DNA AND THE AGE OF INFORMATION
Michael Waterman
Wednesday 4-5:50
RIH 221

Biology is undergoing an unprecedented information explosion that would be impossible without computers. The impact of genomic information is already having profound impact on our lives and we are just at the beginning. Our ability to analyze the vast amount of data that is continuously being generated, including their visualization, has become possible due to the accelerated development of computers. In this seminar we will look at where the information is coming from and how it is being used in several key areas including genetic components of disease, human history, race and privacy. The seminar will be accessible to non-specialists and non-majors in biology and computer science.

Students will read selected material in the relevant areas and will be expected to attend regularly and participate in course activities.

Michael Waterman is University Professor of Biological Sciences, Mathematics and Computer Science and has been at USC since 1982. He is a member of the National Academy of Sciences, the National Academy of Engineering, and is a foreign member of the French Academy of Sciences and the Chinese Academy of Sciences. He is also an elected Fellow of the Los Angeles Institute for the Humanities.

Week 1**. Introduction to the course**. (8/26/2105)

Darwin and Mendel made seminal contributions to our understanding of biology and their insights are the foundation of our current knowledge. Darwin provided the conceptual framework and Mendel gave insight for the mechanisms of inheritance in a simple case. (See Chapter 4, *Origin of the Species* by Charles Darwin.) But the explosion of genetic information could not happen until computers and biotechnology reached its current level.

Week 2. **DNA structure**. (9/2/2105)

Watson and Crick in 1953 proposed the double helix structure of DNA and the modern search for an understanding biology began. Their paper is short and clearly written. Just skip over the x-ray cystography details and try to see where they say, "A pairs T and G pairs C." (*A Structure for Deoxyribose Nucleic Acid*, Watson J.D. and Crick F.H.C. *Nature* **171**, 737-738 (1953) I can tell a few Watson stories and I recommend the book, *The Double Helix* where he gives his account of their work. Rosiland Franklin's

contributions are significant and we'll talk a bit about why she did not share the Nobel Prize.

Week 3. **Reading DNA and the necessity of computers.** (9/9/2015)

In the mid 1970s a practical method of reading (sequencing) DNA was developed. The Sanger method remained the technique of choice through the Human Genome Project (2001). We will discuss why computers have been ubiquitous in sequencing DNA as well as analysis of the resulting sequences.

Week 4. **Polymerase Chain Reaction.** (9/23/2015)

Successful analysis of DNA has required millions of identical copies of the molecules. PCR allows as few as one molecule to be increased into a sample that can be analyzed. This technique revolutionized biology and many other areas such as forensics. The basic idea is straightforward and the trick is making it work. We have the original PCR machine at USC and we will take a look at it.

Week 5**. The Human Genome Project.** (9/30/2015)

This effort resulted in a sequence of the 3 billion letters of the human genome and is one of the greatest scientific projects of all human history. We will take a look behind the scenes of the HGP as well as some of the politics.

Week 6. **Sequencing after 2001.** (10/7/2015)

After the HGP amazing new technologies arose which are called Next Generation Sequencing. The HGP cost around $3 billion dollars. From 2001 to 2012 sequencing cost dropped by a factor of 100,000, plunging the price for an individual human genome sequence to $1000. This will make routine medical applications feasible and affordable.

Week 7. **Handling all that data.** (10/14/2015)

The computational methods to handle Sanger sequencing held up through the HGP but NGS required a really different approach. An elegant and simple mathematical idea from the 19-th century came to the rescue.

Week 8. **Human ancestry and extinct organisms.** (10/21/2015)

DNA lets us study ancestry and evolution much more accurately than previously. The story of the human species arising in Africa and spreading over the globe is being understood. Sequencing DNA of fossils (old bones) is revealing some startling facts about our ancestors.

Week 9. **DNA determinants of human disease.** (10/28/2015)

We humans differ from each other by about 1 letter in 1000 (we differ from chimpanzees about one letter in 100). Those differences, which are in some sense very small, are the basis of some of our differences in appearance and health. Not to be neglected are the microorganisms that live in and on us; there are 10 times as many such cells than human cells within and upon our bodies.

Week 10. **Tweaking life's code**. (11/11/2015)

As described the first week, humans have been selectively breeding animals and plants for their own purposes for tens of thousands of years. This has produced organisms that are genetically modified. However today we face the possibility of altering DNA directly, giving us much more power to alter life's history. Society will be called on to make the rules for what is and is not allowed.

Week 11. **Race**. (11/18/2015)

Information about DNA quickly leads to claims about race. It is not so obvious as to what race actually is, although people say they know it when they see it. We will look at this issue from various viewpoints.

Week 12. **Privacy.** (11/25/2015)

If knowledge of your DNA leads to predictions of health, it could impact your ability to get insurance or even to be employed. Inferences about intelligence or behavior could lead to even worse outcomes. And there is little doubt that your DNA sequence information can become public in a number of ways. We will discuss some of the aspects of this emerging issue.