

How large is the alphabet of DNA?

December 12 2013



Queen bee larvae in royal jelly. Worker bees and the queen have exactly the same DNA sequence, but queen larvae are fed royal jelly which epigenetically modifies their DNA so they grow to be larger and fertile. Credit: Waugsberg via Wikimedia Commons

New sequencing technology is transforming epigenetics research, and could greatly improve understanding of cancer, embryo formation, stem cells and brain function.

The mechanisms which cause certain genes to be switched on or off, and are thought to play a role in cancer development and stem cell differentiation, can now be accurately detected and studied thanks to a new DNA sequencing method.

The technology developed by Cambridge Epigenetix is helping researchers understand modifications to DNA, by detecting 'extra' DNA bases, which until now could not be definitively identified.

There are four standard DNA bases (Guanine, Cytosine, Adenine and Thymine), and the way they are ordered determines the makeup of the genome. In addition to G, C, A and T, there are also small chemical modifications, or epigenetic marks, which affect how the DNA sequence is interpreted and control how certain genes are switched on or off. The study of these marks and how they affect gene activity is known as epigenetics.

The most-studied mark is 5-methylcytosine (5mC), which is formed when molecules of methyl attach to the cytosine base of DNA, a process known as methylation. In 2009, a 'sixth' base, 5-hydroxymethylcytosine (5hmC) was discovered in human DNA, and subsequently two further modified DNA bases, 5-formylcytosine (5fC) and 5-carboxycytosine (5caC) were also identified.

Professor Shankar Balasubramanian of the Department of Chemistry founded Cambridge Epigenetix in 2012 to develop innovative epigenetic research tools that can identify, decode and help elucidate the function of the 'extra' DNA bases.

Standard DNA sequencing methods work by reading the features of the four standard bases, but cannot detect whether a cytosine base has been methylated. In order to address this shortcoming, a method called bisulfite sequencing was developed to detect methylation by adding a bisulfite reagent that converts the non-methylated cytosine bases to uracil, one of the subunits of RNA. By sequencing bisulfite-treated DNA, researchers can identify which cytosine bases were originally methylated and which were not.

However, because 5hmC and 5mC are both resistant to bisulfite treatment, it is impossible to distinguish between these two epigenetic marks using traditional bisulfite sequencing.

The reason this is a key distinction to make is that 5mC and 5hmC are thought to have completely different physiological functions. Research on the link between gene expression and methylation indicates that there are certain sites where methylation causes the gene to be switched off and silenced, whereas hydroxymethylation causes the gene to be switched on.

"Functionally, they have profoundly different meanings, yet we haven't been able to tell the difference between them using typical sequencing methods," said Professor Balasubramanian.

Following the discovery of the fifth and sixth bases, Professor Tony Green from the Department of Haematology encouraged Professor Balasubramanian to think about a new method of sequencing to detect these modifications. Balasubramanian and his PhD student Michael Booth co-invented such a method, known as oxidative bisulfite sequencing.

Oxidative bisulfite sequencing allows researchers to quantitatively measure 5mC and 5hmC at single-base resolution, enabling more accurate DNA sequencing.

The technique works by chemically oxidising 5hmC to 5fC, which like cytosine is susceptible to bisulfite treatment. Once the oxidative bisulfite reaction is complete, 5hmC and cytosine will appear in the sequence as thymine, so that the only cytosine bases remaining in the sequence are truly 5mC.

"In one reaction, you can get an accurate representation of methylation

without having to factor in the 'contamination' from hydroxymethyl C," said Professor Balasubramanian. "What our research group and Cambridge Epigenetix are doing is bringing this capability to go beyond the standard four letters of the genetic alphabet in a way that benefits from all the general innovation brought from 'next generation' [sequencing technology](#), such as the Solexa/Illumina approach."

Research studies indicate that dynamic regulation of DNA function by these epigenetic marks is essential for normal foetal development and plays an important role in cancer, neurological disorders and other diseases. In addition, it is thought that DNA modification plays a central role in stem cell reprogramming.

"Reprogramming the way DNA functions is fundamental to all living systems," said Professor Balasubramanian. "It's remarkable that for so long, we weren't aware of these other modifications in human DNA. If we've found four more bases since 2009, then who are we to argue that nothing else is there?"

More information: www.cambridge-epigenetix.com/

Provided by University of Cambridge

Citation: How large is the alphabet of DNA? (2013, December 12) retrieved 4 May 2024 from <https://medicalxpress.com/news/2013-12-large-alphabet-dna.html>

| |
|--|
| <p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p> |
|--|