical Safety and Pollution Prevention. EPA Registers Innovative Tool to Control Corn Rootworm. N.p., 15 June 2017. Web. 9 July 2017. <https://www.epa.gov/newsreleases/epa-registers-innovative-tool-control-corn-rootworm>.

### Sequencing wild wheat

More than 10,000 years ago, farmers in the Fertile Crescent domesticated wild wheat grasses, paving the transition to agriculture and settlements. Specific strains of wild emmer wheat were preferentially selected for traits like larger grain size, higher numbers of grain and ease of harvesting. Over time, this produced the modern varieties of durum (pasta) and bread wheat, which today provide nearly 20% of the calories consumed by humans around the globe.

The genome of wild emmer wheat was recently sequenced, identifying genes involved in domestication. While a ripe head of wild wheat is easily shattered into individual grains, modern wheat remains attached to the stalk – an important trait for harvesting. It appears that loss of function mutations in two genes have led to nonshattering stalks. These are recessive mutations, meaning modern wheat is homozygous for the mutations at both genes.

Today's economically important wheat strains lack beneficial traits (such as drought tolerance and disease resistance) that are common in their wild ancestor. Identifying the DNA variants responsible for these traits will speed future breeding efforts to produce hardier crops that produce more grain with less water.

**REFERENCE:** Avni R. et al. Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* (2017) 357:93-7 doi: 10.1126/science.aan0032.



### Improving the common bean

The common bean is one of the most important grain legumes for many people around the world. It is a staple food crop for people in tropical regions, and is also an important source of household income in Africa and Latin America. The common bean has a number of cultivated varieties including kidney, pinto, navy, green and wax beans.

The common bean breeding programs seeks to genetically improve the plant so that fewer seeds produce a more abundant and nutritious harvest requiring less water and fertilizer.





## Mutation protection

A 234-year old oak tree in Switzerland is revealing new information about how plants protect themselves from genetic mutations and sustain such long lives. Researchers sequenced the plant's genome from several different branches of the tree, called the "Napoleon Oak" because it was young when the French leader marched through Lausanne in 1800. Because copying the genome before cell division often introduces mutations, scientists expected DNA in newer branches would be significantly different from older branches. Surprisingly, they found the overall number of mutations was much lower than expected, based on the number of cell

divisions estimated to have occurred between older and newer branches.

Animals isolate the cells that ultimately will become eggs and sperm, subjecting them to a smaller number of overall divisions (and a correspondingly lower number of accumulat-

ed mutations) In contrast, the cells that produce a plant's reproductive parts arise from the same stem cells that produce stems, roots and leaves. These findings, published as a bioRxiv preprint awaiting peer review, add to a growing body of evidence that plants somehow shelter their stem cells from mutations, helping long-lived plants remain healthy as they age.

**REFERENCE:** Sarkar N. et al. Low Rate of Somatic Mutations in a Long-Lived Oak Tree. (2017) bioRxiv doi: 10.1101/149203.

In a step toward meeting that aim, researchers performed a genome-wide association study (GWAS) to discover regions of the common bean genome that affect agriculturally important traits. GWAS involves scanning markers across the entire genome of many individuals (in this case, individual common bean cultivars) to find genetic differences that could be connected with a particular physical trait. In this case, researchers compared the genetic makeup of 280 cultivated bean varieties grown in four locations across the United States. The analysis examined 150,000 DNA variants and identified candidate genes linked to six important crop traits: days to flower, days to maturity, growth habit, lodging (when the stalk or entire plant bends), canopy height (the amount of the plant growing aboveground), and seed weight. Overall, scientists in the project detected about 30 genetic regions potentially related to those traits, providing a framework for further study of common bean improvements.



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

**REFERENCE:** Moghaddam S.M. et al. Genome-Wide Association Study Identifies Candidate Loci Underlying Agronomic Traits in a Middle American Diversity Panel of Common Bean. The Plant Genome (2016) 9 doi: 10.3835/plantgenome2016.02.0012.



## What about the flavor?

Anyone who has eaten a fresh garden tomato knows: the flavor of a tomato from the grocery aisle can't compare. As commercial tomatoes have been bred for larger harvests, better disease resistance, and firmer skins to withstand shipping and storage, the complex flavor still found in heirloom varieties has been lost. Now, in an effort to bring deliciousness back to supermarket tomatoes, a group of researchers have identified the genetic factors that contribute to tasty tomatoes. They hope that this information might help tomato breeders recover the lost flavors.

Flavor is a complex mix of taste and smell. In tomatoes, that means a mix of acids and sugars contributing to taste and volatile chemicals contributing to smell. In this project, researchers examined flavor-associated chemicals in 398 modern, heirloom and wild tomato varieties. Consumer input helped identify which chemicals most contribute to flavor. Using whole genome sequencing and a genome-wide association study, the regions of the tomato genome that influence the amounts of those chemicals were identified. Scientists found that as breeders preferentially selected larger tomatoes, sugar content in the fruit decreased while acid content remained the same. They suggest that simply reducing tomato size will increase sugar content, which consumers prefer. While that solution seems simple – even obvious – tomato growers, who are paid by the pound rather than by the tomato, may be reluctant to begin growing tiny fruit, which are more expensive and time-consuming to pick.

On a positive note, sugar flavor can be enhanced by those volatile chemicals that influence aroma. In their analysis, the researchers identified genetic regions related to 25 aromatic compounds that contribute to tomato flavor. Out of those 25 compounds, 13 are reduced significantly in modern tomatoes. While reducing one of those chemicals wouldn't noticeably change the flavor, reducing half of the important flavor aromas has significantly altered – or arguably ruined – how commercial tomatoes taste. The researchers suggest that plant molecular breeders could use genetic markers associated with these aromatic compounds to breed commercial tomatoes that retain the flavorful alleles.

**REFERENCE:** Tieman D. et al. A chemical genetic roadmap to improved tomato flavor. Science (2017) 355:391-4 doi: 10.1126/science.aal1556 .





# 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.



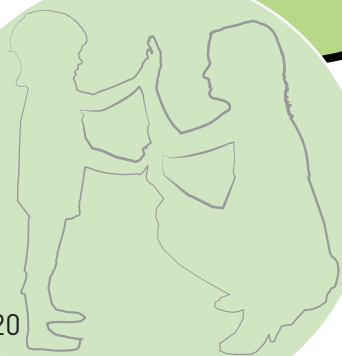
## Lifestyle:

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.

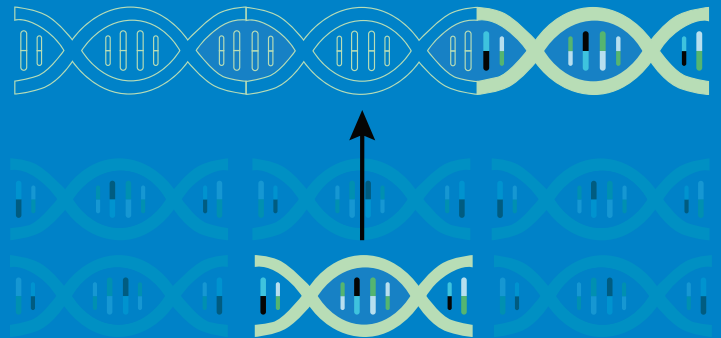
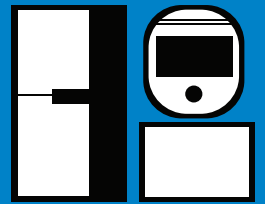
# How does a human genome get sequenced?

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 analyze the genome further.

## COURSE OF STUDY CONNECTED TO GUIDEBOOK TOPICS

Course	Objective and Applicable Subheading	Linking Scientific Concept
Biology	<b>1</b> 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 Analyses, Recombinant DNA and Genetic Engineering, Synthetic Biology, Pharmacogenomics
	<b>2</b> 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 6) RNA and Protein Analysis, Gene Therapy and RNAi, Stem Cells
	<b>3</b> Formulate an evidence-based explanation regarding how the composition of deoxyribonucleic acid (DNA) determines the structural organization of proteins.	DNA Sequencing, RNA and Protein Analyses, Bioinformatics, Application, Recombinant DNA and Genetic Engineering, Synthetic Biology, Therapeutic Approaches: Gene Therapy, Copy Number Variation, Personal Genome Analysis, Genome Editing
	<b>3a</b> 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, Genome Editing. See also The Progress of Science Timeline (pg 6)
	<b>3b</b> 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 Analyses, 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, Genome Editing
	<b>3c</b> 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
	<b>4</b> 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
	<b>10</b> 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]
	<b>11</b> 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
	<b>11a</b> 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
	<b>11b</b> 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
<b>Biology</b>	<b>11c</b> 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
	<b>12</b> 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
	<b>12a</b> 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
	<b>13a</b> Engage in argument to justify the grouping of viruses in a category separate from living things.	Infectious Disease
	<b>14</b> 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
	<b>15</b> 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
	<b>16</b> 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
<b>Anatomy and Physiology</b>	<b>3a</b> 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
	<b>6a</b> 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
	<b>6b</b> 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
	<b>9a</b> 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 threatened and endangered species.	Comparative Genomics, Recombinant DNA and Genetic Engineering, Agriculture
AP Biology <i>Big Idea 1 Evolution</i>	<b><i>The process of evolution drives the diversity and unity of life</i></b>  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 Analyses, Bioinformatics, Comparative Genomics
<i>Big Idea 2 Free Energy and Molecular Building Blocks</i>	<b><i>Biological systems utilize free energy and molecular building blocks to grow, to reproduce, and to maintain dynamic homeostasis</i></b>  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
<i>Big Idea 3 Information</i>	<b><i>Living systems store, retrieve, transmit and respond to information essential to life processes.</i></b>  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 Analyses, 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, Genome Editing
<i>Big Idea 4 Biological Systems</i>	<b><i>Biological systems interact and these systems and their interactions possess complex properties</i></b>  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	<b>5</b> 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
	<b>6</b> Discuss valid and essential information for the safe use of consumer goods and health products.	Agricultural, Cancer, Noninvasive Prenatal Diagnosis, Personal Genomic Analysis, Pharmacogenomics
	<b>10</b> Determine the causes of disability and premature loss of life across life stages.	Cancer, Identifying Genetic Influence on Disease
Technology Education	<b>26</b> Explain uses and advantages of databases.	Bioinformatics
	<b>27</b> Apply appropriate techniques for producing databases.	Bioinformatics
Agriscience	<b>10</b> Determine characteristics and functions of plants. Explain how agricultural crops can be utilized as alternative fuel sources.	Agricultural applications
Forensic and Criminal Investigations	<b>7</b> Describe presumptive and confirmatory forensic tests. <i>Examples: blood type comparison, DNA testing</i>	Criminal Justice and Forensics
	<b>8</b> 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	<b>10</b> Recognize legal responsibilities, limitations, and implications within the health care delivery setting. <i>Examples: Patients' Bill of Rights, legal documentation requirements, Health Insurance Portability and Accountability Act (HIPAA)</i>	Genetic Information Nondiscrimination Act, Personal Genome Analysis
Health Informatics	<b>5</b> Describe legal and ethical regulations as they relate to health informatics. <i>Examples: Patients' Bill of Rights, legal documentation requirements, Health Insurance Portability and Accountability Act (HIPAA)</i>	Genetic Information Nondiscrimination Act, Personal Genome Analysis
Introduction to Agriscience	<b>16</b> Analyze biotechnology to determine benefits to the agriculture industry. <i>Examples: improved productivity, medical advancements, environmental benefits</i>	Agricultural Applications, Bioinformatics, Recombinant DNA and Genetic Engineering, Genome Editing
Introduction to Pharmacy	<b>9</b> Identify classifications of selected drugs. Examples: analgesic, antibiotic, antiemetic	Personalized Medicine, Pharmacogenomics
	<b>11</b> Differentiate among drug interactions, drug reactions, and side effects.	Personalized Medicine, Pharmacogenomics
Introduction to Biotechnology	<b>1</b> Trace the history of biotechnology. Describing both scientific and non-scientific careers, roles, and responsibilities of individuals working in biotechnology.	Agricultural Applications, Bioinformatics, Criminal Justice and Forensics, Diagnosing Chromosome Disorders, DNA Sequencing, Pharmacogenomics, Genome Editing. See The Progress of Science (pg 6)
	<b>4</b> Correlate key cellular components to function.	See HudsonAlpha iCell® (pg 6)
	<b>5</b> Describe the process of meiosis and the cell cycle, including the hereditary significance of each.	Cancer, Diagnosing Chromosome Disorders, Noninvasive Prenatal Diagnosis, Stem Cells
	<b>8</b> 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. 26

## COURSE OF STUDY CONNECTED TO GUIDEBOOK TOPICS

Course	Objective and Applicable Subheading	Linking Scientific Concept
Introduction to Biotechnology	<b>9</b> Describe the structure and function of deoxyribonucleic acid (DNA), including replication, translation, and transcription.  Applying the genetic code to predict amino acid sequence  Describe methods cells use to regulate gene expression.  Defining the role of ribonucleic acid (RNA) in protein synthesis	Recombinant DNA and Genetic Engineering, Studying the Genome to Understand the Sequence  Bioinformatics  Cancer, Comparative Genomics, Epigenetics, RNA and Protein Analyses, Therapeutic Approaches  Recombinant DNA and Genetic Engineering, RNA and Protein Analyses, Therapeutic Approaches
	<b>11</b> Describe factors such as radiation, chemicals and chance.	Cancer, Infectious Disease, Genome Editing
	<b>13</b> Differentiate among major areas in modern biotechnology, including plant, animal, microbial, forensic, and marine.  Describing techniques used with recombinant DNA	Agricultural Applications, Bioinformatics, Criminal Justice and Forensics, DNA Sequencing, Infectious Disease  Agricultural Applications, DNA Sequencing, Synthetic Biology
	<b>14</b> Explain the development, purpose, findings, and applications of the Human Genome Project.  Analyzing results of the Human Genome project to predict ethical, social and legal implications  Describing medical uses of gene therapy, including vaccines and tissue and antibody engineering.  Using computer bioinformatics resources to provide information regarding DNA, protein, and human genetic diseases	Comparative Genomics, Copy Number Variation, DNA Sequencing, Identifying Genetic Influence in Disease, Personalized Medicine, Pharmacogenomics, Studying the Genome to Understand the Sequence  Criminal Justice and Forensics, Genetic Information Nondiscrimination Act, Personalized Genomic Analysis  Cancer, DNA Sequencing, Infectious Disease, Recombinant DNA and Genetic Engineering, RNA and Protein Analysis, Genome Editing  Bioinformatics, Cancer, Comparative Genomics, Copy Number Variation
	<b>15</b> Describe the replication of DNA and RNA viruses, including lytic and lysogenic cycle.	Infectious Disease
Plant Biotechnology	<b>1</b> Identify career opportunities associated with plant biotechnology.	Agricultural Applications
	<b>14</b> Describe the ecological and economic importance of plants.	Agricultural Applications
	<b>16</b> Explain the historical significance of plant biotechnology.	Agricultural Applications, Comparative Genomics; See also The Progress of Science Timeline (pg 6)
	<b>17</b> Describe methods of genetic engineering.	Agricultural Applications, Genome Editing





# FOUNDATIONAL CONCEPTS AND APPLICATIONS

## KEY TECHNOLOGIES

### DNA Sequencing

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

Today, the need for even greater sequencing capability at a more economical price has led to the development of new technologies based on different chemistries and refined for accuracy and speed. These “second generation” approaches reduce the necessary volume of reagents while dramatically increasing the number of simultaneous sequencing reactions in a single experiment. They are capable of producing nearly 150 times more sequence than the first generation systems, at 1/150th the cost. For example, the cost of sequencing all three 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](http://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 two to three 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.

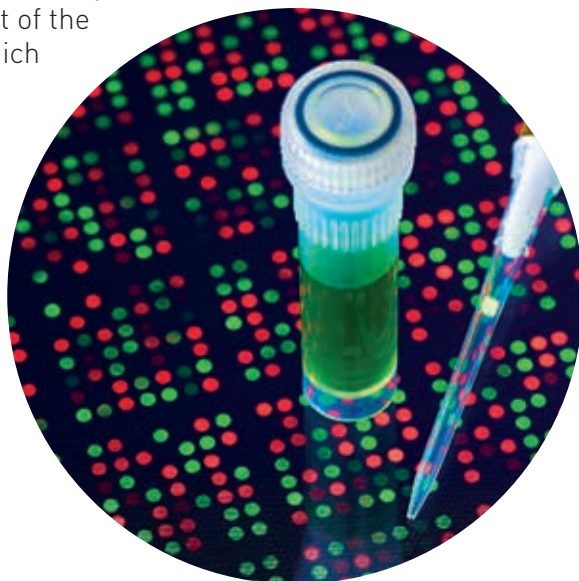
Second-generation sequencing technology has recently been extended to also identify RNA expression across cells. Scientists have shown that this approach, known as RNA-seq, yields more precise results than microarray analysis. It is expected that RNA-seq will become the standard tool for measuring genome-wide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

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.

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.

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





# KEY TECHNOLOGIES

## 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, one terabyte is equal to 1,000 gigabytes

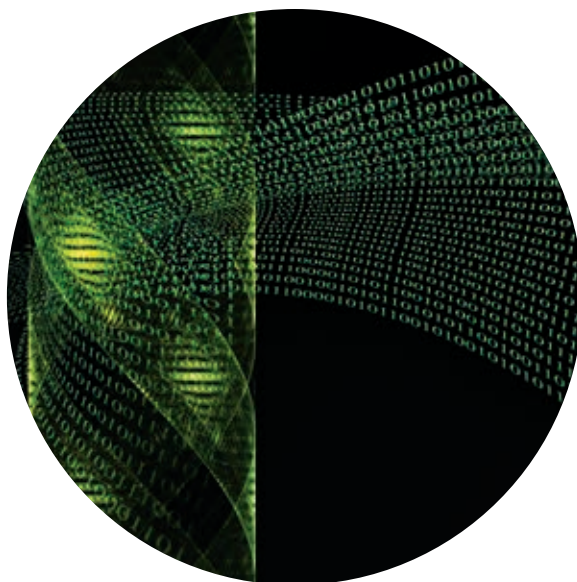
Understanding the meaning of all that information is a daunting challenge. Deciphering the data requires a biological knowledge of what to look for, algorithms (computer programs) capable of detecting interesting features and computers powerful enough to perform complex analyses efficiently and rapidly. 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 as ENCODE and the 1,000 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](http://www.hudsonalpha.org/available-educational-kits).



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.

### Genetically Modified (GM) Crops

More than 13 million farmers across 25 countries

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](mailto:edoutreach@hudsonalpha.org).

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.



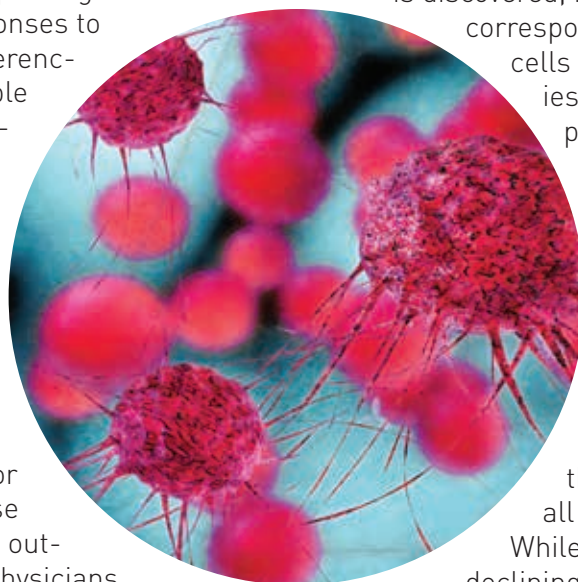
# APPLICATIONS

## 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 step-wise accumulation of additional 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.hudson-alpha.org/available-educational-kits](http://www.hudson-alpha.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.

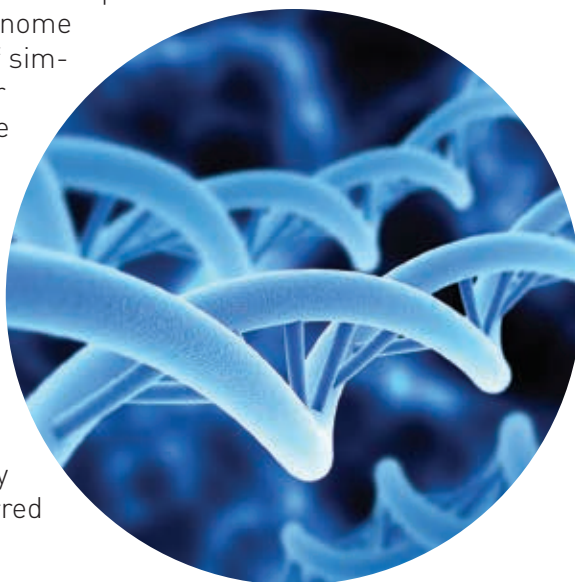
While humans may seem to have little in common with organisms such as fruit flies, roundworms or mice, they are all composed of cells that must take in nutrients and remove waste, interact with neighboring cells and the outside environment, and grow and divide in response to specific signals. To varying 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.



## APPLICATIONS

### 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 (1,000+ 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.

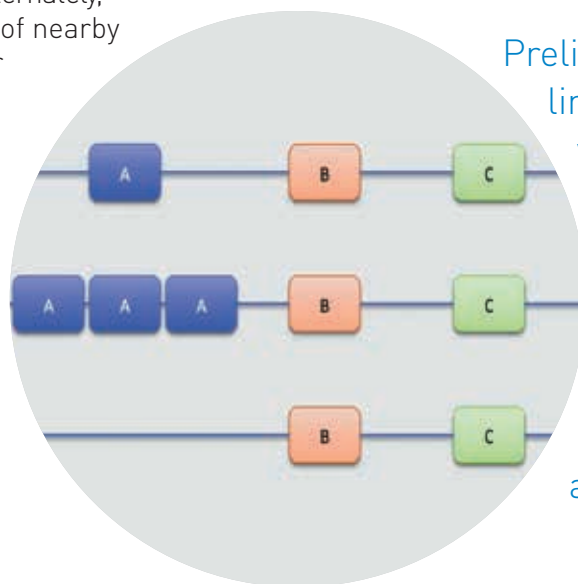
It came as a surprise to many scientists just how much DNA variation is due to copy number changes. Previous studies based primarily on SNPs suggested that any two randomly selected human genomes would differ by 0.1%. CNVs revise that estimate: the two genomes differ by at least 1.0%. While this may not seem like a major increase, remember that the human genome is composed of approximately three billion nucleotides, so the estimated number of nucleotides that vary between two random individuals has increased from three 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](mailto: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.

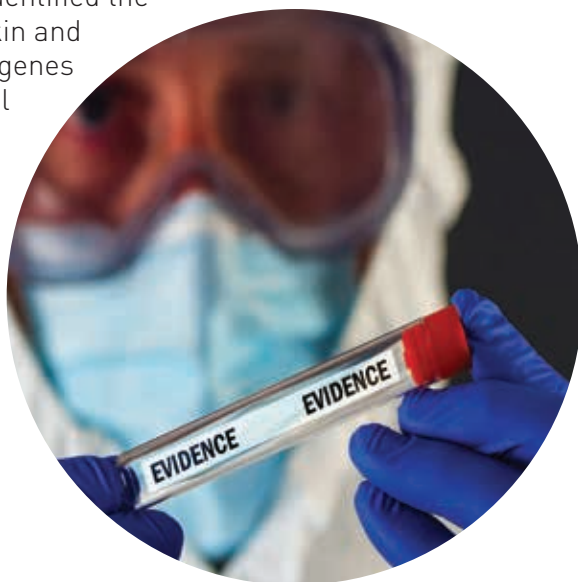
DNA analysis is also used to suggest ancestral origins; there are several companies offering Y-chromosome and mitochondrial DNA studies to determine, for example, to which of the ancient tribes of Britain a man belongs or whether a man or woman has African, Native American or Celtic DNA markers. It is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

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

DNA profiling is a critical component of the 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.

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.

This “forensic molecular photo fitting” may one day serve as a genetically-based police sketch




## APPLICATIONS

### 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 (aCGH), allow a finer level of resolution, with the ability to identify sub-microscopic chromosome changes.


Array CHG has replaced karyotyping as the standard chromosome diagnostic tool to detect abnormalities in chromosome number, microdeletions and other chromosome imbalances. It is used in both prenatal and postnatal settings. Depending on the specific array, it can detect chromosomal imbalances as small as 1,000 bp in size. The use of array CGH has significantly increased the diagnosis of chromosomal imbalances among individual with clinical anomalies.



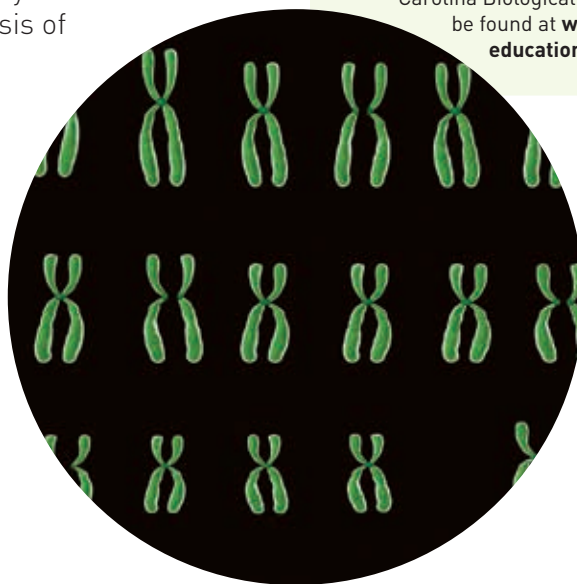
Chromosome studies, their behavior in cell division, the formation of egg and sperm and the concept of karyotyping are regularly discussed in Biology classes (Standards 4, 11, 12 and related sub-standards). Karyotyping to diagnose chromosomal disorders is examined in the Career/Tech course Intro to Biotechnology (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](http://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](http://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, pg 48).

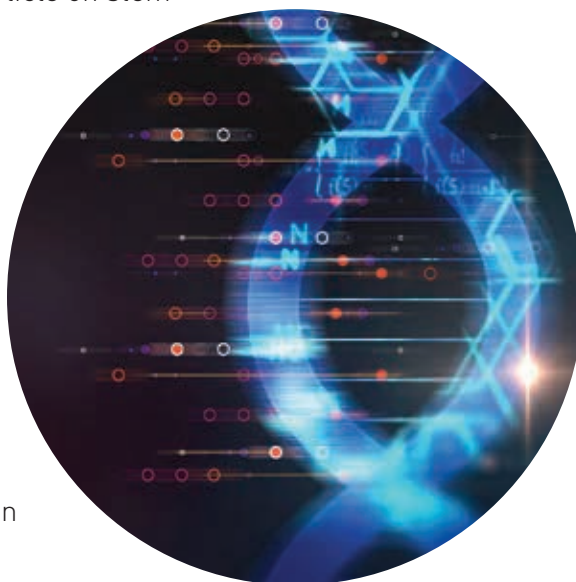
Studies of identical twins suggest that at birth, twins share similar patterns of epigenetic modification. As they age and are exposed to different diets and environments, the twins' patterns become markedly different, leading to altered activation and silencing patterns.

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 its 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.



## APPLICATIONS

### 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

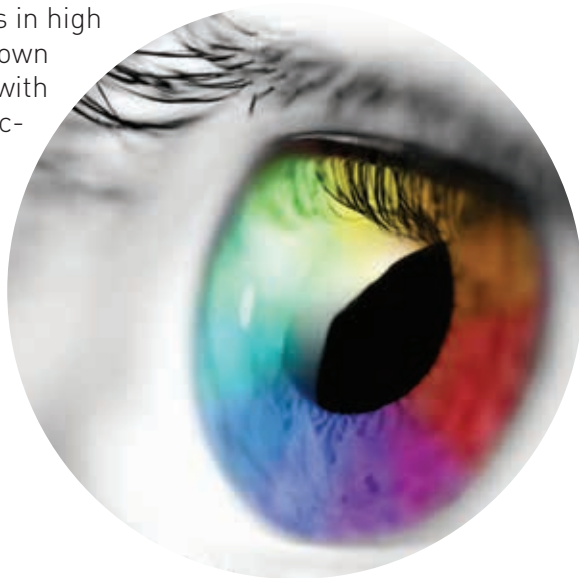
In 1907, Charles and Gertrude Davenport developed a model for the genetics of eye color. They suggested that brown eye color is dominant over blue eye color. This would mean that two blue-eyed parents would always produce blue-eyed children but never ones with brown eyes. For most of the past 100 years, 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% 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.



# APPLICATIONS

## Genome Editing

Tools that directly modify genetic sequences allow scientists to explore the functional impact of DNA mutations as well as engineer changes that create drug-producing bacteria, disease-resistant crops or life-saving genetic therapies. These approaches are analogous to the “find and replace” feature in word processing software – they scan the genome for a specific sequence and then make a targeted modification within that sequence.

All genome editing tools rely on some sort of programmable nuclease – an enzyme guided to a DNA sequence of interest in order to cut across the DNA strand. This cut triggers a DNA repair process that can knockout (disrupt) the genetic instructions or replace them with a different set of information. There are three programmable nucleases – zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and the clustered regularly interspaced short palindromic repeat Cas-9 system (CRISPR-Cas9). The CRISPR-Cas9 system has proven especially powerful and has quickly become a mainstay of genome editing protocols in microbes, plants and animals.

Genome editing tools have provided researchers with an unprecedented ability to modify cells. This leads to applications in basic research, agriculture and human disease therapy. However, programmable nucleases are not without their challenges. They can be difficult to correctly deliver to the appropriate cells, the frequency of DNA modification can be very low, and they occasionally cut DNA at “off target” sites. Techniques that increase specificity and efficiency have been developed to address these challenges, but additional tweaking is still needed.

Genome editing has shown immense promise – the first wave of modified crops and human therapies are moving from the laboratory to the field and clinic. A number of regulatory and safety hurdles must still be cleared, and there are many ethical, social and policy implications waiting to be addressed.

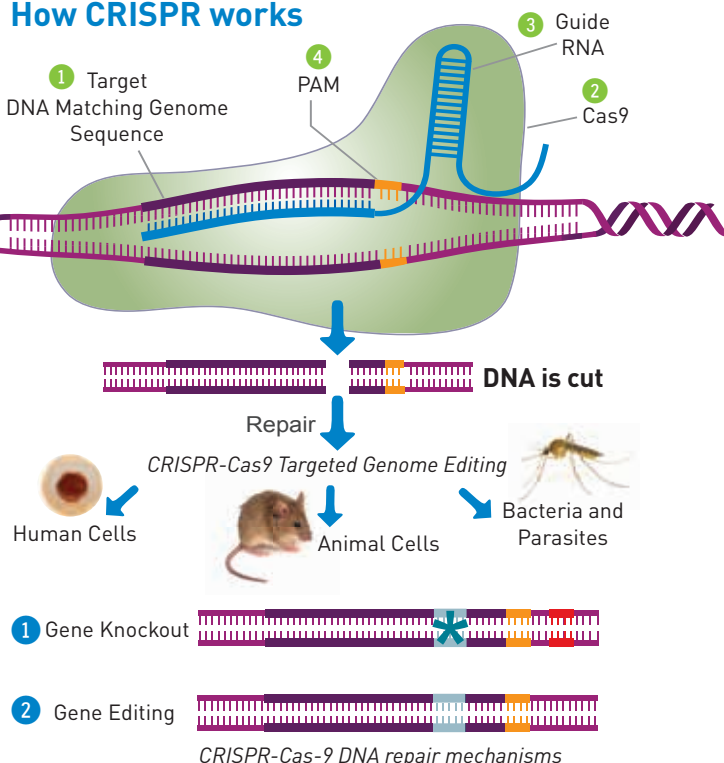
The CRISPR-Cas9 method has many applications in basic research, agriculture, drug development and eventually treating humans with genetic diseases.

The concept of genome editing can be explored as part of Biology standards 3, 3a and 3b, showcasing how a change in the DNA sequence impacts the corresponding protein and trait. This would also be an appropriate conversation in AP Biology (Enduring Understandings 3.B, 3.C and 3E). Lastly, the impact of genome editing has a place in classroom conversations for Introduction to Agriscience (standard 16), Introduction to Biotechnology (standards 1, 9, 13 and 14) and Plant Biotechnology (standard 17).

### There are four key components of a CRISPR-Cas9 system:

- 1. Target DNA:** This is the region of the genome to be modified.
- 2. Cas9:** This bacterial enzyme unzips and cuts the target DNA. To date, most approaches use the Cas9 protein found in the bacteria *Streptococcus pyogenes*.
- 3. 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.
- 4. 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.

### How CRISPR works





## Identifying Genetic Influence on Disease

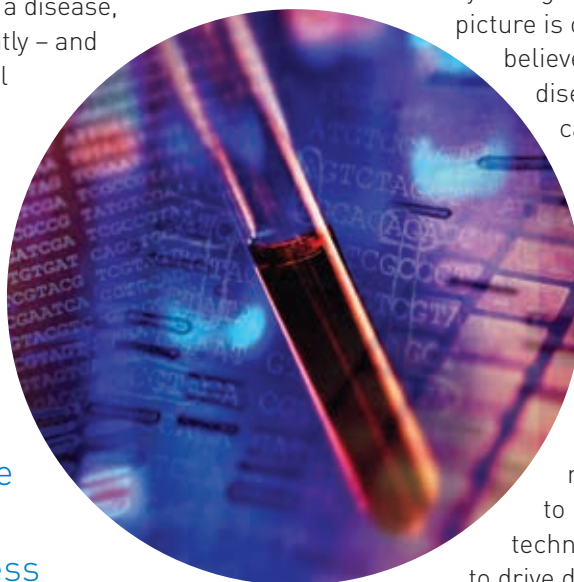
Much progress has been made in identifying the genetic causes of single gene diseases such as cystic fibrosis, phenylketonuria and Huntington disease. This has led to more accurate risk analysis, better testing approaches and, in some instances, more effective methods of treatment. Even though there 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](http://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 had 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.

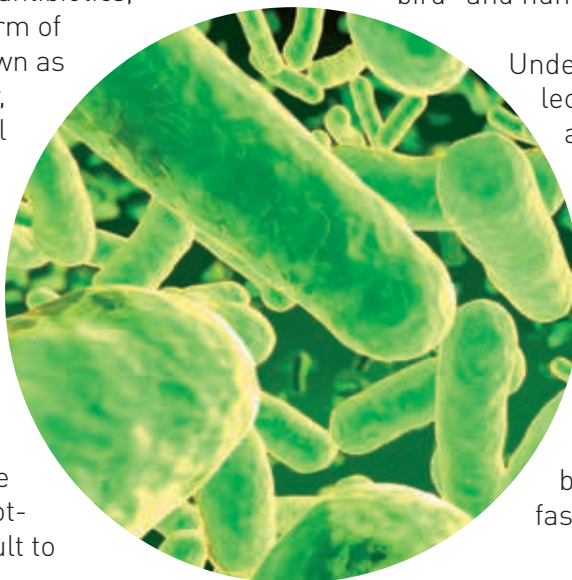
## APPLICATIONS

### 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 takeover 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 choose 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?





## APPLICATIONS

### 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

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 was 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.

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.





## 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 46).

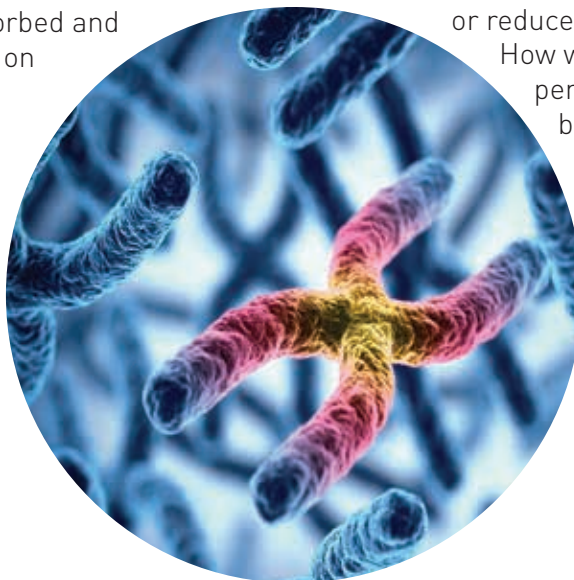
This newfound knowledge is rapidly moving into the clinical setting. At the forefront are a series of drugs such as Gleevec®, Herceptin® and Iressa® known to be most effective in people with a specific genetic profile (set of genetic variants).

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).

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.

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?



# APPLICATIONS

## 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 levels, leading to increased cell growth and tumor formation. In these cases, the anti-cancer drug Herceptin® is added to the patient's treatment plan where it increases the efficacy of chemotherapy.

Molecular testing is needed because only 25% of breast cancer patients will see any benefit from Herceptin – the rest should be given another treatment. In a similar manner, 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 31)
- modifying bacteria by adding genes that produce enzymes used in industry (Chymosin — used for making cheese)
- producing therapeutic products such as human insulin (Humulin®), blood clotting factors (rFVII) and components of the immune system (Enbrel®)
- developing biosensors to identify toxins in the water, soil or air



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

Silencing an overactive gene is a related form of therapy that at times utilizes recombinant DNA. More information about this approach, known as RNAi, can be found on pg 51.

## APPLICATIONS

### Stem Cells

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

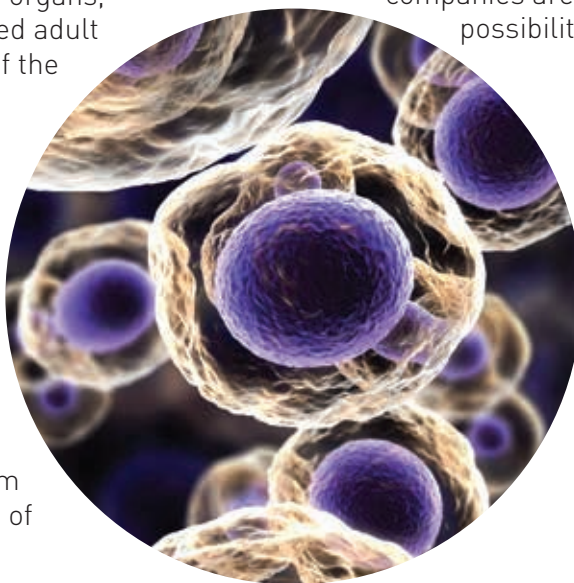
Multiple types of stem cells have been identified or developed. Embryonic stem cells (ES cells) were the first category discovered. These cells are fully pluripotent, but only found in young embryos (the stage of human development from conception to eight weeks gestation). Because the process to collect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

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).

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.



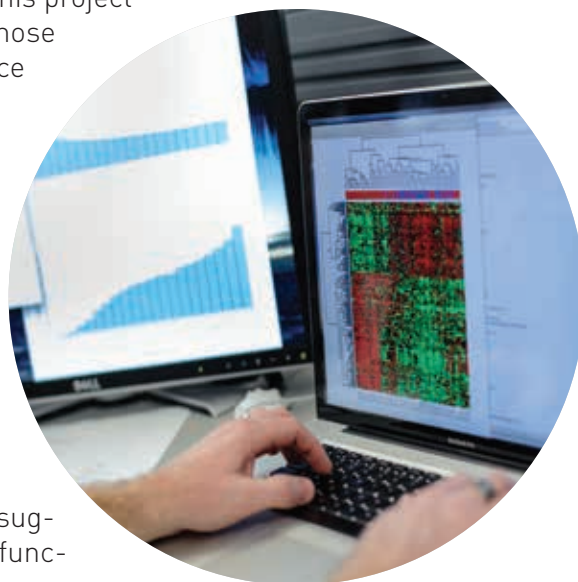


## Studying the Genome to Understand the Sequence

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.



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 non-coding 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 recently revised their Exploring DNA Extraction lab available through AMSTI Science in Motion. Students choose from a variety of plant and animal samples and common household reagents. Given their knowledge of DNA and the biochemical nature of cell structures, they are tasked with creating a lysis buffer that maximizes the amount of DNA extracted. Students explain how their lysis solution impacts the extraction process through text, diagrams or models.

This lab could be incorporated in multiple places throughout a curriculum and supports a deeper grasp of the structure of DNA. It is a useful activity prior to discussing recent advances such as HapMap, ENCODE or the 1,000 Genomes project. The freely available Progress of Science timeline ([timeline.hudsonalpha.org](http://timeline.hudsonalpha.org)) provides information about these and other discoveries associated with the development of genomic sciences.

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.

## APPLICATIONS

### 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

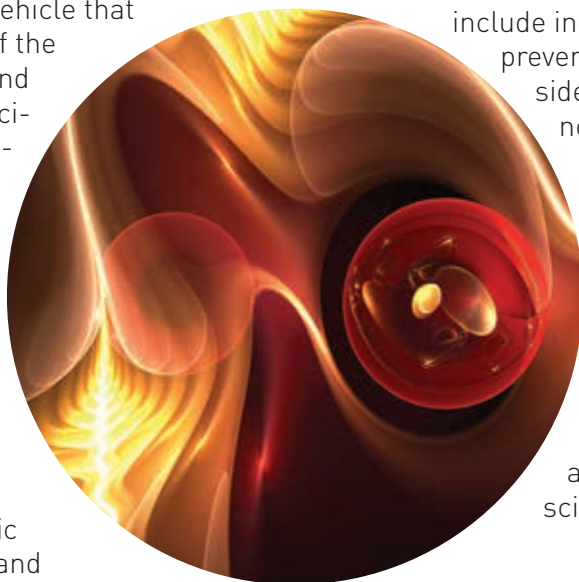
The concepts behind synthetic biology link to Biology COS standard 3, particularly as it relates to understanding of the elements of DNA code necessary for biological function. Discussion of synthetic biology also connects to AP Biology (Enduring Understanding 3.A and 3.E) and the CTE Intro to Biotechnology (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.

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 uncharted scientific territory.

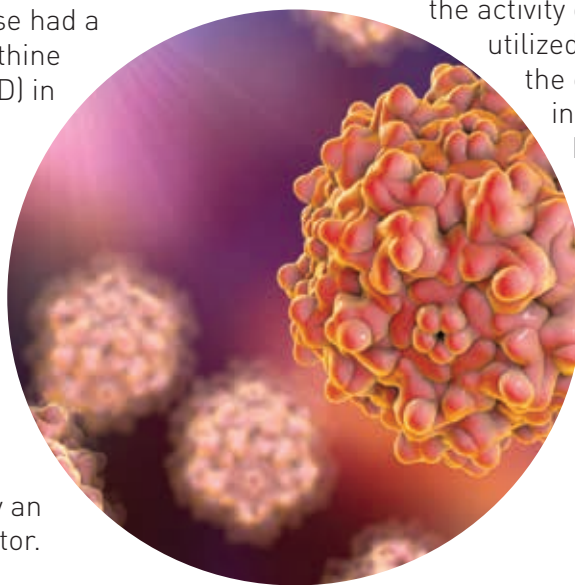


## Gene Therapy

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 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.

## 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 References (pg 9)

1. Edwards, Monica and Abadie, Laurie. "How Stressful Will a Trip to Mars Be on the Human Body? We Now Have a Peek Into What the NASA Twins Study Will Reveal" NASA.gov. N.pg., 30 July 2017. Web. 07 July 2017.
2. Liu B. et al. IL-33/ST2 signaling excites sensory neurons and mediates itch response in a mouse model of poison ivy contact allergy. *PNAS* [2016] 113:E7572-9 doi: 10.1073/pnas.1606608113.
3. US Food and Drug Administration. "FDA Allows Marketing of First Direct-to-consumer Tests That Provide Genetic Risk Information for Certain Conditions." N.pg., 6 Apr. 2017. Web. 9 July 2017.
4. Marouli E. et al. Rare and low-frequency coding variants alter human adult height. *Nature* [2017] 542:186-90 doi: 10.1038/nature21039.
5. Corey D.R. Nusinersen, an antisense oligonucleotide drug for spinal muscular atrophy. *Nature Neuroscience* [2017] 20:497-9 doi: 10.1038/nn.4508.
6. Liscovitch-Brauer N. et al. Trade-off between Transcriptome Plasticity and Genome Evolution in Cephalopods. *Cell* [2017] 169:191-202 doi: 10.1016/j.cell.2017.03.025.
7. Markt S.C. et al. Sniffing out significant "Pee values": genome wide association study of asparagus anosmia. *BMJ* [2016] 355:i6071 doi: 10.1136/bmj.i6071.
8. Mallarino R. et al. Developmental mechanisms of stripe patterns in rodents. *Nature* [2016] 539:518-23 doi: 10.1038/nature20109.
9. Main, B.J. et al. The Genetic Basis of Host Preference and Resting Behavior in the Major African Malaria Vector, *Anopheles arabiensis*. *PLOS Genetics* [2016] 12:e10063032 doi: 10.1371/journal.pgen.1006303.
10. Kuhnemund M. et al. Targeted DNA sequencing and in situ mutation analysis using mobile phone microscopy. *Nature Communications* [2017] 8:13913 doi: 10.1038/ncomms13913.

### Image Credits

Unless otherwise stated, images were purchased or provided by: Adobe® Stock, Pixabay®, Pexels®, Shutterstock®, Unsplash, RightsLink and HudsonAlpha Institute for Biotechnology

pg 9 J. F. Broekhuis

pg 10 Chromosomal Neighborhoods - Krijger P and de Laat W. *Nature Reviews Molecular Cell Biology*. 2016 doi:10.1038/nrm.2016.138.

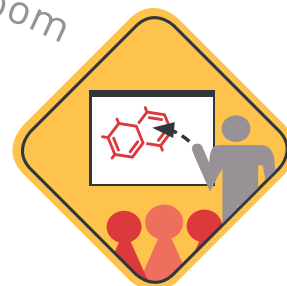
pg 10 Replication Fork - Graham J.E. et al. *Cell* 2017 doi:10.1016/j.cell.2017.05.041.



MIDDLE SCHOOL

# GPS

navigating genetics content in your classroom



## Genetics Performance Standards For Middle School Teachers

This series of one-day workshops emphasizes the genetics components of the Alabama 7th grade Life Science course of study, specifically standards 12, 13 and 14. HudsonAlpha educators will lead teachers through a full day of utilizing models, cutting edge genomics research and pedagogical strategies to bring real-world genetics content to their classrooms.

**Teachers will leave the workshop with 8 hours of professional learning, digital resources and kits to take back to their classroom.**

**Huntsville - September 19, 2017**

**Mobile - October 12, 2017**

**Auburn - January 30, 2018**

**Birmingham Area - February 21, 2018**

**Register for FREE to attend  
Middle School GPS at**

**[www.hudsonalpha.org/middle-school-gps](http://www.hudsonalpha.org/middle-school-gps)**

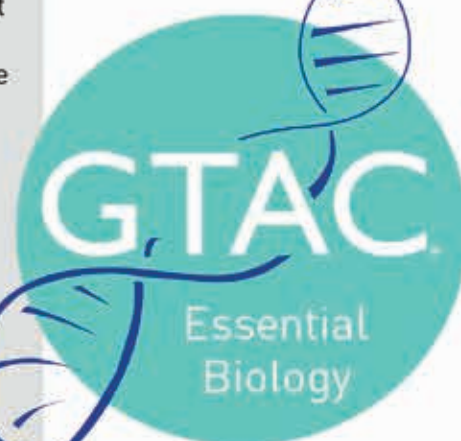


## Genetic Technologies for Alabama Classrooms:

**GTAC: Essential Biology** is a yearlong professional development experience focused on the genetic content found in a general Biology course. The program begins with a one-week intensive professional development academy held at HudsonAlpha and continues through the following school year with additional learning opportunities, planning and content support. Teachers Receive:

- 40+ Professional Learning Units
- Toolkit of equipment and resources
- Updated genetics content knowledge
- Stipend, upon completion of post-workshop deliverables

**Registration opens January each year.**



**GTAC: Advanced Concepts** is a week long professional development academy focused on the applications of the science concepts encountered in an advanced life science course such as AP or IB Biology. Selections for this competitive application based program are based on courses taught and educators' previous experiences.

- Hear from scientists involved in cutting edge genomic research.
- Use modern biotech equipment and laboratories.
- Develop classroom plans and have support implementing.

**Applications open December each year.**



601 Genome Way Huntsville, AL 35806 | 256.327.0462 | [www.hudsonalpha.org/GTAC](http://www.hudsonalpha.org/GTAC)

## Contact

### Neil E. Lamb Ph.D.

Vice President for Educational Outreach  
HudsonAlpha Institute for Biotechnology  
601 Genome Way  
Huntsville, AL 35806  
[nlamb@hudsonalpha.org](mailto:nlamb@hudsonalpha.org)  
[hudsonalpha.org](http://hudsonalpha.org)

