Science Highlights

November 2018 | by Ann A. Kiessling, PhD at the BEDFORD RESEARCH FOUNDATION

GENE EDITING OF HUMAN EMBRYOS

What is "gene-editing?"

Is gene editing human embryos a positive scientific breakthrough for human health?

Or misuse of a powerful research tool?

This "Highlights" will outline the basic biology behind gene editing, followed by a description of the process in general and in human embryos, specifically.

The Basics

The genetic information of humans, collectively termed the "human genome," is contained within 22

Human genome: ***All of the genetic information needed for the embryonic development and adult function of a human being.***

chromosomes plus either 2 "X" chromosomes in girls, or 1 "X" chromosome and 1 "Y" chromosome in boys. Each chromosome is two long strings of four deoxyribonucleic acid (DNA) units (Adenosine, Cytosine, Guanosine, Thymidine; A, C, G, T) attached to each other in a sequence specific for

Chromosome: ***a long string of genes attached end to end and then folded with proteins in a specific way.***

that gene. The two long strings are held together by attractions between the units, i.e., A in one string is attracted to T in the opposite string.

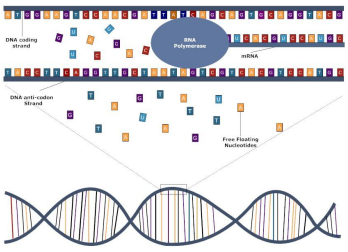
Each species has its own number of chromosomes, e.g. the genome of the laboratory mouse is divided among 20 chromosomes, even though the total amount of DNA is the same in each mouse and human cell, approximately 5 picograms. An onion also has 20 chromosomes, but they are an order of magnitude larger than human or mouse, with approximately 50 picograms of DNA per cell. So the amount of genetic information does not correspond to the complexity of the organism.

Importantly, there are two copies of each chromosome present in all cells (total of 46), except for sperm, which have only one copy. This becomes an important fact for gene editing.

Gene: ***A specific sequence of A, C, G, T units that instruct the sequence of amino acids that comprise a specific protein. Humans have 20- to 25 thousand genes***

The specific sequence of A, C, G, T units is interpreted in two steps. First by assembling a copy of the gene sequence to serve as a template, and secondly by stringing amino acids together to generate the sequence specified by the template. The process is an engineering marvel that takes place billions of times every day in cells throughout the body. The gene's code for a specific protein is the order of combinations of three of the four A, C, G, T units specific for each amino acid.

Picture a train moving down a track. The two rails for the wheels are the two strands of DNA comprised of series of A, C, G, T units attached to each other and attracted to the opposite strand for stability. As the engine wheels pass by the units, the sequence for the emerging single stranded template is "read" in the car behind that contains a stockpile of A, C, G, U ribonucleic acid (RNA) units that are strung together in the same sequence as the gene. Behind the train car synthesizing the template is another car full of amino acids and the enzymes that string them together to create the protein specified by the gene.



The accuracy of the cellular machinery to "translate" gene sequences into the amino acid sequences for functioning protein molecules is both extraordinary and essential for normal cell functions.

For example, consider the gene that codes for a protein responsible for tissue rejection, beta 2 microglobulin, **B2M**. It is part of chromosome 15. The single stranded RNA template copy of the gene is 1675 units long and the protein it codes for contains 119 amino acids. The B2M protein begins with the string of eight amino acids linked together like beads: Methionine- serine- arginine- serine- valine- alanine- leucine- alanine… which corresponds to gene sequence ATG TCT CGC TCC GTG GCC TTA GCT… This example also illustrates that each amino acid can have more than one triplet code, e.g. the triplet codes for serine are both TCT and TCC. This provides an important buffer for the specified amino acid should a T- to C- mutation occur in one of the codes for serine. And it also illustrates why everyone's DNA sequences are not identical.

Moreover, to illustrate the importance of faithful replication of each chromosome every time a cell multiplies, if one unit were lost in the middle of the above sequence for B2M, e.g. a C, a "frame shift" would occur, and the sequence would become ATG TCT CGC TCG TGG CCT TAG… This sequence would code for methionine- serine- arginine- serine- tryptophan- proline- followed by the "stop" triplet, TAG. Hence, no B2M protein would be synthesized as a result of a deletion of a single C.

Not only is translating the code into specific proteins essential to normal cell function, creating accurate and complete copies of all 46 chromosomes each and every time a cell multiplies is also essential to normal body functions.

Protein translation: The process of stringing together amino acids according to the sequence of A, C, G, T units in the gene

Because each amino acid is specified by a triplet sequence, if even one A or C or G or T is accidentally eliminated during replication of the chromosome, it would result in a "frame shift" in the triplet codes, as described in the B2M example.

This possibility is thought to be the reason for the second copy of each chromosome — as insurance that at least one copy of each gene will be available for the cell to use for essential processes. This is not the case for the X and Y chromosomes in males, which is why the disease hemophilia occurs in men. The genes that code for the proteins responsible for blood to clot following an injury are on the X chromosome. There is no back-up in men for mutations in X-chromosome genes, so such mutations result in loss of key blood clotting factors, hence hemophilia.

But there are extensive "gene repair" systems in every cell to correct mutations as they occur. Most mutations are probably due to the complexity of the enzyme systems themselves, others result from the relentless bombardment from gamma rays experienced by everything on earth. Every change in a gene sequence can be termed "gene editing," whether or not it is repaired.

Naturally Occurring Gene Edits

No two individuals have exactly the same gene sequences because multiple sequences code for the same amino acid. This is the basis for DNA tests to prove paternity or predict ancestry. Most of the gene variations do not change the proteins they code for, but some do, such as genes for eye and hair color and height.

Gene edit: A modification of a specific sequence of A, C, G, T units that instruct the sequence of amino acids that comprise a specific protein. The edit may or may not alter the amino acid sequence and the protein.

Therefore, fertilization of an egg, pollination of a flower, introduce gene edits in the offspring because of variations in the gene sequences of the two cells uniting.

Still other gene edits occur because of "transposable elements," first described in corn by Barbara McClintock (1), Nobel Laureate in 1983. Such "transposable elements" are common in all life forms, approximately 45% of the human genome is transposable elements and their location in individual genomes is highly variable.

The most well-studied gene edits in humans are those that cause cancer, such as the breast cancer gene, BRCA, on chromosome 13. It codes an important enzyme in DNA repair. A mutation that results in a "frame shift," as described above, results in no BRCA protein expression. Hence, its function to repair spontaneously occurring DNA mutations is inhibited, resulting in cells containing mutated DNA that lack the controls that limit cell multiplication, leading to uncontrolled cell expansion, the definition of cancer.

A more recently studied naturally occurring gene edit is the 32 gene unit deletion in CCR5 on chromosome 3. The mutation results in loss of CCR5 protein on the surface of HIV target cells,

CCR5: A member of the C-C chemokine receptor family that codes for the docking protein for the HIV virus on the surface of HIV target cells.

rendering them resistant to HIV attachment and infection. This mutation naturally occurs in approximately 1.5% of humans.

Gene Edits for Research

Early gene editing experiments were accomplished by mating individuals with different traits. Two well known examples are Mendel's famous red peas crossed to white peas to yield pink peas, and Mr. Little's Fancy Mice, popular in the early 1900's, bred for coat color, formed the basis of the Jackson Laboratory's inbred mice to study genetic diseases.

Nobel Laureate Mario Capecchi (link to his AES talk on the BRF site) systematically studied the function of mouse genes by mutating them into silence, so called "knock-out" mice. This was accomplished by flooding cultures of mouse embryonic stem cells with strands of synthetic DNA that could replace the

Knock-out mouse: A form of gene editing that resulted in mutation of specific genes to silence them in order to discover their importance to normal functions, such as fetal development in the mouse.

normal gene with an edited copy during DNA replication. The edited gene sequence was designed to not guide the synthesis of the normal protein. Such gene edited cells were combined with early mouse embryos, ultimately becoming part of the tissues of the mouse, including occasionally sperm and eggs. Males with gene edited sperm were mated to females with gene edited eggs to produce offspring containing two copies of the edited, non-functioning genes. Although laborious and time-consuming, this approach has yielded highly valuable information about the normal functions of thousands of genes.

In the past 20 years, other less time consuming methods of silencing genes, or increasing their expression, have been developed, all with the goal of understanding their function in health and disease.

In 2013, the most recent method for gene editing was popularized by scientists at Stanford and MIT. It is an adaptation of a naturally occurring defense mechanism that bacteria have against the viruses that invade them. Termed CRISPR/Cas, it is a complex between a protein that can cut DNA strands and a synthetic single-stranded RNA with a sequence of A, C, G, U that matches the gene being

CRISPR/Cas: "Clustered Regularly Interspaced Short Palindromic Repeats" is a term that describes DNA sequences in the viruses that infect bacteria. The immune system of bacteria includes a family of proteins (CRISPR-associated, Cas) that recognize CRISPR sequences and degrades them. The enzyme, Cas, needs to bind to a specific RNA sequence of 120 units, which can be synthesized synthetically, in order to degrade the DNA. These two components also function well in cell types other than bacteria, and so have become a useful tool for cutting DNA, resulting in either small deletions, or successful insertions of new synthetic DNAs. Both outcomes create an edited (mutated) gene.

targeted. The simplicity and specificity of the system have rapidly led to a wide variety of applications among scientists world-wide.

Such targeted DNA cuts can edit the gene sequences so they no longer code for a functioning protein, analogous to the natural CCR5 mutation, or opening the DNA strands can allow the incorporation

of synthetic DNA sequences into the cut site. This raises the exciting possibility of being able to repair defective human genes.

Gene Edits for Treatment of Disease

Most scientists have applied the CRISPR/Cas system to specific tissues or to stem cells. For example, it is theoretically possible to repair the X-chromosome mutations in liver cells so normal blood clotting factors can be produced by the liver (2).

Bedford Research scientists are applying the technology to edit B2M gene sequences in unfertilized eggs which are subsequently activated for stem cell derivation (link to article on website).

But more recently other scientists have applied CRISPR/Cas technology to human embryos. Last year a Portland Oregon research team reported their efforts to repair a mutation in the gene MYBPC3 known to be associated with acute heart failure in young men (3). The 30-member team created embryos with sperm from a man carrying the mutated gene in half of his sperm. (It is important to note that this experiment is not possible in Massachusetts because the stem cell bill (MGLc 111L) specifically prohibits the creation of embryos for research purposes only.)

At the time of fertilization of eggs with the mutant sperm, the Oregon scientists also injected the CRISPR/Cas agents designed to home to the gene mutation and insert "normal" DNA sequences. They reported the repair was successful in some embryos, but not all. Other research teams in New York and Australia replied to the report with their own interpretations of the results and all groups agreed much more work is needed to understand how to reliably edit genes in early human embryos.

Gene Edits for Enhancement

Earlier this week, a Chinese scientist reported the birth of twin girls whose genomes had been modified to silence the CCR5 gene (4). The birth was reported to be one of a series of human embryo experiments designed to render the offspring resistant to infection by HIV and to prove the principal that gene editing was possible — and perhaps beneficial— in human embryos. The work was not reported in a scientific format, so few scientists have had the opportunity to review the data in detail.

Several ethical concerns with this report, <u>if true</u>, have been raised. The gene editing was not performed to correct a known, serious medical issue in the embryos. It was performed to enhance resistance to HIV. A highly controversial idea.

But a more practical problem with the work is the possibility of "off-target" gene edits. Much research has been devoted to discover, and eliminate, the random edits that may occur at other than the gene locations being specifically targeted by the CRISPR/Cas reagents. It is these potentially deleterious unintended consequences that must be addressed in order to protect the offspring produced.

The Future of Gene Editing

Gene editing is common in nature, and forms a basis for evolution itself. Some edits are positive, such as the mutation that leads to resistance to HIV infection; other edits are negative such as the ones causing hemophilia and heart failure.

Humans have been gene editing plants, animals, and each other by cross-breeding for millennia. Technology companies like Monsanto have harnessed natural plant gene editing systems to modify plant genes such as rice to produce Vitamin A and corn to resist infestation by worms. Termed "genetically modified organisms" such plants have caused heated world-wide debate, despite the fact that cross-breeding plants modifies many more genes in one breeding cycle than the Monsanto technology.

Nonetheless, every new technology that can have effects in subsequent generations — of plants, animals, or microorganisms — needs to be carefully reviewed and implemented in stages that can be monitored for unintended, deleterious consequences.

We cannot put this genie back in the bottle, but with reasoned approaches, humans can optimize the benefits and mitigate the dangers posed by gene editing.

Ann A Kiessling, PhD

(1) Carnegie Institute, 1948;
(2)hemophilianewstoday.com/crispr-cas9-hemophilia
(3) Nature, 2017, 548:413; (4) https://www.technologyreview.com/s/612472/rogue-chinese-crispr-scientist-cited-us-report-as-his-green-light/