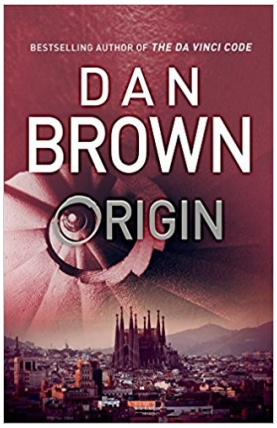
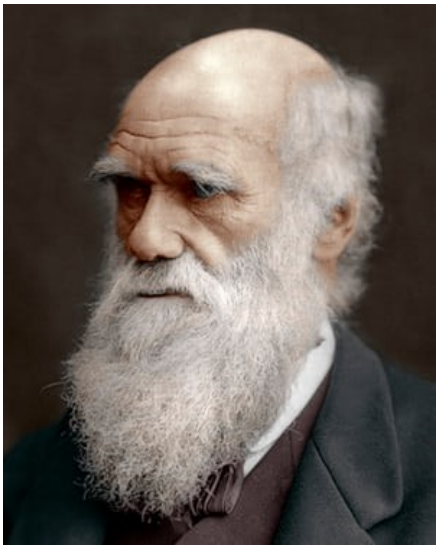
What Charles Darwin didn't know



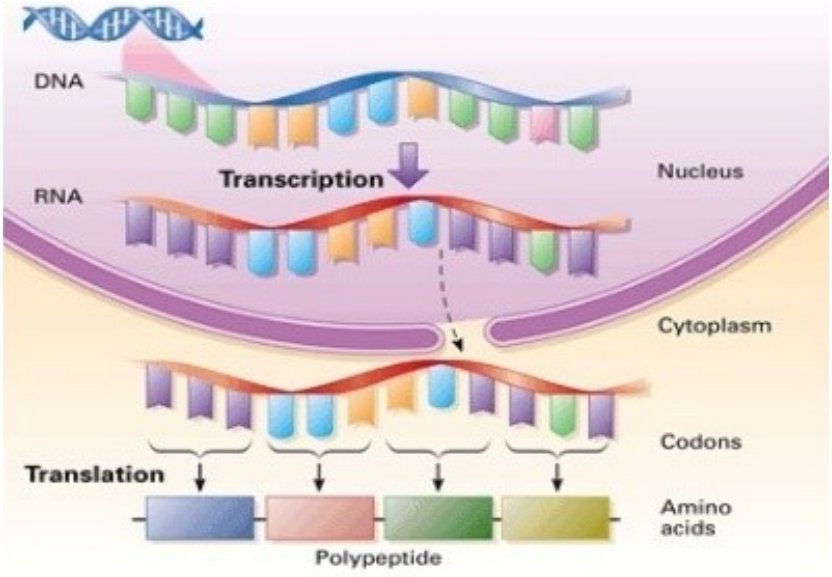
"What's the difference between a pattern and a code?" This is a key question in Dan Brown's latest novel "Origin", with a theme of how life on earth began and where it is heading. According to the hero Robert Langdon, patterns occur everywhere in nature, but codes – by definition – must carry information. He goes on to say that codes are the deliberate inventions of intelligent consciousnesses. Some critics reckon this goes too far; for example, the development of spoken language, definitely a code in this sense, could hardly be said to have begun deliberately, though there's certainly intelligence behind it.



When Charles Darwin developed his theory of evolution by natural selection in the mid-1850s, he based his ideas on natural variations in isolated species populations together with the results of other people's deliberate selective breeding experiments. He drew conclusions, still largely accepted and further developed today, about what we now call genetic inheritance, but he had no idea what the actual physical mechanism of this inheritance was, because the modern science of genetics was unknown. The founder of genetics was Gregor Mendel, an Augustinian monk in what is now the Czech Republic. He carried out experiments with varieties of peas in his monastery garden at around the same time as Darwin was writing "On the Origin of Species", worked out how certain characteristics were passed on, and coined the terms dominant and recessive to describe what we now call genes. His work was largely overlooked until the start of the 20th century.

The genetic code, which is the information of heredity, is carried by the biological molecule DNA. It is an extremely long molecule with two external backbones of phosphate (related to bones and minerals) and sugar units twisted into a double helix – a shape rather like a spiral staircase or a curved ladder laid out on the slide of an enormously tall helter-skelter. Each ladder rung connecting the phosphate-sugar rails is a pair of small molecules called nucleobases, one attached to each rail. Although each molecule of human DNA has over 3 billion such base-pair rungs, there are only 4 different bases, known as C, A, G, and T, the initial letters of their chemical names; and each rung is always a pairing of A with T, or C with G, never any other combination. This means a DNA molecule can make two identical copies by splitting down the middle of the rungs and each half assembling a matching set of bases on a new second rail. This is how your DNA reproduces itself and is passed on to your descendants.

But the real power and significance of the genetic code is that the detailed sequence of bases controls how proteins are constructed from their small amino acid building blocks. Each of the 20 different amino acids in proteins corresponds to three successive bases along the DNA chain, for example CGA means glycine and CAT means histidine. One of the most remarkable features of this biological code is its efficiency: it uses only 4 characters, in contrast to the 20–30 characters in most western alphabets and thousands of characters in written Chinese.



All humans share 99.9% of the same DNA sequence, a higher proportion than most other species. What we have in common identifies us as human, and the small differences are what make you unique (unless you're an identical twin). The entire sequence of over 3 billion bases in the human genetic code was finally established in a huge research programme, the Human Genome Project, at the turn of the millennium. Its achievement is helping in the development of medical treatments for genetically inherited diseases.

Some surprises came out of this work. Less than 2% of the DNA base sequence is actually used in sections that code for protein production (the genes). The function of much of the rest is still a mystery, though some provides self-repair mechanisms and other maintenance features.

So, if the genetic code carries information for a purpose, does it imply deliberate design? As Robert Langdon says, "That's the paradox." The characters in Dan Brown's "Origin" seem to be moving towards an odd position blending atheism with a mystical New Age spirituality and nature worship – is this true of the author too? The Director of the Human Genome Project at the time of its completion is less ambiguous in his viewpoint: Francis Collins, now the head of the US National Institutes of Health, moved from an atheist stance to a personal Christian faith largely as a result of his research in biology and medicine, and the developing story of DNA and the human genetic code confirmed him in this. Having begun with a Dan Brown quotation, I'll end with one from President Bill Clinton, in the speech he and Francis Collins prepared together for the announcement of the first draft of the human genome in June 2000: "Today we are learning the language in which God created life", to which Collins added "we have caught the first glimpse of our own instruction book, previously known only to God."

Bill Clegg

Modified version of the article in the Tyne Valley Express for July–August 2018