

DNA Coding Theory, DNA Library Design, and Group Testing

Vladimir Ufimtsev

Deoxyribonucleic acid (DNA) is an acid within the nucleus that is enriched with genetic codes used by living organisms for their physical development. Investigation into this substance was initiated by scientists at the Cavendish laboratory in the early 1950s. This investigation into the molecular structure of DNA suggests that DNA molecules are double-helical strands of phosphates and sugars that are bound together by perpendicular base-pairs. We can use a branch of Mathematics that was initiated around the same time as DNA research, known as Coding Theory, to study the similarity between different DNA strands. Codes built on spaces of DNA sequences can be implemented in Biomolecular computing and could have other important applications. The insertion deletion metric (built from the longest common subsequence), commonly used for DNA codes, will be discussed as well as a newer metric based t-stem similarities (Dyachkov, Macula, Rykov, Ufimtsev 2008). Applications of DNA Computing to screening experiments will also be given.