The Human Genome Project

The Human Genome Project resulted in the first draft sequence of the 3.2 billion base pair human genome in 2001. That first genome cost approximately 4 billion dollars. Today, NextGen DNA sequencing technologies make it possible to sequence a human genome for ~$1,000.

The ability to easily and quickly sequence whole genomes has lead to many new insights into the molecular basis of human diseases – and is paving the way to an era of personalized medicine. But in addition to focusing on uncovering the genetic basis of an individual’s disease, researchers are now beginning to develop technologies that will be used to repair specific genes.

How do we edit the human genome? ... by directing various DNA binding proteins to specific DNA sequences – where they then cut the DNA using nuclease domains that have been fused to their DNA binding domains. Three different classes of DNA binding proteins are being engineered for use in this way: Zinc Finger Nucleases, TALENS, and Cas9 (CRISPR associated protein 9).

Inactivating the CCR5 gene as a functional cure of HIV/AIDS.

The CCR5 protein is an HIV co-receptor. It cooperates with the cellular CD4 protein to allow the initial docking of the HIV virus onto T-cells, and subsequent infection. Curiously, approximately 15-20% of the northern European population are homozygous for this 32 bp deletion in their CCR5 gene – making them resistant to HIV infection. Approximately 1% of this population is homozygous for this mutation – and resistant to HIV infection.

Sangamo Biosciences have developed a zinc finger nuclease that is targeted to the naturally occurring 32 bp deletion in their CCR5 gene. As a result, this deletion site, which is located midway through the gene, changes the translation reading frame. Therefore, regardless of which protein you use to edit the human genome, it must be able to bind specifically to a unique 16 bp sequence of DNA.

Inactivating the CCR5 gene as a functional cure of HIV/AIDS.

The Berlin Patient

Tim Brown is believed to be the only person who has been functionally cured of HIV. After having controlled his HIV for many years with antiretroviral therapy, Tim was “lucky” in that he developed acute myeloid leukemia (AML) – requiring bone marrow stem cell transplantation. The donor for his transplantation was also homozygous for a naturally occurring 32 bp deletion in their CCR5 gene. As a result, this transplantation not only cured Tim’s AML... it also cured his HIV. The transplanted T-cells were resistant to HIV infection due to the missing CCR5 protein. Four years after the transplant, Tim remains free of both cancer and HIV.

So, why don’t we use a zinc finger nuclease to target the CCR5 gene in HIV patients – with the goal of inactivating it and making the patient’s endogenous T-cells resistant to further infection? That is exactly what is being done in a Phase 2 clinical trial conducted by a group from the University of Pennsylvania School of Medicine, the Albert Einstein College of Medicine and Sangamo Biosciences (a biotech company specializing in the development of therapeutic zinc finger nucleases).

The CCR5 Gene with the 32 bp Deletion

The CCR5 Protein

CCR5 is a chemokine receptor – that also functions as a co-receptor for HIV entry. CCR5 resembles a G-protein-coupled receptor, made up of seven transmembrane alpha helices. The structure of CCR5 has only recently been solved – as a complex with maraviroc – to bind a zinc atom to stabilize the protein in a conformation that is not able to interact with the HIV gp120 protein.

The eleven amino acids encoded by the 32 bp deletion is indicated on the model to the left. This deletion site, which is located midway through the gene, changes the translation reading frame. Therefore, the protein product translated from the gene containing this deletion is truncated – as a result of the out-of-frame STOP codon encountered 31 codons after the deletion site.

Zinc Finger Proteins and Zinc Finger Nucleases

Zinc Finger Nucleases are sequence specific DNA binding proteins. Each finger is composed of a short alpha helix and a 2-stranded beta sheet. Two helidines from the helix and two cysteines from the beta sheet coordinate bind a zinc atom to stabilize this protein motif. Each finger recognizes and binds to three consecutive base pairs in double-stranded DNA.

By linking 6 zinc fingers together, we could target a unique 18 bp sequence of DNA. But most natural zinc finger DNA-binding proteins have only 3 consecutive fingers. Can you guess why?

To target a unique site in the human genome, researchers have created two different 3-finger proteins – each one targeting a different 9-bp sequence, separated by 5 bp. Each 3-finger protein is equipped with one half of a FokI nuclease domain. The FokI only functions as a homodimer. So, when both 3-fingered proteins bind to their targets, the functionally active nuclease homodimer forms and makes a double-stranded cut in the DNA. Once the cut is made, an error-prone DNA repair system will try to repair the damage - and introduce mutations into the genome at that site.

How long is a statistically unique site in the human genome?

If you want to direct a nuclease to a unique DNA sequence on the human genome, the nucleotide sequence must be long enough to be statistically unique. You can easily calculate how long this sequence must be by asking yourself:

How many possible 1 base sequences are there? Answer: 4
How many possible 2 base sequences are there? Answer: 4 x 4 = 16
How many possible 3 base sequences are there? Answer: 4 x 4 x 4 = 64
How many possible 4 base sequences are there? Answer: 4 x 4 x 4 x 4 = 256
How many possible 5 base sequences are there? Answer: 4 x 4 x 4 x 4 x 4 = 1024

Continue this line of reasoning until the number of possible sequences exceeds the number of nucleotides in the human genome ~ 3.2 billion. The answer is 16 nucleotides.

What could possibly go wrong?

Life is complicated – especially when we are talking about editing a specific gene in the cells of a patient. Some complications to worry about include:

1. How do we get the ZFN into the cells?
2. What about cuts at “off-site targets”???
3. What fraction of T-cells must be modified in order to have a therapeutic effect?

References and Resources

4. The CBM's Science Olympiad CCR5 resources: http://cbm.msoe.edu/scienceOlympiad/module2015/