2016

Decoding the Double Helix: Frederick Sanger and Sanger Sequencing

Creighton A. Jackson
Parkland College

Decoding the Double Helix: Frederick Sanger and Sanger Sequencing

Creighton Jackson

Che 101-004, Mullen, Parkland College

Introduction

Frederick Sanger (13 August 1918 – 19 November 2013)

A British biochemist and is the only recipient of two separate Nobel Prizes in the field of Chemistry. His first was awarded for his work deciphering the structure of the protein insulin. This discovery would go on to influence research on proteins and DNA, as it determined that proteins had definite structures and led to the theory that DNA directed the construction of proteins. His second was awarded for his work on the Sanger method of DNA sequencing and was shared with Walter Gilbert.

DNA Sequencing

- The field of DNA sequencing began with Watson & Crick's discovery of the double-helix structure of DNA.
- Sanger's work with sequencing the amino acid structure of insulin provided direction to DNA researchers, who began considering how DNA directed protein formation (Marks retrieved 2016)
- Early methods of sequencing DNA focused on methods for labeling specific nucleotides. (Sanger 1977)
- Sanger built on this method in 1977. (Sanger 1977)
- In 1977, the genome of the bacteriorphage φX174 (172,282 bases) was fully sequenced. (Marks retrieved 2016)
- The sequencing process was refined and eventually automated.
- As improvements were made to the Sanger method, more and more complex genomes were sequenced. In 1995, the genome of *Haemophilius influenzae* was sequenced (1,830,137 bases). (Marks retrieved 2016)
- Further improvements led to the production of the first human genome in 2001 using Sanger Sequencing. (Venter 2001)
- Advances in DNA sequencing have improved techniques and reduced costs to less than a 10,000th of what large-scale Sanger Sequencing cost initially.

Modern Sequencing Methods

- Massively Parallel Signature Sequencing
- Polony Sequencing
- 454 Pyrosequencing
- Illumina Sequencing
- SOLiD Sequencing
- Ion Torrent semiconductor sequencing
- ...and many others

Sanger Sequencing (Sanger 1977)

- Sanger developed the "plus and minus" method, which utilized DNA polymerase to transcribe specific regions of DNA under controlled conditions.
- In 1977, Sanger proposed a new method for sequencing DNA using chain terminating inhibitors to stop the transcription of DNA when a particular base would be implemented.
- By doing this for each of the four bases, then analyzing and organizing the resultant chain fractions, the overall sequence of the DNA can be determined.

Method (Simplified)

- Prepare a bulk DNA sample and four solutions of normal dNTPs and polymerase.
- Add one modified ddNTP to each solution
- Add a sample of the DNA to be sequenced to each solution and incubate, allowing the DNA to replicate.
- Denature the DNA samples using heat.
- Organize the chain fragments by size using gel electrophoresis.
- Vizualize DNA bands using autoradiography or UV light.

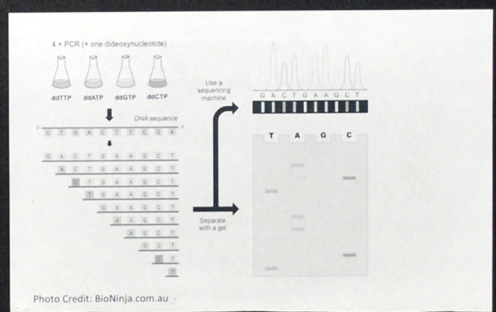

Photo Credit: BioNinja.com.au

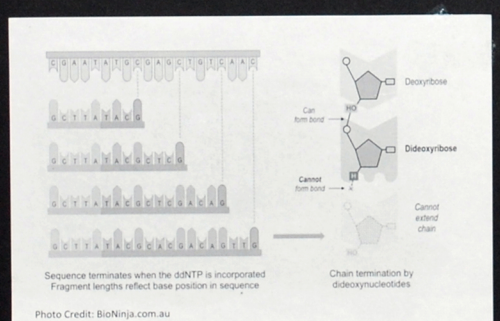

Photo Credit: BioNinja.com.au

Chain-Terminating Inhibitors

Chain-terminating inhibitors are analogues of deoxynucleosidetriphosphates (dNTPs) that contain no 3′-hydroxyl group. This prevents DNA polymerase from continuing the transcription past the base where the inhibitor is incorporated

- 2′,3′-dideoxythymidine triphosphate (ddTTP) was commercially available
- 2′,3′-dideoxyadenosine triphosphate (ddATP) had been prepared by another research team
- 2′,3′-dideoxyguanosine triphosphate (ddGTP) and 2′3′-dideoxycytidine triphosphate (ddCTP) were synthesized based on method for ddATP

Improvements to Sanger Sequencing

- Dye-terminator Sequencing (Smith 1986)
  - Use of fluorescent dyes to tag chain-terminating ddNTPs, allowing sequencing to take place in a single reaction, rather than 4 separate reactions.
  - Improves efficiency of the reaction and enables automatic reading through optical systems.
- Automation
  - Process has been automated with the development of DNA sequencers.
  - Automated interpretation of sequencing output is not as accurate as human judgment when it comes to recognizing sub-optimal results. (Smith 1986)
- Microfluidic Sanger Sequencing (Kan 2004)
  - "Lab-on-a-chip"

Biography (*A life of Research* 1992)

- Sanger was born 13 August 1918 in Rendcomb, England.
- His father, Frederick Sanger, was a general practitioner
- He graduated with his School Certificate (roughly equivalent to modern high school) a year early.
- Spent his intervening year before university working in a lab with Geoffrey Ordish, his chemistry master.
- Attended St. Johns College at Cambridge in 1936; studied natural sciences, focusing on chemistry, biochemistry, and physiology.
- Did not excel in mathematics and physics like many of his classmates.
- Graduated in December 1940 and married Margaret Joan Howe. They had 3 children together
- Received his doctorate in 1943 for his thesis "The metabolism of the amino acid lysine in the animal body."
- Both of his parents died of Cancer in his first two years of university.
- Registered as a conscientious objector to military service twice before being granted an unconditional exemption.

Photo Credit: MRC, Laboratory of Molecular Biology

References

A Life of Research on the Sequences of Proteins & Nucleic Acids: Fred Sanger in Conversation with George Brownlee. Biochemical Society, Edina – Film & Sound Online. 9 October 1992

Kan, Cheuk-Wai, et al. (2004, November). DNA Sequencing & genotyping in miniaturized electrophoresis systems. *Electrophoresis*, **25** (21-22).

Lara, M. (n.d.). The Path to DNA Sequencing: Life & Work of Fred Sanger. Retrieved November 28, 2016, from http://www.whatisbiotechnology.org/exhibitions/sanger/path

Sanger, F., Nicklen, S., & Coulson, A. R. (1977, December). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.*, **74**(12).

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Et al. (n.d.). Fluorescence detection in automated DNA sequence analysis. *Nature*, **321**(6071).

Venter, J. C., Et al. (2001, February). The sequence of the human genome. *Science*, **291**(5507).