



جامعة الموصل
كلية علوم الحاسوب والرياضيات

استخدام النمذجة الماركوفية في المعلوماتية الحيوية

فاطمة محمود حسن زمزوم

رسالة ماجستير
رياضيات

بإشراف
أ.د. باسل يونس ذنون

المستخلص

تستقطب حاليا المعلوماتية الحيوية اهتمامات الباحثين لأهميتها البالغة في حياة الكائنات الحية. وفي هذه الرسالة يتم تسليط الضوء على ثلاث جوانب رئيسية من المعلوماتية الحيوية: يتمثل الجانب الاول مشاهدات حقيقية للحامض النووي الرايبى منقوص الأوكسجين DNA للإنسان، وتستخدم بعض المعالجات الرياضية والإحصائية لغرض تحليل هذه المشاهدات والتعرف على بعض سماتها وخصائصها. إذ يتم إجراء تحليل بياني وتُرسم القواعد النيروجينية للحامض النووي ببعدين وثلاثة أبعاد. كما يُدرس الترابط الداخلي بين مواقع الحامض النووي ويجرى تحليلا طيفيا لمشاهدات الحامض النووي. وتُجرى أيضا محاولة للتعرف على مرتبة هذه المشاهدات. ومن خلال هذه الدراسة التحليلية يتبين بان مشاهدات الحامض النووي الرايبى منقوص الأوكسجين DNA للإنسان معقدة البنية التركيبية، ومترابطة مع بعضها البعض وعلى فترات طويلة الأمد. اما الجانب الثاني فيتمثل في مسألة التعرف على الطفرات الوراثية، إذ تعد مسألة التعرف على الطفرات الوراثية من المسائل ذات الأهمية الخاصة بعلم الوراثة لعلاقتها الوثيقة بالتطور الوراثي للكائنات الحية. وفي هذه الرسالة تقدم ثلاثة طرائق إحصائية ورياضية مقترحة لغرض التعرف على الطفرات الوراثية، الأولى تعتمد أساسا على النموذج الماركوفي، والثانية على نماذج الانحدار الذاتي، وأما الثالثة فتعتمد على خوارزمية مقترحة وموظفة على نموذج ماركوف المخفي. وتطبق الطرائق الثلاثة على متابعة الحامض النووي منقوص الأوكسجين DNA للإنسان، ويتبين من التطبيق العملي ان هذه الطرائق تعطي نتائج جيدة في الكشف عن الطفرات الوراثية. واما الجانب الثالث فيتمثل في التعرف على مواقع الجينات. حيث يوظف نموذج ماركوف المخفي للمعالجة الرياضية. وتقترح خوارزمية للتعرف على موقع الجين P53 ضمن متابعة الحامض النووي DNA. ومن التطبيق العملي للخوارزمية المقترحة على متابعة الحامض النووي منقوص الأوكسجين (DNA) للإنسان يتبين الكفاءة العالية لهذه الخوارزمية.

UNIVERSITY OF MOSUL
COLLEGE OF COMPUTER SCIENCES
AND MATHEMATICS



THE USE OF MARKOVIAN MODELING IN
BIOINFORMATIC

Fatima Mahmood Hasan Zamzwm

M.Sc. / Thesis

Mathematics

Supervised By

Prof. Dr. Basil Younis Thanoon

2011 A.D.

1432 A.H.

Abstract

Bioinformatics is currently attracting researchers to the concerns of critical importance in the lives of organisms. This thesis is to shed light on three key aspects of bioinformatics: the first is the real views of nucleic acid DNA of humans, some treatments use mathematical and statistical analysis for the purpose of these observations and to identify some of the features and characteristics. A two-dimensional and three dimensions graphical analysis of actual and nitrogenous bases of DNA is performed. It also examines the inter-connectedness between the sites of DNA being analyzed spectrally for Views DNA. An attempt to identify the rank of these observations is also performed. Through this analytical study it is shown that DNA synthetic complex structure, and interdependent with each other and on long-term periods. The second aspect is to the question of a genetic mutation, it is a question of identify genetic mutations of issues of special interest in genetics to its close relationship with the genetic evolution of living organisms. In this thesis, three statistical methods are proposed for the creation of genetic mutations, the first depends mainly on the form Markovian model, and the second on the autoregressive models, and the third depends on the algorithm proposed employee and the hidden Markov model. By applying the three methods on DNA of humans, it is shown that the practical application of these methods give good results in the detection of genetic mutations. The third aspect of this thesis is to identify the gene locations. Employing the hidden Markov model for the mathematical treatment. An algorithm is proposed to identify the location of the gene P53 in the DNA sequence. It is the practical application of the proposed algorithm on the DNA sequence shown high efficiency of this algorithm.