

# DNAdots™

Simple explanations  
of modern genetic technologies

minipcr®

© 2017 Amplyus, all rights reserved

# The Human Genome Project



## What it is:

The initiative that sequenced the entire human genome

The *Human Genome Project (HGP)* is widely recognized as a tremendous success of government initiative and international collaboration. Launched by the United States government in 1990 with the goal of sequencing the entire human genome, the scale of this project was enormous; it included the sequencing of 3 billion DNA base pairs and 20,000-25,000 human genes. With a \$3 billion budget and occurring simultaneously at 20 institutions across the world, when it was first announced, the Human Genome Project was met with some skepticism by people who thought this publicly funded project was technically beyond reach and financially too costly. By 2000, the HGP announced it had completed a draft of the entire human genome at 99.99% accuracy – five years ahead of schedule.

The HGP was mainly completed using *BAC-based sequencing*. BAC-based sequencing involves generating bacterial artificial chromosomes that carry long stretches of yet unknown sequences in the human genome. The sequenced regions included both coding and noncoding regions, but it did not include the centromeres (center of chromosomes that hold them together) or telomeres (caps at the ends of chromosomes), neither of which contain DNA that codes for protein. Once those chromosomes were sequenced, it fell upon computer scientists to analyze the sequences, arrange them in order, and locate the regions that code for protein.

## The debate:

Many argue that the HGP was completed so efficiently in part because of competition from a private company, the Celera Corporation. In 1998, Celera, run by J. Craig Venter, announced a competing project to sequence the human genome using \$300 million in private funding. The major difference in philosophy of these projects was that Celera wanted to keep the sequence somewhat private, requiring payment to access the sequence and patenting certain genes before the sequence was released. In contrast, the HGP published all sequences allowing for free and open use as soon as they were complete. In fact, the Human Genome Project's Bermuda Agreement, laid down in 1996, ensured that all information from the HGP would be made freely available to all within 24 hours.

This difference made any collaboration between the two projects impossible. Instead, they competed, and in 2000 the race ended in a virtual tie. Both Celera and the HGP were able to simultaneously publish their draft human genome sequences, making a joint announcement of their successes. Despite the fact that Celera did not complete a sequence before the HGP, many argue that if Celera had not put pressure on the public project it would not have been completed in such a timely and effective manner.

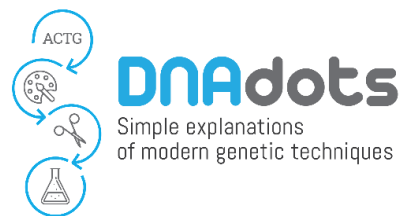
## Whose genome?

When we say “the human genome”, we gloss over the fact that no two humans have the exact same genome. In fact, both sequencing projects used multiple anonymous donors to construct their sequences, and so neither sequence is the genome of a single person. The sequence is what is known as a reference sequence. A reference sequence is recognized as not the only possible sequence, but is instead a starting point that scientists can build from and compare to. Between any two individuals, a comparison of DNA sequences will be remarkably similar; 99.6 – 99.9% of the DNA sequence will be identical. This makes reference sequences incredibly informative. The basic map of all human DNA is the same. Still, of all the genetic variation that exists, what makes us look different from each other, makes us susceptible to different diseases, and makes us genetically unique, is found in that 0.4 – 0.1%. Trying to understand those differences is the major push of sequencing efforts today. Using new techniques, such as next-generation and third-generation sequencing, genomes are being sequenced at increasingly rapid rates. In 2012, the 1000 Genomes Project announced the completed genomes of over 1000 individuals, collected from a diverse sample of ethnic groups from around the world.

## The future:

### Beyond sequencing: Understanding the genome

While the HGP finished years earlier than expected, and new sequencing technology continues to produce results at breakneck speed, some of the promises of the HGP have been slower to materialize. Many researchers made dramatic claims about speeding up the discovery of cures for human disease and ushering in a new era of personalized genetic medicine. While the HGP revealed that there are a little more than 20,000 human genes, as well as the location and sequence of each gene, the challenge is that the sequence of a gene often tells us little about its function. The HGP has allowed scientists to connect many genes to diseases through genome wide association studies (another *DNAdot*) and other methods, but to have a complete picture, a genetic sequence must be placed in the correct physiological, developmental, and even evolutionary context. Connecting this sequence data to the physiological bases for health and disease is an even bigger challenge. The HGP has brought us an incredibly long way from the early sequencing days of the 1990s, but we have learned that it takes a lot more than a DNA sequence to understand complex biology.



### Learn more:

- “Human Genome Project.” *Wikipedia*. [https://en.wikipedia.org/wiki/Human\\_Genome\\_Project](https://en.wikipedia.org/wiki/Human_Genome_Project)
- “An Overview of the Human Genome Project.” National Human Genome Research Initiative (NHGRI). <http://www.genome.gov/12011238/an-overview-of-the-human-genome-project/>
- Chial, Heidi. “DNA Sequencing Technologies Key to the Human Genome Project.” *Nature Education*. <https://www.nature.com/scitable/topicpage/dna-sequencing-technologies-key-to-the-human-828>

## Questions:

### Review:

1. What are some ways in which the Human Genome Project was considered a success?
2. The Human Genome Project did not sequence chromosome centromeres or telomeres. Why do you think scientists avoided these regions?
3. Name two ways the Human Genome Project and the sequencing project done by the Celera Corporation differed.
4. If everyone's genome is different, what makes having a reference sequence useful?
5. What are some ways that having a sequenced genome aids in research of new diseases? And why has the progress in studying new diseases not advanced as quickly as some scientists originally promised.

### Critical thinking:

1. Original estimates suggested that humans may have as many as 100,000 genes, but the Human Genome Project revealed that humans have only about 21,000 genes. The fruit fly has about 17,000 genes, about 80% of what humans have. The puffer fish is thought to have over 27,000 genes, the Norway Spruce over 28,000, considerably more than humans. We generally think of ourselves as more complex organisms than the other organisms on this list. If that is true, what does it say about the relationship between organism complexity and number of genes in the genome?
2. Scientists rarely sequence a person's whole genome. Instead they sequence short sequences related to the issue/question they are interested in. How would the Human Genome Project be useful to scientists taking this approach?

### Discussion:

1. While the Human Genome Project was widely considered a success, it was no doubt expensive, costing a total of 3.8 billion dollars. Some people think that despite its success, such a price tag is too much for a publicly funded project. Critics of the HGP will often cite that a private corporation did the same job in a shorter amount of time for less money. Supporters will note that by the time Celera Corporation began working, the HGP had already advanced the field of sequencing and genomics in ways that Celera could build off of. They also compare the publicly available data of the HGP with the privately held data created by Celera. What do you think? Should governments support such large and ambitious scientific projects in the name of advancing science and the public good? Are they too expensive? Should we let private corporations take the lead?

Answer key: Available to teachers upon request: [dnadots@minipcr.com](mailto:dnadots@minipcr.com)