The 3rd Iranian Conference on Bioinformatics

Abstract Book

5-6 January 2010

Tel: +98 21 44580301-310 (342)
Fax: +98 21 44580417

Address: National Institute of Genetic Engineering and Biotechnology (NIGEB)
http://www.nigeb.ac.ir
http://icb.ut.ac.ir
icb@nigeb.ut.ac.ir

15 و16 دی ماه 1388
تلفن: 21-44580301-310 داخلي 342
دورنگار: 44580417

ادرس: کیلومتر 15 اتوبان تهران-کرج-شهرک علم و فناوری پژوهش.
بلوار پژوهش، پژوهشگاه ملی و مهندسی رزیک و زیست فناوری

فارسی
In The Name of God
Conference Head

- Prof. Abbas S. Lotfi, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran

Organizing Team

Organizing Committee Chair

- Dr. Zarrin Minuchehr, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.

Organizing Committee (Alphabetical)

- Dr. Haydeh Ahrabian, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran.
- Dr. Farzad Didehvar, Department of Mathematics and Computer Science, Amirkabir University, Tehran, Iran.
- Dr. Changiz Eslahchi, Faculty of Mathematics, Shahid Beheshti University, Tehran, Iran.
- Dr. Ali Farazmand, Department of Cell & Molecular Biology, School of Biology, University of Tehran, Iran.
- Prof. Bahram Goliaei, Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran.
- Dr. Ahmad Aleyasin, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.
- Dr. Nayer Azam Khoshkhohlg Sima, Department of Physiology and Proteomics Agricultural Biotechnology Research Institute of Iran (ABRII), Karaj, Tehran.
- Dr. Ali Masoudi-Nejad, Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran.
- Dr. Abbas Nowzari-Dalini, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran.
- Dr. Mehdi Sadeghi, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.
Scientific Committee Chair

- Dr. Armin Madadkar Sobhani, Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran.

Scientific Committee (Alphabetical)

- Dr. Haydeh Ahrabian, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran.
- Dr. Abbas Alemzadeh, School of Agriculture, Shiraz University, Shiraz, Iran.
- Dr. S. M. Amini Nasab, School of Biology, College of Science, University of Tehran, Tehran, Iran.
- Dr. Farzad Didehvar, Department of Mathematics and Computer Science, Amirkabir University, Tehran, Iran.
- Dr. Mansour Ebrahimi, Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran.
- Dr. Elahe Elahi, School of Biology, College of Science, University of Tehran, Tehran, Iran.
- Dr. Esmaeil Ebrahimi, Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran.
- Dr. Changiz Eslahchi, Faculty of Mathematics, Shahid Beheshti University, Tehran, Iran.
- Dr. Ali Farazmand, Department of Cell & Molecular Biology, School of Biology, University of Tehran, Iran.
- Prof. Bahram Goliaei, Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran.
- Dr. Ghasem Hosseini Salekdeh, Physiology and Proteomics Department, Agricultural Biotechnology Research Institute of Iran (ABRII), Karaj, Iran.
- Dr. Massoud Mahmoudian, Department of Pharmacology, School of Medicine, Iran University of Medical Sciences, Tehran, Iran.
- Dr. Mohammad Ali malboobi, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.
- Dr. Ali Masoudi-Nejad, Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran.
- Abbas Nowzari, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran.
- Dr. Hamid Pezeshk, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran.
- Dr. Houman Razi, School of Agriculture, Shiraz University, Shiraz, Iran.
• Dr. Mehdi Sadeghi, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.
• Dr. Ahmad Shafieai, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran.
• Dr. Parvin Shariati, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.
Sponsors

Iranian Bioinformatics Society (IBIS)

National Institute for Genetic Engineering & Biotechnology (NIGEB)

University of Tehran

The Center of Excellence in Biomathematics

Pasteur Institute of Iran
Eastern Mediterranean Health Genomics and Biotechnology Network

Iranian Biotechnology Society

Biotechnology Incubator

Islamic Azad University (North Tehran)

Iranian Proteomics Society
Institute for Research in Fundamental Sciences (IPM)

Tehran Municipality

Oil Industries' Engineering and Construction (OIEC)

Jarf Company

Persian Bioinformatics Systems Company
Special Thanks to:

- Samira Kheitan
- Masoumeh Nouri
- Leila Ghanbari Maman
- Farahnaz Alizanjani
- Shamila Darvishalipoor
- Leila Badiefar
- Zahra Mirzaeezadeh
- Shiva Valaee
- Soheila Davaeifar
- Roghayeh Mozouni
- Camelia Siamaki
- Hamed Hajimohammadi

And all other colleagues at the National Institute of Genetic Engineering and Biotechnology (NIGEB) and Institute of Biochemistry and Biophysics (IBB) which this conference have not taken place without their precious efforts.
Bioinformatics Conference Program
Tuesday 5th Jan. and Wednesday 6th Jan. 2010
National Institute of Genetic Engineering and Biotechnology, Tehran, Iran

Tuesday 5th January 2010
Registration: 8:00-9:00 in the morning

Welcome Session

9:00-9:10  Prof. Abbas Sahebghadam Lotfi Head of the Conference and National Institute of Genetic Engineering and Biotechnology

9:10-9:20  Prof. Bahram Goliaei President of the Iranian Bioinformatics Society

9:20-9:30  Dr. Armin Madadkar-Sobhani Scientific Committee Chair

9:30-10:00  Opening Ceremony : Poster Hall
Tuesday Morning Session A: 10:00 - 12:30

Chair: Haydeh Ahrabian, Farzad Didehvar, Mohammad Ganjtabesh, Mehdi Sadeghi, Hamid Pezeshk,

Keynote 1
10:00-10:30  Hamid Pezeshk
A Semi Markov Model For Prediction Of Protein Secondary Structure

Presentations
10:30-10:45  Mohammad Ganjtabesh
Enumerating The Number of RNA Structures
10:45:11:00  Negin Najafi
Motif Finding in Biological Sequences Using Neural Networks
11 – 11:15  Amir Rahimi
Selection Of Appropriate Scoring Matrix For Finding Templates In Homology Modeling
11:15 – 11:30 Reyhaneh Esmaielbeiki
Predicting Protein-Protein Interactions Using the Mirror Tree
11:30 – 11:45 Mahsa Behzadi
Modeling And Stability Analysis Of Interconnected Regulatory Cycles
11:45 – 12  Reza Hassanzadeh
MCQ-net : An Algorithm for Constructing Phylogenetic Networks

12-14 Pray, Lunch, Posters and Exhibition
**Tuesday Afternoon Session B: 14:00-16:30**

Chair: Elaheh Elahi, Ali Farazmand, Massoud Mahmoudian, Ali Malboobi, Ghasem Hosseini Salekdeh,

**Keynote 2**
14:00 - 14:30  
Mohammad Ali Malboobi

Highly Diverse Plant Acid Phosphatases Are the Products of Both Convergent and Divergent Evolutionary Processes

**Presentations**

14:30 – 14:45  
Reza Ahangari Cohan

Rational site selection for cysteine specific PEGylation Using Bioinformatics: Erythropoietin As a Case Study

14:45 – 15:00  
Yaser Merrikhi

Modeling the effect of external electric field on the structure and Function of FoF1-ATPase Nano-Biomolecular Motors Via Molecular Dynamics Method

15:00 – 15:15  
Mohammad Taghi Shakeri

Assessing Results Of Clustering Gene Expression Data Via Margin Trees Method

15:15 – 15:30  
Nasrin Kahkeshani

Combinatorial Enumeration of RNA Structures

15:30 - 15:45  
Vahid Aslanzadeh

A Proposed Model For Mechanism Of Exon Skipping In Human And Mouse Transcripts

15:45-16:00  
Motahhareh Mohsenpour

Bioinformatics Analysis of Plastid Genomes in Order to Efficient Gene Targeting to Plastomes

16-16:30  
Coffe Break, Exhibition and Posters
Tuesday Afternoon Session C: 16:30-18
Chair: Bijan Bambai, Ali Farazmand, Khosro Khajeh, Ali Masoudi-Nejad, Mohammad Ali Malboobi,

Presentations
16:30-16:45  Mohsen Taheri
Overlap Divide and Conquer: A new Method for Haplotype Inference with Maximum Parsimony
16:45-17:00  Mostafa Ghaderi-Zafrehei
Expression Profiling of Genes Involved in Muscular Activity and Inactivity in Hind Legs of Rats for Comparative Transcriptomics
17:00-17:15  Fatemeh Zare-Mirakabad
SA-PSOMF: An Improved PSO-Based Algorithm For pattern Discovery With Simulated Annealing
17:15-17:30  Ahmadreza Ghaffarizadeh
A Novel Evolutionary algorithm based method for Efficient Quantitative Trait
17:30:17:45  Mohammad Engareh
A New Algorithm for computing Sequences Similarity
17:45-18     Mehdi Mirzaie
A New Scoring Function For Discrimination Of Native Structures From Decoys
Wednesday 6th January 2010

Wednesday Morning SessionD: 8:30 -10:30

Chair: Ahmad Shafiei, Haydeh Ahrabian, Changiz Eslahchi, Ali Masoudi-Nejad, Abbass Nowzari

Keynote 3
8:30-9:00 Changiz Eslahchi
Evaluation of Clustering Algorithms in Protein Interaction Network

Presentations
9 – 9:15 Azim Dehghani Amirabad
Identification Of Novel potential BACE1 Inhibitors as Potential Therapeutics For Alzheimer's Disease Through Virtual Screening and Artificial Neural Network Methods

9:15-9:30 Shahrzad Shaterzadeh-Oskuei
Finding Motif in DNA Sequences Using Hidden Markov Model

9:30 – 9:45 Amir Lakizadeh
Are There Any Differences Between Feed-Forward And Recurrent Neural Networks In Predicting Thermostable Protein Temperature?

9:45 – 10:00 Sepideh Pashami
A Fast Graph Query Algorithm In Biological Network

10:00 – 10:15 Mohammad Taghizadeh
Creating An Exact Model Of 3D Protein Structure For In Silico Site Directed Mutation
10:15 – 10:30   Farinaz Roshani

Propagation of Information in the Genetic Networks

10:30 – 11 Coffee Break, Exhibition & Posters
**Wednesday Morning Session E: 11:00-12:30**

Chair: Ali Karami, Mehdi Sadeghi, Changiz Eslahchi, Ali Masoudi-Nejad, Elaheh Elahi

**Keynote 4**
11 – 11:30  
**Ali Karami**  
The Roots of Systems Biology and Complex Recurrent Interactions: from Henri Poincare to Robert Rosen

**Presentation**
11:30 – 11:45  
Alireza Meshkin  
Modeling and Implementing An Agent-Based System For Prediction of Protein Relative Solvent Accessibility

11:45 – 12  
P. Gifani  
Fractal Analysis of DNA by Nonlinear Genome Signal Processing for Exon and Intron Separation

12 – 12:15  
Mohammad Hossein Ferdosi  
Comparison of Different Softwares for Haplotype Inference Based on Number of Genotype required for a Specific Accuracy

12:15 – 12:30  
Akbar Karami  
Bioinformatics Study of Phenylacetaldehyde Synthase (PAAS)

12:30 – 14  Pray, Lunch, Posters and Exhibition
**Wednesday Afternoon session SessionL: 14:00-16:30**
Chair: Ali Karami, Bijan Bambai, Khosro Khajeh, Mansour Ebrahimi, Elaheh Elahi

**Presentations**

14:00 – 14:30  
*Mansour Ebrahimi*  
Application of Bioinformatics Tools in Protein Engineering; Presenting a Few Applications Employed in Our Labs

14:30-14:45  
*Narges Rahpayma*  
Which Features Are Responsible For Halolysin Proteins Halostability?

14:45-15:00  
*Mehdi Kashani*  
Digital DNA

15:00-15:15  
*Somaye Baseri*  
A Precise Model For Prediction Of In Vitro Plant Somatic Embryogenesis Based On The Microarray Gene Expression Data

15:15-15:30  
*Narges Shamabadi*  
Are There Any Differences Between Features Of Proteins Expressed In Malignant and Benign Breast Cancer

15:30-15:45  
*Esmaeil Ebrahimie*  
Determining The characters of Thermostable Xylanase Enzyme

15:45-16:00  
*Mehdi Esmaeilzadeh*
Racial Classification Using Cephalic Size and Support Vector Machines

15:45-16:00  Saeed Bigdeli
Prediction of Protein Stability Changes by Data Mining and Support Vector Machine

16:00-16:30  Closing Ceremony
Contents

A Semi Markov Model for the Prediction of Protein Secondary Structure ........................................... 30
Enumerating the Number of RNA Structures .................................................................................................. 31
Classification of RNA Bi-Secondary Structures and Identification of Illegal Structures ......................... 32
Motif Finding in Biological Sequences Using Neural Networks ................................................................. 33
Selection of Appropriate Scoring Matrix for Finding Templates in Homology Modeling ....................... 34
Predicting Protein-Protein Interactions Using the Mirror Tree .................................................................... 35
Modeling and Stability Analysis of Interconnected Regulatory Cycles ....................................................... 36
MCQ-net: An Algorithm for Constructing Phylogenetic Networks ............................................................ 37
Highly Diverse Plant Acid Phosphatases Are the Products of both Convergent and Divergent Evolutionary Processes .................................................................................................................. 38
Rational Site Selection for Cysteine Specific PEGylation Using Bioinformatics: Erythropoietin as a Case Study .............................................................................................................................................. 39
Modeling the Effect of External Electric Field on the Structure and Function of FoF1-ATPase Nano-Biomolecular Motors via Molecular Dynamics Method .................................................................................. 40
Assessing Results of Clustering Gene Expression Data via Margin Trees Method .................................. 41
Combinatorial Enumeration of RNA Structures ............................................................................................. 42
A Proposed Model for Mechanism of Exon Skipping in Human and Mouse Transcripts ...................... 43
Position Dependencies in Human and Mouse Cassette Exons and Retained Intron Splice Sites .......... 44
Bioinformatics Analysis of Plastid Genomes in Order to Efficient Gene Targeting Toplastomes .......... 45
Overlap Divide and Conquer: A New Method for Haplotype Inference with Maximum Parsimony ... 46
SA-PSOMF: An Improved PSO-Based Algorithm for Pattern Discovery with Simulated Annealing ..... 48
A Novel Evolutionary Algorithm Based Method for Efficient Quantitative Trait Loci Mapping .......... 49
A New Algorithm for Computing Sequence Similarities .................................................................................. 50
A New Scoring Function for Discrimination of Native Structures from Decoys .................................. 51
Evaluation of Clustering Algorithms in Protein Interaction Networks ................................................. 52
Identification of Novel Potential BACE1 Inhibitors as Potential Therapeutics for Alzheimer’s Disease through Virtual Screening and Artificial Neural Network Methods ........................................ 53
Finding Motif in DNA Sequences Using Hidden Markov Model ......................................................... 54
Are There any Differences between Feed-Forward and Recurrent Neural Networks in Predicting Thermostable Protein Temperature? .................................................................................. 55
A Fast GRAPH Query Algorithm in Biological Network ...................................................................... 56
Creating an Exact Model of 3D Protein Structure for in Silico Site Directed Mutation ....................... 57
Propagation of Information in the Genetic Networks ............................................................................. 58
The Roots of Systems Biology and Complex Recurrent Interactions: from Henri Poincaré to Robert Rosen........................................................................................................................................ 59
Modeling and Implementing an Agent-Based System for Prediction of Proteins Relative Solvent Accessibility .................................................................................................................................... 61
Prediction of Relative Solvent Accessibility with Support Vector Regression Based on Qualitative and Quantitative Features Selection .................................................................................................. 62
Fractal Analysis of DNA by Nonlinear Genome Signal Processing for Exon and Intron Separation ..... 63
Comparison of Different Softwares for Haplotype Inference Based on Number of Genotypes Required for a Specific Accuracy ........................................................................................................ 64
Bioinformatics Study of Phenylacetaldehyde Synthase (PAAS), a Protein Involved in Flower Scent Production, in Rosa Hybrida and Petunia Hybrid ............................................................................. 65
Digital DNA ............................................................................................................................................. 66
Which Features Are Responsible for Halolysin Proteins Halostability? ............................................. 67
Application of Bioinformatics Tools in Protein Engineering; Presenting a Few Applications Employed in Our Laboratories ........................................................................................................... 68
A Precise Model for Prediction of in Vitro Plant Somatic Embryogenesis Based on the Microarray Gene Expression Data ............................................................................................................. 69
Are There any Differences between Features of Proteins Expressed in Malignant and Benign Breast Cancers? ............................................................................................................................................. 70
Determining the Characters of Thermostable Xylanase Enzyme .............................................................. 71
Racial Classification Using Cephalic Size and Support Vector Machines ....................................................... 72
Prediction of Protein Stability Changes by Data Mining and Support Vector Machine ................................. 73

**Tuesday Posters** ................................................................................................................................................. 74

Microarray Data Analysis for Detection and Classification of Viral Infection .................................................. 75
Complex Recurrent Interactions in Systems Biology: from Henri Poincare to Robert Rosen ............................ 76
APOBEC1 3D Structure Modeling By PDBViewer .............................................................................................. 78
Prediction of Relative Solvent Accessibility Using Pace Regression .............................................................. 79
Pattern Matching of Dynein Heavy Chain Family .............................................................................................. 80
Structure-Based New Antifungal Compounds Discovery ............................................................................... 81
Boolean Modeling of T Helper Cell Gene Regulatory Network ........................................................................... 82
The Utility of Chloroplast Matk Gene as Plant DNA Barcode for Phylogenetic Analysis in Zingiberaceae Family ............................................................................................................................................... 83
Study of Knowledge Status of Non-Physician Workers in a Health Care Center ............................................ 84
A Similarity-based Selection Mechanism for Genetic Algorithms to Solve Assignment Problems .............. 85
New Online Facilities for "MIRU-VNTR" Mycobacterium Tuberculosis Data Analysis: Comparing with Offline Analyze ................................................................................................................................................. 86
Absents of Automatic Software for Calculating HGDI Index for Calculating Score of Clustering in Mycobacterium Tuberculosis Epidemiology Studies ....................................................................................................... 87
Design of GMP Software for Analysis of Genome and Amino Acid Sequences ........................................... 88
An Online-Random Algorithm for Motif Finding Problem in Biological Networks ........................................ 89
Designing and bBlasting Primers for Amplification and Separation of Particular Sections of the Genome ......................................................................................................................................................... 90
Molecular Study on Breeding Plans of Iranian Native Golpaygani Cattle by Stochastic Simulation ............ 91
Protein-Protein Interactions Network Topology Concept: Aim at Suitable Agriculture of Tomorrow .......... 92
Suggestion of a More Efficient Model of Artificial Protein According to Probable TIMP-4 Three Dimensional Structure .................................................................................................................................................. 93
Study of Evolution Phylogeny of Sturgeon Fish Using Beluga (Huso huso) as Model ........................................ 94
Microsatellite Profiling Technique and Assignment Programs Serve to Detect Parentage in Caspian Salmon (Salmo Trutta Caspius) ......................................................................................................................... 95
The Assessment of Biosafety in Cry Proteins by Using Allergenic Databases ................................................. 96
Bioinformatics Analysis for Study of PPAR Gamma Gene Expression during Neurogenesis ......................... 97
An Engineered Peptide for Sticky Hands of HIV Virus .................................................................................. 98
A Similarity-based Selection Mechanism for Genetic Algorithms to Solve Assignment Problems ........ 99
Bioinformatics Studies for Determination of Transcription Factor Binding Sites in Promoters of Mouse PPARγ Isoforms .................................................................................................................. 100
Pathogenicity Analysis of G9055A Mutation in Mitochondrial ATPase 6 Gene in Iranian Patients with Familial Brugada Syndrome ........................................................................................................ 101
Alternative Splicing and Plant Functions .................................................................................................... 102
Neutrality Test of Two PAH-BglII and PAH-EcoRI Polymorphisms in the Isfahan Population ................. 103
Introduction of Candidate Substrates for Drosophila Gamma-Carboxylase .................................................. 104
Annotation of Herceptin/Her2 Interactions Using Protein Databases and Molecular Structure ............ 105
Bioinformatics Studies to Identify a Putative Region for Promoter of Mouse PEP .................................. 107
Survey of Data Mining Algorithms in the Protein-Protein Interactions Analysis ....................................... 108
Bioinformatics Approach for Designation of the Most Appropriated shRNA to Knock Down Peroxisomal Protein Gene ......................................................................................................................... 109
The Comparison Results of Two Secondary Protein Structure Prediction Methods on Mitochondrial A458T Missense Mutation in ND5 Protein .............................................................................. 110
Designing a Set of Universal Primers for Detection of β-Lactamase Super Family .................................. 111
Use of Universal Primers to Targeting of 16S rRNA Gene for Identification of Gram-Positive Bacteria in Environmental Samples ........................................................................................................ 112
Evaluation a Set of Universal Primers for Molecular Detection of TEM and AmpC (Dha and Mox)-Type -Lactamase in Clinical Isolates of Escherichia Coli ............................................................................. 113
Evaluation of SaCOL2291 and SaCOL2581 Genes Polymorphisms, New Candidate Genes for Preparation of S.aureus Vaccine ............................................................................................................ 114
Informative SNP Selection by Particle Swarm Optimization

Analysis of Synonymous Codon Usage Bias and Base and Amino Acid Composition in 13 Species of Flaviviridae

Bioinformatic Investigation of Short Tandem Repeat Loci D6S2879 and D6S2880 Associated with HLA-DRB1

Identification of Functional Primers for M2_3_22 STR Marker Next to the HLA-DRB1

Bioinformatics Comparison of Hemagglutinin in Different Genotypes of Measles Virus

Genome-Wide Screening for TnrA Regulated Genes of Bacillus Clausii Associated with a TnrA Box

Application of Nucleotide Substitution Models in Reconstructing Phylogenetic Tree

Comparison of the Alpha-Amylase Proteins from Several Amylolytic Microbial Strains

Barcoding of Life: A Tool for Insect Identification

Comparison of the IturinA Operon in Bacillus Species

Microsatellite Diversity and Population Genetic Structure of Rutilus frisii kutum in Mazandaran coasts

Molecular Characterization and Functional Analyses of a Novel Cold Stress-Regulated and Chloroplast-Targeted Protein from Cereals

The Use of DNA Barcoding for Identification and Classification of Cyst Forming Nematodes

From Protein-Protein Interaction to Neurotransmitter Secretion in the Nervous System

Motif Search in DNA Sequences Using Fuzzy Logic

R Package, a Powerful Tool for Bioinformatics and Its Role in Statistical Analysis of Phylogenetic Trees

Analyzing ChIP-Seq Data

Prediction of Three-Dimensional (3D) Structure and Function of ATAXIN-2 Protein in Rice

Application of GenAlEx and TFPGA Softwares for Identifying Fall and Spring Migrating Caspian Salmon (Salmo trutta caspius Kessler, 1877) by Using AFLP Molecular Marker

In Silico Expression Analysis of a Protein Phosphatase 2C Subfamily in Rice
A Novel View of Nano-Molecular Modeling Study of the Synchronize Interaction between Tamoxifen and Aspirin with Human Serum Albumin .......................................................... 135
Probing the Binding Cavity of Lomefloxacin on HSA Based on its Comparative Docking with Ciprofloxacin ........................................................................................................... 136
Searching Genomes of Retroviridae Family for Riboswitches ........................................ 137
Expressed Sequences Tags (ESTs) Analysis of Vernalization Process in Shoot ApicalMeristem of Triticum Monococcum ......................................................................................... 138
Differences of ftsA Gene Sequence of Two Different Intracellular Bacteria: Wolbachia .......................................................... 139
Phylogenetic and Functional Analysis of Zn/Cd/Pb-ATPase Pump Using Bioinformatics .......... 140
A Molecular Modeling Study of Interaction between Amlodipine and Human Serum albumin Based on Lamarckian Genetic Algorithm by Adopting Three Different Grid Box sizes .......................................................... 141
Bioinformatics Analysis for Design and Construction of a Plant/Bacteria Hybrid SigmaFactor, for Specific Gene Expression in the Plant Plastids ........................................................................................................ 142
Haplotype Inference with Greedy Parsimonious Tree Grow Method .................................. 144
Structural Bioinformatics Approaches to the Discovery of New Antimycobacterial Drugs .......... 145
Optimization of Primer Designing for TAIL-PCR Based Promoter Isolation ......................... 146

Wednesday Posters ........................................................................................................ 147
In Silico Elucidation of Iron Acquisition Genetics in Rosaceaous Fruit Trees .......................... 148
Codon Optimization of cry6 in Order to Increase its Expression Level in Potato with Using Bioinformatics Analysis .......................................................................................... 149
Structureal and Kinetic Studies of native 5-enolpyrovil shikimate -3-phosphate Synthase with Mutant and CP4 EPSPS in order to obtain Glyphosate Tolerance ........................................................................................................ 150
The Study of 3D-Structure of SUMO4 Based on Homology Modeling .................................. 151
Homology Modeling Of Functional Domain of PyocinS2 from the Crystal Structure of ColicinE7 and Studying Of PyocinS2-Immunity Protein Interactions ................................................................. 152
Phylogenic Investigation of Plants Based on Alignment of p5CS Gene Product Sequences ........ 153
Phylogenic Study and Alignment of p5CS Gene Product in Poaceae Species ........................... 154
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application of QTL Databases to Agriculture Biotechnology and Molecular Plant Breeding with Emphasis on Rice (Oryza sativa L)</td>
<td>155</td>
</tr>
<tr>
<td>Modeling the Structure of a Carrier Protein: Human Diferric Transferrin</td>
<td>156</td>
</tr>
<tr>
<td>A Novel View of the Interaction between Cyclophosphamide and Human Serum Albumin: A Molecular Modeling Approach</td>
<td>157</td>
</tr>
<tr>
<td>Codon Usage Optimization of cry1Ab Gene to Enhance Expression Brassica Napus L. Plant Using Bioinformatics Analysis</td>
<td>158</td>
</tr>
<tr>
<td>A Bioinformatics Study of Human Glutathione S-Transferases</td>
<td>159</td>
</tr>
<tr>
<td>Codon Optimization of mtld in Order to Increase Its Expression Level in Brassica napus L. with Using Bioinformatics Analysis</td>
<td>160</td>
</tr>
<tr>
<td>Application of Haplotype Block Partitioning in Genome-Wide Case-Control Association Study</td>
<td>161</td>
</tr>
<tr>
<td>Promoter Analysis of AtMYC2 and AtMYB2 Genes: Cis-acting Regulatory Elements and Their Possible Roles in Controlling Gene Expression Levels</td>
<td>162</td>
</tr>
<tr>
<td>Microarray Data Analysis of Two Transcription Factors &quot;ERF4&quot; and &quot;TFIIIB&quot; in Soybean (Glycine Max) Somatic Embryogenesis</td>
<td>163</td>
</tr>
<tr>
<td>The arsR Gene: A Comparative Phylogenetic and Bioinformatic Analysis to Study Reasons for Different Sensitivity to Arsenic Compounds</td>
<td>164</td>
</tr>
<tr>
<td>Breast Carcinoma, Intratumoural Heterogeneity and Cancer Grading</td>
<td>165</td>
</tr>
<tr>
<td>Comparison of Root Genes Expression Profiling of Wheat under Drought Stress by Expressed Sequence Tags (EST) Analysis</td>
<td>166</td>
</tr>
<tr>
<td>Phylogenetic Analysis of the Enterobacteriaceae Family Based on the Nucleotide Sequences of rpsL Gene</td>
<td>167</td>
</tr>
<tr>
<td>A Report about Choosing the Appropriate Gene Source for Cloning Based on Bioinformatics Studies</td>
<td>168</td>
</tr>
<tr>
<td>A Study on Different Analysis Methods of Real-Time PCR Data</td>
<td>169</td>
</tr>
<tr>
<td>Introduction of Empirical Bayes Approach, as a Novel Reliable Method in Microarray Data Analysis</td>
<td>170</td>
</tr>
<tr>
<td>Designing a Bioinformatic-Based System for Identification of Self-Incompatibility Alleles in the Pyrus Species</td>
<td>171</td>
</tr>
</tbody>
</table>
The Study on Conserved Regions of Pyrroline-5-Carboxylate Reductase Gene (p5CR) in the Important Plant Species ................................................................. 172
Positive Selection in Molecular Evolution of the HSP70 Gene Family ................................................................. 173
Bioinformatics Analysis of Sequence tag Sites (STss) in SMA Locus ................................................................................. 174
Bioinformatics Study of Δ 12-Desaturase Gene in Six Plant Species ................................................................................. 175
Alcanivorax Borkumensis Whole Genome Analisys for P450 Like Proteins ................................................................. 176
System Genetics Approach for Analysing Rat Muscle Expression Data ................................................................................. 177
Molecular Characterization of the Iranian Isolate of Barley Dwarf Virus Rep Protein and Its Similarity with Other Mastreviruses ................................................................................. 178
Potential of MS Excel in Developing Bioinformatics Softwares ................................................................................. 179
Genetic Diseases and Bioinformatics .................................................................................................................................. 180
Structural Studies of Therapeutic Monoclonal Antibodies Using IMGT/3D Structure; Tools and Database .......................................................................................................................... 181
The Nano-Molecular Modeling of the Interaction between Ropinirole Hydrochloride and Human Holo-Transferrin Based on Lamarckian Genetic Algorithm .............................................................................................. 182
A Novel View of the Interaction between Tamoxifen and Human Holo-Transferrin Modeled: A Molecular Modeling Approach ........................................................................................................................................................................... 183
Primer Designing and Restriction Enzyme Detection for Exon 14 of ABCG2 Gene in Iranian Holstein Bulls ....................................................................................................................................................... 184
Comparison of Conserve 16s Ribosomal RNA Sequence for Designing Universal Primer to Detect Obligate Anaerobic Bacteria ........................................................................................................................................ 185
The Use of Bioinformatics in DNA Fingerprinting of Pistachio Cultivars ................................................................................. 186
Analysis of Alternative Splicing Mediated by the PASA Software ................................................................................. 187
Application of Bioinformatics in Genetic Comparison of Caspian Salmon (Salmo Trutta Caspius) Immigrant Forms ........................................................................................................................................ 188
Analysis of Codon Usage Bias of Kinase Genes for Different Species ................................................................................. 189
Missing Value Imputation in DNA Microarrays Based on Conjugate Gradient Method ................................................................. 190
Identification of Alzheimer Related Genes Using Supervised Independent Component ................................................................. 191
Molecular Modeling Investigation of The Interaction between Amlodipine and Human Serum Albumin at Three Different pH: A Comparative with In Vivo like Model ......................................................... 192

A New Algorithm for Protein Pattern Search .................................................................................. 193

Phylogenetic Study on 6 Plants Species Based on Chitinase Gene Alignment ................................. 194

Molecular Epidemiology Application to a Mycobacterium Tuberculosis Transmission by MIRU-VNTR plus web site ............................................................................................................. 195

Prediction the Tertiary Structure of Extracellular Domains of ROR1, an Approach for Preparation Cancer Vaccines ..................................................................................................................... 196

Homology Modeling and Docking Studies of Thyroid-Stimulating Hormone (TSH) ...................... 197

Prediction of Three-Dimensional Structure of Flagellin from E. coli ............................................ 198

The Suitable Binding Sites Finding of the Interaction between Human Serum Albumin and Fluoxymestrone as a Model of Drug Binding Protein .......................................................................................... 199

Is acdS Gene Inherited by Horizontal Gene Transfer? .................................................................. 200

Sequencing and Comparison of the Nuclear Ribosomal Internal Transcribed Spacer Region for Iranian Black Caraway Populations ................................................................... 201

Introducing 1-methyl Malate as a New and Potent Inhibitor of Class A β-Lactamase by Computational Studies ................................................................................................................................. 202

Sarcin/ricin rRNA Motif Detection in PDF Files 3G4S, 3G6E, and 3CXC using FR3D Program ........ 203

Computational Simulation of Gramicidin like Channel as A Peptide Nanotube ........................... 204

Prediction of the Mode of Interaction between Monoterpenes and the Nitroreductase from Enterobacter Cloacae by Docking Simulation ................................................................................. 205

Quantitative Structure Activity Relationship Analysis of Some Isocoumarins as Pancreatic Cholesterol Esterase Inhibitors ........................................................................................................... 206

Modeling the Epithelial-Mesenchymal Transition and its Effect on Tumor Invasiveness and Metastasis ................................................................................................................................................... 207

A Novel View of the Interaction between Human Serum Albumin and Estradiol: A Molecular Docking Approach ................................................................................................................................. 208

TITAN Protein Secondary Structure Assignment by Emphasis on Tight Turns .............................. 209
Molecular Modeling Studies on Cholesterol Esterase Inhibitors: Proposition of New Ligands, and Insight into Possible Additional Effects.................................................................210

Tertiary Structure (3D) Prediction of Extracellular Domains of Ror1 ..........................................................211

Cholesterol Oxidase Gene Analysis Derived from a Native Rhodococcus sp. Strain 501..............212

Modeling of Mutation in Phenylalanine Hydroxylase S231F Using Computational Methods ..........213

Modeling Three Dimensional Structures of WW Domain of WWOX Using Homology and Molecular Dynamics Simulation.................................................................214

Comparison of Dock Techniques in Structural Based Inhibitor Design of HIV-gp41 .......................215

Molecular Dynamics Study of Human Proinsulin-Protein A Fusion System ...........................................216

A Matrix Based Pattern Template for Protein Patterns Induced on Genome .................................217
Orals
Keynote 1:

A Semi Markov Model for the Prediction of Protein Secondary Structure

Pezeshk H.

School of Mathematics, Statistics and Computer Science, and Center of Excellence in Biomathematics, College of Science, University of Tehran, Tehran, Iran

In this work, a bidirectional Hidden Markov Model (HMM) and one of its applications in protein secondary structure is introduced. Our model is based on the assumption that the information on both sides of an amino acid can provide a suitable tool for measuring dependencies. An especial type of HMM known as the Segmental Semi Markov Model (SSMM) assigns a segmental distribution to each state. In an SSMM a segment of observations is emitted from a single state. We use a bidirectional version of an SSMM for the prediction and compare our results with those obtained from ordinary HMMs. We also use the information on a profile of the multiple alignments to improve the precision of the prediction.
Enumerating the Number of RNA Structures

Ganjtabesh M.

Center of Excellent in BioMathematics, University of Tehran, Tehran, Iran

Counting the number of RNA structures is an important combinatorial problem in computational biology. In this paper, we count the number of structures for the simple case of an RNA sequence in which any arbitrary pair of bases is allowed. The only criterion that is considered in our model is the minimum length condition for hairpin loops which is equal to 1. The asymptotic behavior and its correlation with the number of involutions are also presented.
Classification of RNA Bi-Secondary Structures and Identification of Illegal Structures

Ganjtabesh M.

Department of Mathematics, Statistics, and Computer Science, Collage of Science, University of Tehran, Tehran, Iran

RNA structures play an important role in bioinformatics. In this paper, we discuss about the different classes of RNA pseudoknotted structures, especially the class of bi-secondary structures and it’s relation to the other classes. We provide a generalized set of rewriting rules as well as an algorithm for classifying the RNA bi-secondary structures. We also introduce a new order of the RNA structure classes related to bi-secondary structures and identify the illegal structures for different classes of RNA pseudoknotted structures.
Motif Finding in Biological Sequences Using Neural Networks

Najafi N., Zare-Mirakabad F., Ahrabian H., Nowzari-Dalini A.

Biomath Group, Department Matematic Statics & Computer Sciences, University of Tehran, Tehran, Iran.

The identification of over-represented motif in unaligned DNA sequences is a challenging problem in both Computer Science and Molecular Biology. Many machine learning techniques and statistical methods are presented for solving this problem. These existing methods are not exact and it is still desirable to find more accurate ones. In this paper, we propose an adaptive resonance theory (ART) neural network as an alternative method for motif finding. First, this neural network is initialized by the Gibbs sampling strategy and then the input sequences are given to it. Finally, the motifs are reported. The advantage of this approach is that it can be used to simultaneously characterize every feature present in the dataset, thus lessening the chance that weaker signals will be missed. Our algorithm is implemented and tested on different types of real datasets. The results are compared with some other well-known algorithms and the effectiveness of our algorithm is shown.
Selection of Appropriate Scoring Matrix for Finding Templates in Homology Modeling

Rahimi R.¹, Mahdevar Gh.¹, Minuchehr Z.², Madadkar-Sobhani A.¹

¹ Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran

² National Institute of Genetic Engineering and Biotechnology, Tehran, Iran

In protein structure prediction by homology modeling, finding appropriate template and the best alignment between the found template and the target sequences are the most crucial steps. In this regard, choosing suitable amino acid substitution scoring matrix may have a significant impact on both procedures. In this study, the ability of various scoring matrices in finding the right template and the proper alignment suitable for homology modeling was investigated. Correlation of various PAM and BLOSUM scoring matrices with amino acids genetic code and physicochemical properties similarity matrices were evaluated. Sequences of 40 proteins from alpha amylase, protease and lipase families downloaded from UniProt and their structures were obtained from PDB. Sequence and structure alignments were performed for each pairs of proteins in the family. Then the correlation between the structure similarity and sequence alignment score with different matrices were calculated. RMSD was used as an indication of structural similarity and evaluated using MatAlign version 11. The results show that well-known amino acid substitution matrices, PAM and BLOSUM, have high correlation with amino acids genetic code similarity matrix. Although this is an advantage of such matrices in phylogenetics studies, but having low correlation with physicochemical properties similarity matrix, decreases their efficiency in finding the best template and alignment suitable for homology modeling. We have also introduced a new scoring matrix based on physicochemical properties of amino acids and compared it with common matrices.
Predicting Protein-Protein Interactions Using the Mirror Tree

Esmaielbeiki R., Christophe Nebel J.

Faculty of Computing, Information Systems & Mathematics, Kingston University London, UK

Protein-protein interactions are important for every process that takes place within a living cell. Moreover, since proteomic scientists aim to discover all the proteins within the human body and how they interact with each other; the study of protein-protein interactions has become a very active topic in Molecular Biology. For example, modifications in protein-protein interactions change the order of the events that take place within cells that may lead to critical diseases such as cancer. Therefore, knowledge about protein interactions can provide key information for drug design. There are many experimental techniques such as yeast 2-hybrid for predicting protein interactions. However, since they are very slow and expensive, many computational methods have been proposed. Moreover, some of these software-based approaches have shown similar or even higher accuracy when compared to experimental methods.

Among bioinformatics techniques, the mirror tree, introduced by Pazos and Valencia in 2001, is one of the most popular computational approaches to predict protein-protein interactions. It is based on the intuitive assumption that proteins which interact are under co-evolution pressure. Therefore, analysis of co-evolutionary information should contribute to the evaluation of their interaction. For the last eight years, improvements of the original mirror tree algorithm have been proposed in this paper, we provided a comparative study of these different techniques using a large dataset of *Escherichia coli* proteins. We also introduced the mutual information measure as an alternative metric to evaluate the co-evolutionary information shared by a protein pair. Results suggested that our implementation of mirror Tree was particularly accurate and the use of mutual information did not significantly affect the quality of predictions.
Modeling and Stability Analysis of Interconnected Regulatory Cycles

Behzadi M.

LIX, Ecole Polytechnique, Palaiseau Cedex 91128, France.

Biochemical reactions are continually taking place in all living organisms. The complexity of biochemical and biological processes is such that the development of computer models is often essential in trying to understand the phenomenon under consideration. Our aim is to build a generic framework with which one could simulate the behaviour of complex systems of interconnected regulatory cycles. For the simulation of a biological system we use the traditional reaction-rate approach by means of equations describing the system. In this approach, chemical reactions are modelled by ordinary differential equations (ODEs) representing the variations of the concentrations of the substances. In each of the differential equations, we express the kinetics of one reactant as a sum of fractional terms for enzymatic reactions and non-fractional terms for simple reactions. Once the model is constructed, we aim to study the various modes of cell behaviour according to the concentrations of relevant enzymes in enzymatic reactions. Since stable and unstable equilibria play different roles in the dynamics of a system, it is useful and important to be able to classify equilibrium points based on their stability, and this is what we are able to do by simulation and also by mathematical study. By stability analysis, first given equilibrium, we can determine if it is a stable point or not; furthermore through a mathematical study we are able to find the stability and instability regions by changing one or several parameters. As a first try we have constructed a model for the central part of the glycerophospholipid metabolism in the human cell. Understanding cell metabolism evolution and changes is for many scientists more than a challenge; it is the key to a thorough understanding of cell dysfunction and very likely a step toward the elucidation of carcinogenesis along the lines of Warburg’s seminal papers. In this work we perform mathematical analysis of the metabolic pathways which control and command the production of glycerophospholipids through the enzymatic reactions of phosphatidylethanolamine and phosphatidylcholine. Analysis shows that the normal cell stands at very special points of the equilibrium. We also checked our model against a series of experiments and give evidence for the crucial role of phosphatidylethanolamine N-methyl transferase (PEMT) in carcinogenesis. We currently use this approach to study the stability analysis of a complex metabolic network containing several interconnected regulatory cycles such as glycolysis, Krebs cycle, phospholipids pathway and Amino acids.
MCQ-net: An Algorithm for Constructing Phylogenetic Networks

Hassanzadeh R., Mottaghi E., Eslahchi C., Habibi M.

Faculty of Mathematics, Shahid Beheshti University, Tehran, Iran

Phylogenetic networks are models of evolution that represented biological events that are not consistent with tree-like evolution. For example, hybridization between species, lateral transfer of genes, recombination within a population and convergent evolution can all lead to evolution histories that are not adequately represented by a single tree. Sometimes the underlying evolution is tree-like, but the input data contains errors and conflicting or ambiguous signals can make a single tree representation inappropriate. In these situations, phylogenetic network can be particularly useful.

There are a number of methods for constructing various kinds of phylogenetic networks. One of these methods is Q-NET that firstly constructs a collection of circular weighted splits of the taxa set from weighted quartets and then this collection represented by a special kind of phylogenetic network called a "split network ". Such networks can be generated using SplitsTree4. In this work, we present an algorithm "MCQ-NET" based on Monte-Carlo method that is improvement of Q-NET. In order to show its performance, we applied both of these algorithms to some data sets. We illustrate that results of MCQ-NET are better than Q-NET.
Monoesteric phosphatases, commonly known as acid phosphatase (APase) enzymes, catalyze the hydrolysis of phosphoric ester bonds of various substrate types including phosphorylated sugars, lipids, proteins and nucleic acids. These enzymes play central roles in phosphate acquisition through absorption, recycling and scavenging of this essential element from both internal and external resources, particularly in plants as sessile organisms. Exhaustive reiterated sequence searches in protein databases retrieved an inventory of plant APases, particularly those from Arabidopsis thaliana and Oryza sativa as representatives of the dicotyledonous and monocotyledonous plants, respectively. Discrepancy in protein sequence length, conserved motifs, the number of exons, structural features and active site residues were all indicative of remarkable levels of diversity among functionally similar APases. A comprehensive cluster analysis anchored on the Arabidopsis and rice finished genome data sets led to the classification of the plant APases into five discrete groups with distinct family-specific signatures. Based on the inference of orthologous and paralogous relationships among the members of each family, divergence from common ancestors and lineage specific expansions within the protein families were discriminated. Considering no orthology among plant APase families as well as structural and functional analogies, we proposed that these families are the products of convergent evolution among distinct ancestral proteins toward gaining the function of phosphate bond hydrolysis from various molecular types. Subsequent divergent evolution has expanded the protein families to broaden the substrate ranges and, thus, salvaging phosphate from all possible resources.
Rational Site Selection for Cysteine Specific PEGylation Using Bioinformatics: Erythropoietin as a Case Study


R&D Department, Production & Research Complex, Pasteur Institute of Iran

Proteins can be used as therapeutic agents. However, they do not have ideal pharmacokinetic properties. Short half-life, immunogenicity, low absorption from the gastrointestinal tract and physicochemical instability are the main problems with respect to using proteins as drugs. There are several methods for improvement of the pharmacokinetic properties of proteins, such as amino acid substitution, conjugation with proteins and polymers, delivery systems. PEGylation of proteins is the attachment of the poly ethylene glycol (PEG) molecule to proteins. In cysteine specific PEGylation, free cysteine residues are introduced by site-directed mutagenesis. Because of change in the conformation of protein that leads to decrease in activity, the position of the introduced cysteine residue is important. Therefore, rational selection of positions for replacement with the cysteine residue is necessary. Here, we used bioinformatics in order to predict the best PEGylation sites of erythropoietin (EPO) as a case study. First, a comprehensive literature review was carried out in order to determine plausible sites in EPO suitable for PEGylation. Then, three dimensional structures of native EPO and selected cysteine analogues were generated by homology modeling using MODELLER. Considering the surface area accessibility (SAA) of cysteine residues, exposed cysteine analogues were selected for 5-ns molecular dynamics (MD) simulation. Stable cysteine analogues during MD simulation were subjected to screening based on pKa values, SAA, RMSD and distance between -SH groups. After screening, two cysteine analogues were suggested as candidates for cysteine specific PEGylation.
Modeling the Effect of External Electric Field on the Structure and Function of FoF1-ATPase Nano-Biomolecular Motors via Molecular Dynamics Method

Merrikhi Y. 1, Batebi S. 2, Rafii-Tabar H. 3

1 Department of Medical Physics and Biomedical Engineering, Shahid Beheshti University (Medical campus)
2 Department of Physics, Semnan University, Semnan Iran
3 Department of NanoScience, Institute for Research in Fundamental Sciences (IPM), Iran

The production of ATP, the main source of energy in the human body, is carried out by FoF1-ATPase nano-biomolecular motors. Experimental data show a relation between external electric fields and the rate of ATP production. In this paper, molecular dynamics (MD) method is used to investigate the effect of the external electric field on the structure of FoF1-ATPase subunits. Among the nano-motor subunits, the subunits C, betas, and gamma, which play a known significant role in the nano-motor function, are chosen for simulation. Initially, all the subunits are equilibrated at 300 K and 1 Bar for 1 ns. Then, to study the effect of the external electric field, 1ns-simulation is carried out in the presence of a DC electric field at three different magnitudes. The compactness of the following structures is computed during the simulations: the C subunit, the P loop and the HTH motif (two main parts of the beta subunit), the gamma-protrusion (part of the gamma subunit) and a positive charge network around the gamma-protrusion. The results show that applying an external electric field can affect the nano-motor function by compacting the C subunit, the P loop, and the HTH motif and also by decompressing the gamma protrusion and the network around it. The structural changes of the C subunit, the P loop and the positive charge network can lead to an increase in ATP production but the changes in the structure of the HTH motif and the gamma-protrusion can lead to a decrease in the production of ATP.
Assessing Results of Clustering Gene Expression Data via Margin Trees Method

Shakeri M. T., Sabaghian E.

Department of Biostatistics, Mashhad University of Medical Sciences, Mashhad, Iran

The use of clustering methods for the discovery of cancer subtypes has drawn a great deal of attention in the scientific community. While bioinformaticians have proposed new clustering methods that take advantage of characteristics of the gene expression data, the medical community has a preference for using "classic" clustering methods. There have been no studies thus far performing a large-scale evaluation of different clustering methods in this context. Results: We present the first large-scale analysis of nine different clustering methods, hierarchical clustering with single, average, complete and Ward linkages, UPGMA, Diana, K-means, PAM and CLARA methods for the analysis of 5 cancer gene expression data sets. Subsequently, we use the Margin Trees method for assessing the quality of the results of clustering methods. Ultimately, we calculate the quality of the results of the clustering methods via Kappa coefficient between result of clustering methods and result of Margin Trees method for each clustering method. Our results reveal that the PAM, followed closely by CLARA, exhibited the best Performance in terms of recovering the true structure of data sets. Also we found that Partitioning clustering methods (PAM, CLARA and K-means) have better performance than hierarchical clustering methods (hierarchical clustering with single, average, complete and Ward linkages, UPGMA and Diana). Conclusion: The validation technique used in this paper (Margin Trees) can aid in the selection of an optimal algorithm, for a given data set, from a collection of available clustering algorithms.
Research in Biology is usually thought to be based on experimentation with materials, while in Mathematical Biology experimentation is of a theoretical nature. Mathematical Biology involves the application of physical principles to biological systems. A major advantage of applying Mathematics to biological systems is the ability to construct mathematical models. Such models are mathematical systems that attempt to represent the complex interactions of biological systems in a way simple enough for their consequences to be understood and explored. Traditionally models that allowed biologists to see a problem in a simplified way have been physical systems constructed to exhibit simple biological properties which could be analyzed. In this paper, we studied the combinatorics of helical structures of RNA sequences. RNA is described by its primary sequence of nucleotides A, G, U and C together with the Watson-Crick (A-U, G-C) and (U-G) base pairing rules specifying which pairs of nucleotides can potentially form bonds. There are several ways to represent these RNA structures. We chose the diagram representation, which is particularly well suited for displaying the crossings of the Watson-Crick base pairs. A diagram is a labeled graph over the vertex set \([n] = \{1,\ldots,n\}\) with vertex degrees at most 1, represented by drawing its vertices 1,\ldots,n in a horizontal line and its arcs \((i,j)\) in the upper half-plane. The vertices and arcs correspond to nucleotides and Watson-Crick (A-U, G-C) and (U-G) base pairs, respectively. We categorized diagrams according to the 3 parameters: The maximum number of mutually crossing arcs, \(k - 1\), the minimum arc-length, and the minimum stack-length. We derived the generating function of RNA structures with pseudoknots and enumerated all \(k\)-noncrossing RNA pseudoknot structures.
A Proposed Model for Mechanism of Exon Skipping in Human and Mouse Transcripts

Aslanzadeh V. 1, Sadeghi M. 2, Zare-Mirakabad F. 3, Haghazari A. 1, Hafezparast M. 4

1 Department of Agronomy and Plant Breeding, University of Zanjan, Zanjan, Iran
2 National Institute of Genetic Engineering and Biotechnology, Tehran; And School of Computer Science, Institute for Studies in Theoretical Physics and Mathematics (IPM), Tehran, Iran
3 Center of Excellence in Biomathematics, School of Mathematics, Statistics, and Computer Science, University of Tehran, Tehran, Iran
4 Department of Molecular Medicine, University of Malaya, Kuala Lumpur, Malaysia

Pre-messenger RNA splicing is a fine-tuned process that generates multiple functional variants from individual genes. There is ample evidence that prediction of human splice sites can be refined by analyzing the nucleotides surrounding splice sites. This could mean that nucleotides at splice sites harbour information for the splicing process. The most frequent type of AS is the cassette exon (exon skipping). Using mutual information (MI), we analyzed possible position dependency of cassette and constitutive exon splice sites in human and mouse genomes. We showed that MI in the cassette exon splice sites was surprisingly higher than MI in the constitutive exon splice sites, and also MI between 3´ and 5´ splice sites of the cassette exons was stronger than the observed MI for constitutive exons. We also demonstrated that exonic-exonic and exonic-intronic interactions in cassette exons splice sites are stronger than constitutive exons and consequently cassette exons have more potential to form secondary structure. Based on these results, we proposed a new model for mechanism of exon skipping in human and mouse genomes. This model shows that skipping of exons not only depends on splicing regulatory elements, but secondary structure formation in the splice sites also acts as an important factor in control of exclusion and inclusion of cassette exons from pre-mRNA. Results provide further information for better understanding and computational predication of cassette exons.
Position Dependencies in Human and Mouse Cassette Exons and Retained Intron Splice Sites

Aslanzadeh V. 1, Sadeghi M. 2, Haghnazari A. 1, Zare-Mirakabad F. 3, Gaderian M. 1

1 Department of Agronomy and Plant Breeding, University of Zanjan, Zanjan, Iran
2 National Institute of Genetic Engineering and Biotechnology, Tehran; And School of Computer Science, Institute for Studies in Theoretical Physics and Mathematics (IPM), Tehran, Iran
3 Center of Excellence in Biomathematics, School of Mathematics, Statistics, and Computer Science, University of Tehran, Tehran, Iran

Alternative splicing is a critical post-transcriptional event leading to an increase in transcriptome diversity. Bioinformatics and experimental approaches have unraveled a large number of parameters that contribute to the regulation of splicing. One of the important factors that contribute to splice-site selection is local RNA secondary structure formation. For analysis of secondary structure formation ability in retained intron (RE) and cassette exon (CE) splice sites, using mutual information, we analyzed possible position dependencies in RE and CE splice sites. Intron retention and exon skipping are processes that one intron and one exon can alternatively use in different messages, respectively. Datasets were downloaded from the USCS genome browser and mutual information was then applied to the data. The results showed that both RE and CE splice sites had more potential to form secondary structure than non-retained introns and constitutive exons, respectively. We obtained the same results for mouse data and demonstrated that mutual dependencies in mouse and human genomes are evolutionary conserved. We concluded that secondary structure formation can play an important role in intron retention and exon skipping phenomena in the genome. These results provided further information about the importance of secondary structure formation in alternative splice site selection and prediction of cassette exons and retained introns by computational methods.
Bioinformatics Analysis of Plastid Genomes in Order to Efficient Gene Targeting To plastomes

Mohsenpour M., Babi eian Jelodar N., Tohidfar M.

1 University of Agricultural and Natural Resources, Sari, Iran.
2 Agricultural Biotechnology Research Institute of Iran (ABRII)

Chloroplasts are suitable hosts for foreign genes. Insertion of foreign genes into plastid genomes occurs through homologous recombination between flanking sequences of the gene in plastid vectors. Detection of intergenic regions for foreign gene insertion is the first stage of constructing chloroplast vectors. Use of databases and bioinformatics softwares provides fast and reliable methods to achieve the required sequences for the design of such vectors. In this study, using analysis plastid genomes in bioinformatics databases, we succeed to select a region of plastome as foreign gene targeting sequence to chloroplast genome that including the appropriate length for homologous recombination and integration in plastome, specific plastid origin of replication, targeting gene to inverted repeat region of plastid genome and restriction enzymes unique sites in the center of flanking region for foreign gene integration. Because of the flanking region sequence similarity percent the efficiency of the transgenic will increase plastome dramatically, so a pair of primeries was designed to isolate the plastome flanking regions. These primers had 100% similarity in all plant plastomes, but the length and sequence of their amplification in different genomes were variable. This fragment from different plant plastomes was analyzed for GC Content, sequence complexity, Absolute complexity, nucleic acid distribution and Similarity analysis. Furthermore, the variety of these sequences in different plants was determined by using nucleotide align, restriction endonuclease enzymes recognition sites and tree analysis. Flanking region sequence from tobacco, cotton, corn, lettuce, tomato, carrot, and even the canola and lemon plastomes ,the plastome sequences of which are still not available in Gene Bank were isolated and cloned. These recombinant plasmids can be used as appropriate basic and specific plastid vectors for chloroplast transformation. Thus, in this study we succeed in designing high efficient specific plastid transformation vectors by using data obtained from bioinformatics analysis, not only for plants the their plastome of which were completely sequenced, but also for other plant species that do not have plasmoe sequences in databanks the.
Overlap Divide and Conquer: A New Method for Haplotype Inference with Maximum Parsimony

Taheri M. 1, Sadeghi M. 2, Ghasem Aghaee N. 3

1Department of Computer Science, Shikh bahaee university of Isfahan, Isfahan, Iran.
2National Center of Genetic and biotechnology, Tehran, Iran.
3Faculty of Computer Engineering-Isfahan, Isfahan, Iran

Haplotype inference one of the most important problem in bioinformatics due to its abundant use in analysis of large molecular data and providing a method for mapping inheritance disease to some patterns in SNP sequence of DNA. Since direct acquisition of haplotype information from biological data is too costly, and time consuming, this fore, computational methods for solving the haplotype inference problem is usually used. In this paper, we introduced a new method for haplotype inference by maximum parsimony condition and named it as OWPMR. OWPMR, based on divide and conquer with overlap partition. OWPMR is used for solving HIPP with N genotype and M, and the SNP for each genotype, breaking problem to a Sub Problem with size W SNP, after solving each sub matrix, merging results of all sub matrices and forming final result. Consequently, each sub matrix has an overlap SNP column with a W-1 sub problem after in the genotype matrix. In OWPMR, W or size of overlap segment, as a partition criterion, makes a parallel framework for solving HIPP and introduces this method as one of the first approaches with parallel execution capability in solving of HIPP. For evaluation purposes, we compared results of OWPMR with results of three softwares: PHASE, 2SNP, GERBIL from the perspective of accuracy. High speed of OWPMR and relatively good quality of the results obtained, indicate the efficiency of the new algorithm. OWPMR for the purpose of solving the following sub matrix uses the Greedy PTG method, since this method also improved the method of PTGand one of the tree methods for solving HIPP. OWPMR is from order O (sm2n2), when m is number of SNP and n is number of Haplotype.
Expression Profiling of Genes Involved in Muscular Activity and Inactivity in Hind Legs of Rats by Comparative Transcriptomics

Ghaderi-Zefrehei M.

Department of Agriculture, Yasouj University, Iran

DNA microarray technology allows simultaneous determination of the expression level of thousands of genes. It creates rich information in terms of the expression profile of each gene under study. In this study, data involving gene expression profile of rat’s hind limb muscle in three different experiments was used to investigate differentially expressed genes. The first experiment included 35-days-cage-controls versus 35-days-suspension, the second; experiment consisted of 14-days-cage-controls versus 14-days-suspension and the third experiment involved 14-days-suspension + 1 day-reloading versus 14-days-suspension + 5 days-reloading. In total, matrix of expression of all three experiments contained 1184 genes profiled in 39 arrays. A nested linear model with an appropriate design matrix was defined in LIMMA for gene expression. Different contrasts were carried out within different and across all experiments to find genes which showed different expression profiles. Moderate t-statistics was used to indicate mode of differentially expressed genes and rank them. Parallel inference for genes posed the course of dimensionality in terms of adjustment for multiple testing, e.g., control family-wise error rate (FWE) or probability false discovery rate (pFDR). Three thresholds (cut off) for of FDR (5% #10% #20%) were used to obtain with differentially expressed genes. The results indicated that the third experiment led to many number of genes, which demonstrated differential expression, pattern when compared to either experiment one and two. At the 5% FDR threshold, in experiment three, 138, turned out to be differentially expressed which majority of genes turned out up-regulated style (73 genes). In experiment three, mode of the up-regulated style was dominant at the same level of threshold with respect to other experiments. Just one gene (which encodes Glutathione-S-transferase) showed a differentially expressed pattern over all three experiments, at a20% FDR threshold. These results show that based on moderated t-statistics criterion, the third experiment (14 days-suspension + 1 day-reloading versus 14 days-suspension + 5# days-reloading) had profound effects on the expression pattern of all probs. These different experiments affected genes, which are mostly likely involved in ATP production and fatty acid metabolism. Hence, the implications of these results are that the supply of energy source is of great importance for strenuous work and exercise when compared to protein and other mediator nutrient molecules. These results have significant impacts on adaptive animal and human physiology.
SA-PSOMF: An Improved PSO-Based Algorithm for Pattern Discovery with Simulated Annealing

Zare-Mirakabad F. 1, Nowzari Dalini A. 1,2, Ahrabian H. 1,2, Hashemifar S. 2

1 Center of Excellence in Biomathematics, School of Mathematics, Statistics, and Computer Science, University of Tehran, Tehran, Iran.

2 Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran

The task of transcription factor binding site discovery from the upstream region of gene, without any prior knowledge of what look likes, is very challenging. In this paper we propose a modified Particle Swarm Optimization (PSO) with Simulated Annealing (SA) technique to identify pattern instances in multiple biological sequences. The experimental results on yeast Sachoromyces Cerevisae transcription factor binding sites, demonstrate that the proposed SA-PSO has a better ability to jump out a local optimum and is more effective than conventional PSO. Presented method is working analogous to YMF, MEME and AlignACE algorithms.
A Novel Evolutionary Algorithm Based Method for Efficient Quantitative Trait Loci Mapping

Ghaffarizadeh A., Eftekhari M., Esmailizadeh A.

Young Researchers Club, Azad Arak University, Arak, Iran

In this paper a novel approach for mapping quantitative trait loci (QTL) is introduced. This approach uses one of the recently developed evolutionary algorithms (EAs) called extinction EP (EEP) algorithm as a search technique to select important markers of chromosomes in a QTL mapping problem. In this algorithm information about inclusion or exclusion of markers used for estimating traits are coded into the chromosomes of EEP. In this study, each column of recorded data is a chromosome marker that is considered as a feature, which is relevant or irrelevant to qualitative traits (last column of data). Two subset selection criteria, ridge regression (RR) and partial least squares regression (PLSR) are employed separately as fitness criterion in EEP in order to select the best subset of features. Finally, the intersection of two subsets obtained in the previous steps is presented as the final result of our proposed method. The results show that obtaining the intersection of two aforementioned criteria gives a better accuracy in terms of yielding relevant markers. Experimental results on a QTL dataset with 99 chromosome markers in which 7 markers are the most important ones, confirm the efficiency of our proposed method. The approach will be useful in the context of genomic data sets that contain a large number of highly correlated features.
A New Algorithm for Computing Sequence Similarities
Engareh M. ¹, Shiri M. E. ², Ghaeminejad S. ³, Mohajerani H. ¹

¹ Azad University of Arak, Arak, Iran
² Amirkabir University of Tehran, Iran
³ Tafresh University, Tafresh, Iran

Many algorithms have been developed to search protein sequences. Which use various methodologies and protocols for computation of sequence similarities? This paper deals with mobile frame pattern algorithm, which is designed and offered with dynamic programming in order to compute secondary structure of proteins. Unlike BLAST, an alignment search tool the output of which is parsed hardly for finding sequence similarities, the proposed algorithm is a faster and more effective alignment concept for finding similar sequences in proteins and nucleotides. The proposed algorithm finds all possible similarities between two sequences through frame pattern detection, comparing patterns with each other and finally determines similarity percentage of sequences. Time complexity of the algorithm is equal to O (N²) at the worst condition. The algorithm was implemented in visual basic (VB) and its output was a detailed report. The program is accessible through a site, as well.
A New Scoring Function for Discrimination of Native Structures from Decoys

Mirzaie M. ¹, Sadeghi M. ², Eslahchi C. ¹, Pezeshk H. ³

¹Shahid Beheshti University, Tehran, Iran
²National Institute of Genetic Engineering and Biotechnology, Tehran, Iran
³Tehran university, Tehran, Iran

Three dimensional structure of a protein determines its function. This structure is determined by its amino acid sequence. The current approaches to protein structure prediction are based on thermodynamic hypothesis according to which the native state of protein is at the lowest free energy state. Thus, it is expected that energy of a native structure to be less than that of a decoy structure. In this study, we introduce a new method based on force in fold recognition. The force is calculated by the energy between atom pairs. We use knowledge based energy function. Such energy functions are obtained by the use of experimentally determined structures. After calculating the energy function, we see that these functions are U shaped in some regions and so we can assume that the energy function follows Hook’s law. The force is calculated using the relation between force and energy. We compare the force imposed to each atom of a protein with the corresponding atom in the other structures. We then assign larger scores to those atoms with lower forces. The native structure is assumed to be the one with the highest score in the data set. To evaluate the performance of our model, we apply it to several decoy sets. We also compare the results of our model to those obtained by the minimum energy model.
Understanding the dynamics of biological networks is one of the main problems in biological systems. Several groups have produced a large amount of data with regard to protein interactions. Recently, high-throughput methods have been developed to obtain a global description of the interactome. The protein interaction network can be represented as a graph where nodes correspond with proteins and edges with pairwise interactions. An inevitable result of this wealth of data is the need for efficient and accurate automated tools to predict protein complexes. Several clustering methods have been applied to the protein interaction graph to detect highly connected sub graphs and predict them as protein complexes. These algorithms rely on very different approaches. Each of them requires specifying several parameters, some of which may drastically affect the results. In this study, we give a short conceptual description of five algorithms: Markov clustering (MCL), restricted neighborhood search clustering (RNSC), protein complex prediction (PCP), clustering based on maximal clique (CMC), and k-connected finding algorithm (CFA). We also present a comparative assessment of these algorithms. For such a comparison, a test graph was built on the basis of 1142 complexes annotated in the Mips database and 62 complexes were annotated in the Aloy database.
Identification of Novel Potential BACE1 Inhibitors as Potential Therapeutics for Alzheimer's Disease through Virtual Screening and Artificial Neural Network Methods

Dehghani Amirabad A., Khodagholi F.

Neuroscience Research Center, Shahid Beheshti University of Medical Science, Tehran, Iran.

Secretase, also called BACE1 is an aspartic-acid protease important the the pathogenesis of Alzheimer’s disease (AD), and therefore one of the promising drug targets in therapeuting AD. In this study, we introduced three BACE1 potential inhibitors by means of computer-aided drug design protocols involving virtual screening with docking simulation and artificial neural network (ANN) methods. A library consisting of 14000 compounds was generated through similarity search with 80% similarity to GSK-188909, one of the BACE1 inhibitors that has low IC50 and good in vivo activity. This library was screened against the X-ray crystal structure of the BACE1 enzyme through docking with the FlexX software and 13 hits were obtained, which could bind the enzyme with a lower free binding energy than that for GSK-188909. In addition, the ability to pass the blood brain barrier is one of the important factors in neurodegenerative diseases like AD. However, approximately %99 of neuroprotective agents do not have this ability. So, we used the ANN procedure to build a powerful model for prediction permeability of new BACE1 inhibitors by the NeuroSolution 5.0 software. Data showed that compounds with the pubchem IDs: 21085486, 21085569, 11421009, with molecular formula of: C23H26N2O, C19H24F2N2O, C29H40F2N2O4, respectively, are promising candidates for the development of new drugs against AD because of possessing simple structures, having low free binding energies to BACE1, and strong ability to penetrate through the blood brain barrier (logBB - common descriptor to measure permeability of drugs through BBB - was -0.0017, -0.03, 0.87 for each of the three above compounds, respectively).
Finding Motif in DNA Sequences Using Hidden Markov Model

Shaterzadeh-Oskuei Sh., Nowzari-Dalini A., Ahrabian H., Zare-Mirakabad F.

Center of Excellence in Biomathematics, School of Mathematics, Statistics, and Computer Science, University of Tehran, Tehran, Iran.

Unraveling the mechanisms that regulate gene expression is a major challenge in Biology. An important task in this challenge is to identify regulatory elements, especially the binding sites in DNA for transcription factors. These binding sites are short DNA segments that are called motifs. Recent advances in genome sequence availability and high-throughput gene expression analysis technologies have allowed for the development of computational methods for motif finding. As a result, a large number of motif finding algorithms have been implemented and applied to various motif models. In this paper, we propose a new algorithm, which is capable of extracting a set of motifs without any prior knowledge, from a number of functionally related DNA sequences, using a DNA hidden Markov model. In the literature, the well-known presented program in this respect is called YEBIS, which uses the hidden Markov model for finding motifs. This program applies consensus sequences for constructing the hidden Markov model. Since, representing motifs in the profile model is better than consensus sequences, we use profiles, instead. This algorithm is implemented and tested on different types of real datasets. The results are compared with YEBIS and the effectiveness of our algorithm is shown.
Are There any Differences between Feed-Forward and Recurrent Neural Networks in Predicting Thermostable Protein Temperature?

Lakizadeh A., Ebrahimie M., Ebrahimie E.

Bioinformatics Research Group, Green Research Center, Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz.

Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran
Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran

Engineering thermostable enzymes have received many research interests due to their significant applications in various industrial processes. Thus, understanding the features involved in thermal stability is very important and here we compared two neural networks modeling (feed-forward and recurrent) approaches with regard to 2938 enzymes consisting of mesoand thermophilic proteins, to determine the features involved in this process. To avoid overlap between training and testing sets, we randomly divided records into 10 parts each consisting of 294 records (except the last group with 292 records). The Processes were repeated 10 times and the accuracy for each repeat and total accuracy were calculated. To investigate the effects of the feature selection on the neural networks behavior, all models were also run with feature selection (stepwise regression criteria). The results showed that in the feed-forward neural network, the best overall accuracy (0.90) was obtained when the hidden layer was 3 and the number of neurons in the input layers was 50, 20 and 10. of hidden layers was 3, while its accuracy in predicting true, false and overall records were 0.824, 0.935 and 0.892, respectively. In recurrent (Elman) neural networking, the figures were 0.809, 0.946 and 0.894, respectively. The results showed there is no significant difference (p > 0.95) between the feed-forward and Elman neural networks and no difference found when stepwise regression feature selection was used. When the best networks were run on another dataset of 60 proteins with known temperatures, the best accuracy in determining the right temperature (97.83%) was obtained in the Elman network and the worst one was in the recurrent with 2 hidden layers. The findings confirmed that both networks are suitable for determining the temperature of thermostable proteins and feature selection can be used to lower the burden of the system.
A Fast GRAPH Query Algorithm in Biological Network
Pashami S., Ghiasvand F. A., Ahrabian H., Nowzari Dalini A.

Center of Excellence in Biomathematics, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran.

Studying biological networks could raise new insights in biological systems. One of the interesting problems in this field is finding a sub-network with known structure such as protein interaction network. This problem is known as graph query problem. In this paper, a novel algorithm for solving graph query problem has been presented. Regarding structure of biological networks, pruning techniques developed to improve time performance of presented branch and bound algorithm. Another version of the graph query problem is searching for the magnitude amount of sub-networks in a given biological network. Since all sub-networks are searched in one biological network, presented algorithm optimized, by saving intersection parts of several runs of algorithm. In this respect we combine all subgraphs in one decision tree and all inputs through one run of the algorithm are processed. As a result, the presented algorithm has great improvements in comparison with existing algorithms. One of the applications of this algorithm is searching for motifs in biological networks. Using this algorithm, A software developed for querying candidate motifs in random networks. The software evaluated on biological networks of E. coli and S. cerevisiae, and also on non-biological networks: social and electronic network. The performance of the software is better than all previous methods up to size 7, and in opposite of many other algorithms because of polynomial memory space of the algorithm, finding large motifs is possible with this software.
Creating an Exact Model of 3D Protein Structure for in Silico Site Directed Mutation
Taghizadeh M., Minuchehr Z., Madadkar-Sobhani A., Goliaei B.

Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran.

In the protein data bank (PDB), there are many 3D protein structures that are a result of site-directed mutagenesis studies and there are also wild type 3D structures available for them. For such structural studies, researchers need to have a closely similar structure and for this reason crystallography of these mutated structures should be carried out. So there is a need for finding a protocol to create very similar models for those studies. As a first step, in the way of finding such a protocol, we modeled lysozyme mutated proteins for which 3D structures were available, using wild type lysozyme as template. PDB IDs of the wild type lysozyme and mutated ones, which were being used were 1REX, 2GQI, 2BQM, 2BQK and 2BQH. In this study we used MODELLER and created 100 models for each mutated protein and selected the best models based on distinct evaluating functions: DOPE-Score and MolPDF. Then we used two different side chain modelers, SCWRL and OPUS-ROTA, for getting better in side chain modeling. For 16 generated models, we calculated RMSD in comparison with related crystal structures. This study shows that RMSDs of models chosen with MolPDF and those chosen with DOPE-Score are relatively equivalent. Average RMSD for back bone of the MolPDF selected models was 0.37 Angstrom. Average RMSD for side chains was 1.16 Angstrom and for all atoms was 0.85 Angstrom. The effect of side chain modelers was not positive and with the use of such programs the RMSD of side chains increased as a result of very high homology of the query and template. MODELLER could find side chain conformations better than the side chain modelers. A previous study on crystal structures in 2005 shows that there is a RMSD variation of approximately 0.84 Angstrom between similar structures that are crystallized in different environments, which is also near to our RMSDs. Accordingly, the use of homology modeling instead of crystallography for the above mentioned cases seems to be rational.
Propagation of Information in the Genetic Networks

Roshani F.

Department Physics, Alzahra University, Tehran, Iran

Random Boolean networks (RBNs) were introduced as a simple model for gene regulatory networks. This network is specified by its topology and its dynamical rules. The topology is given by the nodes and the links between these nodes. The dynamical rules describe the time evolution of the state of the nodes. The state of a node is 1 or 0 and it is a function of the state of the node that has a common link to it. There are three phases for these networks: frozen, chaotic and critical. One can study the propagation of external information given by regulatory input variables into a random Boolean network. In particular, this allows us to identify variables which are completely determined by this external information. All of those variables in the network, which are not directly fixed, are called cores and the statistical properties of these cores are considered. Consequently we have introduced a genetic algorithm that is used in optimization problems and based on chromosomes and genes constructions. Using this algorithm we design a Boolean network that has a minimum core and obtain the statistics of the cores and time evolutions of the network.
Keynote 4:

The Roots of Systems Biology and Complex Recurrent Interactions: from Henri Poincaré to Robert Rosen

Karami A., Gifani P.

Research Center of Molecular Biology, Baqyatalah University of Medical Science. Tehran, Iran.

Complex Systems & Cybernetic Control Lab, Department of Biomedical Engineering, AmirKabir University of Technology, Tehran, Iran.

Understanding the structure and dynamics of the complex intercellular web of interactions that contribute to the structure and function of a living cell is the challenging issue for modern biology in the twenty-first century. Most biological characteristics emerge from complex interactions among the cell’s ingredients. There are many evidences of complex systems both individually and functional whole in different hierarchies of molecular biology. Various network types of interaction are not independent; they form a network of networks that is responsible for the behavior of the cell globally.

Systemic view of biology leading to systems biology era of life science is not a new approach in scientific discipline. By checking out the actual roots of systems biology you can find the strong backbone of this interdisciplinary research field which now attracts many people from scientific and technological sectors.

In the systems biology era of the life sciences conceptual guidelines become more important than experimental or computational issues. Recent studies on complex molecular networks, give the impression that they could have a modular hierarchy sets of interacting genes and proteins that carry out specific tasks. On the other hand when such modules interact with each other, a system level dynamics will be emerged which couldn’t be seen in any sub-system individually. From systems science point of view and by the use of nonlinear dynamical system theory, it means that each module has its own dynamics and interaction manifolds. When they hierarchically hook together, depending on different stable and unstable attractors of each module, a new organization of interaction with intertwined interacting region will be emerged which represent a complicated higher level complex manifold. Therefore, conceptual methodology to investigate how such higher level complex manifold emerge from integration of lower level modules is the goal of systems biology.
We discuss about the actual systems biology roots form complexity and nonlinear dynamical system point of view which explain how systems biology, falls directly in the line of reflection carried out by Henri Poincare and Robert Rosen throughout their recurrent approach. This complex recurrent approach will be argued to figure out how simple signaling pathways can be embedded in networks using positive and negative feedback to generate more complex behaviors bring a distributed control mechanism and may show properties that are difficult to predict.

The far from thermodynamic equilibrium phenomena and higher order balances between opposite processes which one process could generate order (negative entropy or negentropy) provided it was coupled to a second process that produced more disorder (entropy) will be discussed to explain why such coupling or integration of processes is consequently essential for life.

This is the recurrent dynamic embedded information organization of living systems which bring out self-organization phenomena to select a trade-off between robustness, fragility, performance, adaptability and resource demands. Consequently, any advancement in the general understanding of complex recurrent interaction will have a prospective impact in preparing a conceptual strategy for systems identification in systems biology era of life science.
Modeling and Implementing an Agent-Based System for Prediction of Proteins Relative Solvent Accessibility

Meshkin A. 1, Ghasem Aghaee N. 1, Sadeghi M. 2

1 Department of Computer Engineering, University of Sheikh Bahaei, P.O. Box 81797-35296, Isfahan, Iran.

2 National Institute for Genetic Engineering and Biotechnology, P.O. Box 14155-6343, Tehran, Iran.

In this paper, an agent based system for prediction of relative solvent accessibility (RSA) of proteins is proposed. Since, it is believed that the 3D structure of most proteins is defined by their sequences, utilizing data mining methods to extract hidden knowledge and information from protein sequences, is unavoidable. Due to the inherent heterogeneity and distribution in data that are used to predict RSA, time consumption and high costs of central data mining on large training data, the necessity of an agent based architecture for predicting RSA, seems to be essential. The system is logically and functionally divided into four layers, solving the tasks of "data fusion", "feature selection", "model building" and "knowledge discovery and prediction". The outcomes of the system design phase under the Prometheus methodology and the complete Characteristics of the agents are discussed. The prediction activity results from the interaction of a set of agents that have been hosted on several levels. The experimented results on the Manesh dataset point to the validity of the approach. The proposed system can autonomously dig out all the valuable knowledge about which physicochemical features are highly correlated with the solvent accessibility of proteins without human supervision, which is of great importance to biologists and their future researches.
Prediction of Relative Solvent Accessibility with Support Vector Regression Based on Qualitative and Quantitative Features Selection

Meshkin A.

Department of Computer Engineering, University of Sheikh Bahaei, P.O. Box 81797-35296, Isfahan, Iran.

As for the vital role and functionality of proteins in the human body and functional dependency of a protein to its three dimensional (3D) structure and immense time and economic costs of experimental methods in determining protein structure, predicting relative solvent accessibility (RSA) has great importance, because it leads to better prediction of the native structure of a protein. From a logical standpoint, the proposed method can be divided into two main steps: the first one provides subset selection of quantitative features based on selected qualitative features and the second, dedicated to training a model with selected quantitative features for RSA prediction. The experimental results on the Manesh dataset show that the proposed method is improvement in the average prediction of accuracy and training time. The proposed method can dig out all the valuable knowledge about which qualitative features are deemed more significant in prediction of RSA with respect to the position of residues in a given window, without human supervision, which is of great importance for biologists and their future researches.
Fractal Analysis of DNA by Nonlinear Genome Signal Processing for Exon and Intron Separation

Gifani P. 1, Karami A. 2

1Complex Systems & Cybernetic Control Lab, Department of Biomedical Engineering, AmirKabir University of Technology, Tehran, Iran
2Baqiyatallah University of Medical Sciences, Tehran, Iran

One of the structural features of the whole genome is the long-range correlation or scale-invariant property of DNA. This phenomenon implies that occurrence of a small segment of nucleotides depends on large scale segments. Such long-range correlation is directly related to power-law and fractal structure of the DNA sequence. In this paper, we investigated this global feature of the DNA sequence by calculating the fractal scaling exponent of the numerical signal which is produced by converting a DNA string to a numerical value by a sequence to number mapping algorithm. By this approach, we have a numerical signal for a DNA string which could be analyzed by signal processing algorithms. Based on the fractal structure of the DNA sequence, we implemented detrended fluctuation analysis to calculate the fractal scaling exponent of this signal. The results implied that optimal fractal scaling exponent (FSE) of coding sequences was significantly lower than sequences that were primary noncoding, indicating the presence of long-range correlations in functional sequences. The statistical property of results represents meaningful separation of two groups of exons and introns in a database of 195 genes. This study show that optimal FSE for both exons and introns indicate the presence of long-range correlations and the fractal nature of the genome signal. More importantly, the FSE of coding sequences (exon) was significantly lower than sequences that were primary noncoding (intron).
Comparison of Different Softwares for Haplotype Inference Based on Number of Genotypes Required for a Specific Accuracy

Ferdosi M. H. ¹, Esmailizadeh A. ¹, Sadeghi M. ², Katanforoush A. ³

¹Shahid Bahonar University of Kerman, Kerman, Iran
²National Institute of Genetic Engineering & Biotechnology (NIGEB), Iran
³Institute of Biochemistry and Biophysics (IBB), Tehran, Iran

Since single nucleotide polymorphisms (SNPs) are responsible for 90 percent of differences among species and are used widely to evaluate biodiversity in animals, SNP diversity is more limited than haplotype diversity, so the latter is used to evaluate diversity in different species. Haplotype determination in the laboratory is very expensive, so in most cases genotypic data are used to infer haplotypes. Several algorithms are available for haplotype inference. The aim of this study was to estimate the number of genotypes required to get a specific accuracy in respect of the number of SNPs and methods of haplotype inference. For this purpose, many populations were simulated using MS software. These populations were varied in terms of the number of SNPs and individuals. Haplotypes were inferred with different softwares and haplotype error rate for each method was evaluated. These data were fitted to suitable a model by using the curve fitting toolbox of the MATLAB software. Numbers of genotypes needed to get a specific accuracy were estimated by using this model for each method. In this study, the different methods for haplotype inference based on number of genotypes needed to get specific accuracy were compared. When the number of SNPs and predetereminated accuracy were low, the difference between various methods was not considerable, but in the case of, high number of SNPs and high accuracy, the difference between the methods increased.
Bioinformatics Study of Phenylacetaldehyde Synthase (PAAS), a Protein Involved in Flower Scent Production, in Rosa Hybrida and Petunia Hybrid

Karimi A., Ashrafi E., Ebrahimie E.

Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran

Rose flowers produce and emit the aromatic volatiles 2-phenyl acetaldehyde (PAA) and 2-phenyl ethyl alchohol (PEA), which have a distinctive flowery/rose-like scent. Previous studies of rose have shown that, similar to Petunia flowers, PAA is formed from L-phenylalanine via pyridoxal- 5'-phosphate-dependent L-aromatic amino acid decarboxylase. Rosa hybrida cv. Fragrance Cloud sequence (RhPAAS) is homologous to Petunia phenylacetaldehyde synthase (PhPAAS). Since there is not much experimental data available about different structural properties of that PAAS protein, in the present investigation, we studied the different structural properties of the PAAS protein in rose and petunia using bioinformatics tools. The features of the first, secondary and tertiary structures of this protein were compared between Petunia and Rose. The results indicated that the frequency of negatively charged, positively charged, leucine, and frequency of the ser-leu, pro-glu, phe-ser and the, thr-thr dipeptides in petunia are more than those in rose. In contrast, in petunia, the frequencies of hydrophobic and hydrophilic residues, α-helix, β-sheet, β-strands of petunia are lower than those in rose. The features achieved in this study may also provide useful clues for designing scent production pathways using protein engineering techniques.
Digital DNA

Kashani M.

*Digital Clone Company, Irland*

Philosophically, the invention can be best appreciated by studying the works of Charles Darwin. See, Darwin's Dangerous Idea, Daniel C. Dennett, New York, 1996. Applying the theory that Darwin formulated for his explanation of the Origin of Species, it is apparent that this concept can also be used to describe so many different people having so many different needs and strategies that must be met by their respective computers and software. The principle of natural selection is the unifying insight into the inventor's approach to his information processing apparatus and method.

By having a large variety of inheritable skills (through storage in computer systems), which constitute recorded procedures in the invention; these different procedures will tend to have different payoffs for different individuals and subgroups of the user population. Under Darwinian Theory, these different individuals and subpopulations would tend to diverge, each pursuing their favored sort of excellence until, eventually, there is a distinct division. With biological systems, transference of inheritable characteristics implemented by DNA mutations and modifications is possible only from parent to child and even then, is frequently "hit or miss". Mutation in this field is often viewed as random, unstructured and survival appears to be by pure chance and by no means permanent, as far as a species is concerned.

However, by using the inventor's methods, this random and unstructured divergence is eliminated since the entire optimized skill set is "inheritable" by one computer communicating with another computer or with the user.

In essence, this method provides a new modality for increasing man's knowledge exponentially in volume and in speed of transference from one individual or subpopulation to another (Digital DNA).
Which Features Are Responsible for Halolysin Proteins Halostability?

Ebrahimie E. 1, Ebrahimie M. 2, Rahpayma Sarvestani N. 2

1 Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran

2 Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran

Extreme halophiles not only tolerate but require high concentrations of salt for optimal growth. Understanding the halophilicity mechanism of extremely haloarchaeal is an inevitable task and is the first step toward engineering halostable crops. Hence, we examined protein features were contributing to the halo-tolerance of halophilic organisms. More than 400 features compared for halophilic and non-halophilic proteins using various screening, clustering, decision tree, generalized rule induction models, and molecular phylogenetic relationship procedures to search for patterns of salt stability. The N-terminal amino acid of all halophilic halolysins was met, whereas other amino acids were found in other proteases and termitases. Eighty-three protein features were shown to be important in feature selection model-ing, and anomaly index (with one peer group) of 2.42 declined to 1.87 with only important features; however, no changes were found in the numbers of groups when K-means and twoStep clustering modeling were performed on datasets with/without feature selection filtering. The depth of the decision trees models varied from 1 (in C&RT with/without the feature selection) to 5 (in C5.0 with 10-fold cross-validation branches). The performance evaluation of decision tree models had the same values and the best correctness percentage was recorded with the C5.0 and CHAID models. We did not find any significant difference (p < 0.95) in the percent of correctness, performance evaluation, and the mean correctness of various decision tree models with/without feature selection, but the number of peer groups in the clustering models was reduced significantly.
Application of Bioinformatics Tools in Protein Engineering; Presenting a Few Applications Employed in Our Laboratories

Ebrahimie M. 1, Ebrahimie E. 2

1Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran.
2Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran.

Bioinformatics uses various algorithms and methods to explore huge amounts of biological data in order to help us understand biological mechanisms. In our labs, research groups use bioinformatics tools to investigate and understand why certain biological processes are working and what are the most important features contributing to their functions. Of special importance to our researchers are some enzymes and proteins responsible for salinity and drought stresses and thermostability. Different approaches have been employed, which can be classified as follows: a). statistical analyses to understand the significant differences among normal and desired proteins (halophilic or thermostable), b.) Feature selection algorithms to define the most important features contributing to desired protein activities, c). Neural network modelings and tools to train and test different networks in order to correlate between features and protein characteristics and use these networks to predict desired abilities. The results of some research groups have been presented briefly here.
A Precise Model for Prediction of in Vitro Plant Somatic Embryogenesis Based on the Microarray Gene Expression Data

Baseri S. ¹, Ebrahimie E. ², Ebrahimie M. ³

¹Department of Statistics, College of Statistics, Shiraz University, Shiraz, Iran

²Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran

³Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran

Microarray is a novel technology facilitating simultaneous measurement of thousands of gene expression levels. Microarray data process is a critical stage requiring the novel statistical methods which can deal with a large number of genes (variables, up to 30000) and small number of replications. In this situation, the statistical methods based on the abundance of the observation are not applicable. Modeling and prediction of the microarray is one of the major applications of statistics in field of science. This prediction and modeling allows scientists to smartly manipulate to appear the phenomena based on gene expression profiling. Using modified empirical estimator (Crude), the microarray data of somatic embryogenesis in soybean were carried out. Also a predictive modeling for percent of somatic embryogenesis response of explants based on genes expression was presented.
Are There any Differences between Features of Proteins Expressed in Malignant and Benign Breast Cancers?

Shamabadi N., Ebrahimie M., Ebrahimie E.

Bioinformatics Research Group, Green Research Center, Department Of Crop Production & Plant Breeding, College Of Agriculture, Shiraz University, Shiraz, Iran.

Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran.

Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran.

Breast cancer as the most common cancer among women is the second leading cause of death in women; So far many approaches have been used to analyze and detect benign and malignant forms of cancer and thus understanding the features involved in proteins expressed by various types of breast cancers is crucial. Herein, we compared features of proteins expressed in malignant, benign cancers using various screening techniques (anomaly detection, feature selection), clustering methods (K-Means, TwoStep cluster) decision tree models (C&RT, CHAID, Exhaustive CHAID, QUEST, C5.0), and generalized rule induction (GRI) models to search for patterns of similarity in each group. We found that in all proteins the N-terminal amino acid was Met and 57 out of 838 proteins’ features ranked as important (p > 0.05) in feature selection modeling. The number of peer groups was 2 with 1 anomalous record in each group and no changes were found in the numbers of clusters when K-Means and TwoStep clustering modeling were performed on datasets with/without feature selection filtering. The depth of the trees generated by various decision tree models varied from 5 (in the Quest model) to 2 (in the C5.0 model) branches. The performance evaluation of the decision tree models tested here showed that C&RT was the best and the CHAID model was the worst. We did not find any significant difference in the percent of correctness, performance evaluation, and mean correctness of various decision tree models when feature selected datasets were used, but the number of peer groups in clustering models was reduced significantly.
Determining the Characters of Thermostable Xylanase Enzyme

Aghagolzadeh P. 1, Ebrahimie E. 2, Ebrahimie M. 3,

1Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran

2Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran

3Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran

Xylanase enzymes have received attractable researches as environmental friendly proteins to reduce the amount of persistent organic chemicals used for bleaching paper pulp. However, such bio-bleaching requires enzymes with better thermostability and understanding the features involved in enzyme thermostability is crucial. Herein we examined features contribute to the thermostability of all xylanase proteins by using various screening techniques, clustering methods, decision tree and generalized rule induction models. 181 (from 1775) features shown to be important in feature selection modeling, no change found in the number of peer groups, clusters and decision tree branches but the number of anomalous records decreased in this modeling. No significant difference found in the percent of correctness, performance evaluation and means correctness. When the value of the count of Ile- Pro was less than 1.500, the average optimum temperature of xylanase enzymes was less than 57.47 °C, otherwise the temperature was higher than 77.68 °C. The frequency of His - Cys (value of 0.002) and the frequency of Glu – Ile (value of 0.004) used to classify the proteins’ optimum temperature less than 57.47 °C and higher than 77.68 °C, respectively. One of the decision tree models used to predict optimum temperature of 30 other proteins with known temperature with the overall accuracy of 91%, while its accuracy in predicting temperatures less and higher than 57 °C were 84% and 95%, respectively. The models used here are powerful tools to determine xylanase thermostability features and paves the road to engineer new thermostable mutants.
Racial Classification Using Cephalic Size and Support Vector Machines

Esmaeilzadeh M.

Azad University Shirvan Beranch.Iran

Congregate different racial and people immigration in cities in the additive recent century has caused the need for racial classification to identify genetic traits of each race. This article used support vector machine (SVM) to classify Fars race living in Mashhad and Shirvanian Kurmanj race. Face cephalic size is used as input parameters. For this purpose used the size of the 60 children of Kurmanj race and 49 children of Fars race in ages 11 to 14 years. Achieve 97 percent accuracy in the validation and accuracy of test results sorting, demonstrates the success of this method for correct diagnosis racial.
Prediction of Protein Stability Changes by Data Mining and Support Vector Machine

Bigdeli S. ¹, Madadkar-Sobhani A. ², Minaei B. ¹

¹School of Computer Engineering and Information Technology, University of Qom, Qom, Iran

²Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran

³Department of Computer Engineering, Iran University of Science and Technology, Tehran, Iran

Prediction of protein structures and their properties is one of the imperative problems for researchers in the field of structural bioinformatics. One particular case is the prediction of protein stability upon structural changes, which is sought in designing new proteins and their analogues. In order to design stable mutants, accomplishing a high accuracy in this prediction of protein stability changes (ΔΔG) is crucial. Several methods for tackling this problem have been introduced in recent years. One common trait of all these methods is the exploitation of ProTherm, a consequential database of thermodynamic data on protein stability changes. In this study, we used support vector machine (SVM) to predict protein stability changes upon single mutation, based on the protein sequence information alone, which is useful in conditions that there is no three-dimensional structure available. Using ProTherm as database and data mining methods, we built a model and a dataset based on SVM, which can predict sign and value of ΔΔG separately. Cross-validation was used to compare our method over previous ones. A correlation of R²=0.76 between our predicted values and experimental data, and accuracy of 86% for identifying the sign of ΔΔG was achieved.
Tuesday Posters
Microarray Data Analysis for Detection and Classification of Viral Infection

Nazari K. 1, Karami A. 2, Mahdavi Amiri N. 3*

1 Faculty of Mathematics, Sharif University of Technology, Tehran, Iran
2 Bbaqyiatallah University of Medical Sciences, Tehran, Iran
3 Faculty of Mathematics, Sharif University of Technology, Tehran, Iran

DNA microarrays consist of DNA microscopic points that are attached to a solid surface such as glass, plastic or silicon chip and formed an array. The pieces of fixed DNA considered as a searcher. In an expriment, we can use thousands of searchers. Therefor, any microarray consists of the same number of genetic tests that the exprrement performed for all of them in parallel. Whit this ability, arrays have speeded up to the biological investigations. Microarray technology can be seen as a continued development of southern blotting.

However, the most important stage in this technology, analysis of data that is gettig and we are in need reliable bioinformatics tools for the analysis of such data whith high certitud degree.

Infectious diseases, from the begining of human life always have been with human and have preduced the basic difficulties for him. One of the most important usages of microarray thechnology is the possibility of testing for the presence of thousands micro-organism in environmental and clinical samples only in a single exprrement that is resulted in recognition in good time of pathogens. Thereby we take an important step in cureing of diseases. As noted above, the most important stage in microarray is data analysis. We present E-Predict algorithm and DetectiV package, for microarray based species identification. We demonststrate the application of E-Predict and DetectiV to viral detection in a large, publicy available dataset and show that DetectiV performs better than E-Predict. DetectiV is implemented as a package for R - powerful, open source software for statistical programming - that containing visualization, normalization and significance testing functions.
Complex Recurrent Interactions in Systems Biology: from Henri Poincare to Robert Rosen

Karami A. ¹, Gifani P. ²

Research Center of Molecular Biology, Baqyatalah University of Medical Science. Tehran, Iran
Complex Systems & Cybernetic Control Lab, Department of Biomedical Engineering, AmirKabir University of Technology, Tehran

Understanding the structure and dynamics of the complex intercellular web of interactions that contribute to the structure and function of a living cell is the challenging issue for modern biology in the twenty-first century. Most biological characteristics emerge from complex interactions among the cell's ingredients. There are many evidences of complex systems both individually and functional whole in different hierarchies of molecular biology. Various network types of interaction are not independent; they form a network of networks that is responsible for the behavior of the cell globally. Systemic view of biology leading to systems biology era of life science is not a new approach in scientific discipline. By checking out the actual roots of systems biology you can find the strong backbone of this interdisciplinary research field which now attracts many people form scientific and technological sectors. In the systems biology era of the life sciences conceptual guidelines become more important than experimental or computational issues. Recent studies on complex molecular networks, give the impression that they could have a modular hierarchy sets of interacting genes and proteins that carry out specific tasks. On the other hand when such modules interact with each other, a system level dynamics will be emerged which couldn't be seen in any sub-system individually. From systems science point of view and by the use of nonlinear dynamical system theory, it means that each module has its own dynamics and interaction manifolds. When they hierarchically hook together, depending on different stable and unstable attractors of each module, a new organization of interaction with intertwined interacting region will be emerged which represent a complicated higher level complex manifold. Therefore, conceptual methodology to investigate how such higher level complex manifold emerge from integration of lower level modules is the goal of systems biology. We discuss about the actual systems biology roots form complexity and nonlinear dynamical system point of view which explain how systems biology, falls directly in the line of reflection carried out by Henri
Poincare and Robert Rosen throughout their recurrent approach. This complex recurrent approach will be argued to figure out how simple signaling pathways can be embedded in networks using positive and negative feedback to generate more complex behaviors bring a distributed control mechanism and may show properties that are difficult to predict. The far from thermodynamic equilibrium phenomena and higher order balances between opposite processes which one process could generate order (negative entropy or negentropy) provided it was coupled to a second process that produced more disorder (entropy) will be discussed to explain why such coupling or integration of processes is consequently essential for life. This is the recurrent dynamic embedded information organization of living systems which bring out self-organization phenomena to select a trade-off between robustness, fragility, performance, adaptability and resource demands. Consequently, any advancement in the general understanding of complex recurrent interaction will have a prospective impact in preparing a conceptual strategy for systems identification in systems biology era of life science.
Introduction and background AIDS as a worldwide concern has been the centre of focus in many studies. One possible approach to controlling this disease can be through disturbing the retroviral cDNA structure by adenosine deaminase enzymes leading eventually to the termination of viral replication. One of these enzymes is APOBEC1 named after its role in editing the mRNA for the protein ApoB. Although, 3D structure of this protein has not been empirically identified. In this study, the 3D structure of APOBEC1 has been modeled by the pdbviewer. Method APOBEC1 was aligned as the first step using BLAST program and the most similar protein was recognized as APOBEC 3G (PDB code E1U3) with more than 65% resemblance. The file for APOBEC1 and APOBEC 3G alignment was received using CLUSTALW. The alignment file was submitted to swissmodel and the pdb format of modeled file was obtained and loaded on to the pdbviewer in order to illustrate the 3D structure. Finally, the structure was energy minimized. Conclusion the 3D structure of APOBEC1 was defined with 6 helices and 7 strands and the total energy of the structure was decreased from 1310.395 kj/mol to -14405.844 kj/mol.
Prediction of Relative Solvent Accessibility Using Pace Regression

Meshkin A.

Department of Computer Engineering, University of Sheikh Bahaei, P.O. Box 81797-35296, Isfahan, Iran.

In this paper, a new approach for prediction of protein solvent accessibility is presented. The prediction of relative solvent accessibility gives helpful information for the prediction of native structure of a protein. In recent years, (RSA) several RSA prediction methods including those that generate real values and those that predict discrete states (buried vs. exposed) have been developed. We propose a novel method for real value prediction that aims at minimizing the prediction error when compared with existing methods. The proposed method is based on pace regression (PR) predictor. The improved prediction quality is a result of features of the PSI-BLAST profile and the PR method, because pace regression is optimal when the number of coefficients tends to infinity. The experimental results on the Manesh dataset show that the proposed method is an improvement in average prediction accuracy and training time.
Pattern Matching of Dynein Heavy Chain Family
Kheitan S. ¹, Minuchehr Z. ¹, Madadkar-Sobhani A. ², Darvishalipoor S.¹, Steipe B. ³

¹ National Institute of Genetic Engineering and Biotechnology, Tehran, Iran
² Bioinformatics Department, Institute of Biochemistry and Biophysics, University of Tehran, Iran
³ Department of Biochemistry, University of Toronto, Toronto, Canada

Dyneins are large molecular motors which drive a variety of functional cellular movements in eukaryotic cells by their heavy chain components. In this paper, all characterized dynein heavy chains (DHCs) are searched for conserved patterns. Discovered patterns were specialized to pick up cytoplasmic and axonemal DHCs in database scanning separately. To evaluate the quality of generalized patterns we used them as screening tools against Swissprot and TrEMBL databases and also constructing the phylogenetic tree. Unannotated sequences which matched to the conserved pattern can be supposed as putative DHCs. The identified pattern can use as the determinant criterion for DHCs and assign axonemal or cytoplasmic type of them.
Structure-Based New Antifungal Compounds Discovery
Ferdosiyan M. 1, Sardari S. 2

1 Department of Biochemistry, science and research branch, Islamic Azad University, Tehran, Iran
2 Drug Design and Bioinformatics Unit, Biotechnology Research Center, Pasteur Institute of Iran, #69 Pasteur Ave., Tehran, 13164, Iran

Background: Ligand-based drug design is one of the virtual screening methods of finding new hits and lead compounds via two different strategies: the similarity search approach and the pharmacophore search method. The "similar structure - similar property" principle states that structurally similar molecules are expected to exhibit similar properties or biological activities. A query compound is used to search and compare with every compound in a typical database to find those compounds that are most similar to it. It can be performed based on 2D and 3D descriptors. Similarity searching method requires two components: first, descriptors that can be used to compare molecules, and second, a similarity coefficient which provides a way of assessing the degree of similarity between two compounds. Many different similarity coefficients have been developed. The Tanimoto coefficient is the most widely used. In the present study, we decided to introduce several new antifungal compounds based on 2D fingerprints and Tanimoto similarity coefficient. Materials and methods: Amphotericin B is an effective antifungal drug to treat systemic fungal infection, thus it was selected as a query to perform 2D similarity search in PubChem using Tanimoto Index above 70%. The molecules which had less complexity than Amphotericin B were selected, some molecular descriptors that are important in the drug design process (such as simple counts properties, hydrophobicity, rotatable bond Ccount, topologica polar surface) were calculated from 2D structures using Chemsketch and ChemBioDraw software applications for amphotericin B and the selected similar molecules. Based on structure-property relationship (SAR) we selected the molecules which had more similar properties to amphotericin B, amongst the top-scoring compounds, to purchase and then we examined their effect on the growth of the test fungi. Result: the number of molecules that were obtained from 2D similarity search was approximately 2000 compounds. We selected 400 compounds with less complexity to calculate molecular descriptors. Then, we screened the compounds according to similar properties in our scoring system. Ultimately we chose 9 available molecules to determine their in vitro biological activity. Conclusion: Similarity searching offers several advantages: 1) We can control the size of the output. 2) A numerical score is given for every compound in the database to define output ranking. We have demonstrated that the selection algorithm is efficient as it yielded 4 active compounds among the 9 candidates tested on the growth of Candida albicans.
Boolean Modeling of T Helper Cell Gene Regulatory Network

Vaezpour S. Y., Dehghan M.

Amirkabir University of Technology, Computer Engineering and Information Technology Department, Tehran, Iran.

Recent remarkable progress in Molecular Biology has led to a complete map of the genomes of many organisms and the identification and classification of the proteins is well under way. The next major challenge is to determine all the interactions between genes, proteins and other cellular components and to integrate this knowledge into a system-level understanding. It is now widely recognized that the networks of interaction and regulation between cellular entities are highly complex and their understanding needs a concerted effort between experiment, modeling and theory. Genes and gene products interact and form networks on several levels. At the genomic level, a class of proteins called transcription factors can activate or inhibit the transcription of genes into mRNAs. Since these transcription factors are themselves products of genes, the ultimate effect is genes regulating each other’s expression by forming so called gene regulatory networks. Since the 1960’s, methods from mathematics and physics have been used to describe and simulate small gene networks more stringently. The simplest dynamic models – synchronous Boolean network models - were used as a model for gene regulatory networks. Boolean networks are based on the assumption that binary on/off switches functioning in discrete time steps can describe important aspects of gene regulation. In synchronous Boolean network models, all genes switch states simultaneously. We give a description of Boolean Modeling and then we focus on the T helper cell Boolean model.
The Utility of Chloroplast Matk Gene as Plant DNA Barcode for Phylogenetic Analysis in Zingiberaceae Family

Valiellahi E., Mirshamsi A., Kavousi H. R., Farsi M.

Department of Biotechnology and Plant Breeding, College of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

DNA barcoding is a taxonomic method that uses short orthologous DNA sequences to identify an organism as belonging to a particular species. This method has been proposed and initiated to facilitate biodiversity studies. The MaturaseK gene (MatK) of chloroplast is highly conserved in plant systematics which is involved in Group II intron splicing. The size of the gene is 1500 bp in length, located within the intron of trnK. of this study is to evaluate generic, species variation and phylogenetic relationships of Zingiberaceae family by using the chloroplast matK gene sequences available from genbank In the, MatK gene from Zingiberaceae was taken from the National Centre for Biotechnology Information (NCBI) for the analysis of phylogeny. The family of Zingiberaceae comprises 47 genera with medicinal values. The sequence alignments were performed by Vector NTI 9.0 and prediction of transition/transversion rates and phylogenetic analyses were carried out by the MEGA software. The objective. The results indicated that the Zingiberaceae genera Afromonum, Alpinia, Globba, Curcuma and Zingiber show polyphylogeny. These results support the claim that DNA barcoding is a powerful tool for taxonomy and biogeography with utility for identifying cryptic species, biogeographic patterns and resolving classifications at the of genus and species levels. From this study, it could be concluded that the matK gene is a good candidate for DNA barcoding of the plant family Zingiberaceae.
Study of Knowledge Status of Non-Physician Workers in a Health Care Center

Hoseinpourfard M.

Management research of Health School, Baqiyatallaah University of Sciences, Tehran, Iran

Introduction: Many organizations do have not enough information about individuals. Hence, in this regard, the role of health care centers and hospitals in the communities responsible for health promotion are increasingly important. Effective knowledge management by focusing on solutions that encompasses the whole system, including the most important tools for this information must be considered. Therefore, the use of knowledge management for increasing the efficiency of these centers as a tool for productivity would beneficial. Materials and Methods: This investigation involves a cross sectional study to evaluate status and level of knowledge about non-physician staff in a health care center. Data collection was conducted by a questionnaire. Average of values and weight were calculated for all questions and standards. The end result is a numeric between 1 and 5 that shows the position of knowledge by this description: less than 2: poor knowledge, 2 to 3: relative knowledge existence, 3 and 4: acceptance knowledge, and more than 4: desirable. Results: Among the non-physician staff, 72 were selected randomly and questionnaires were distributed among them. Number of completed questionnaires collected, was 47 cases (20 persons filled incompletely, and 5 did not fully answer the questions). Total score of the observed health care center was equal to 2.77, namely the non-physician staff of this center had respective knowledge regarding their work. Conclusion: According to the score list, raised problems, the desired center needs to implement infrastructure of knowledge projects in the organization and new personnel and youth should be appropriately training programs had served during work time even understanding knowledge factors as an important investment to take organizational foundation.
A Similarity-based Selection Mechanism for Genetic Algorithms to Solve Assignment Problems

Rajabalipour Cheshne Gaz A. A., Rajabalipour Cheshne Gaz H.

Kerman Agricultural and Natural Resource of Kerman, Technical and Engineering Discipline

MIT University of Malaysia

In this paper, we illustrate a cellular structure mixed with genetic algorithms for solving assignment problems, which have more than one feasible or optimum solution. Considering similarity among individuals in a population, we use two dimensions cellular automata in order to place the individuals onto its cells to make the locality and neighborhood on a Hamming distance basis. This new structure in association with the use of the genetic algorithm on it, as a 2D cellular automaton Hamming GA, introduces locality for genetic selection and local knowledge for their selection process on cells of the 2D cellular automata. The cellular selection based on structure can ensure maintaining population diversity and fast convergence in genetic search and improve convergence performance during the genetic search.
New Online Facilities for "MIRU-VNTR" Mycobacterium Tuberculosis Data Analysis: Comparing with Offline Analyze

Asgharzadeh M., Samadi Kafil H.

Department of Medical Bacteriology, School Of Medicine, Tarbiat Modares University, Tehran, Iran

Infection by Mycobacterium tuberculosis (TB) is a leading cause of death in the world. Molecular studies are so important in controlling TB transmission in many regions. MIRU-VNTR is one of the best introduced techniques for the study of TB transmission. In this study, we used MIRU-VNTR, a PCR based method for the study of strains isolated from patients from the Northwest of Iran. Data analysis and comparisons for detecting possible relationships is critical in MIRU-VNTR. Previously, we used the MVSP software or other similar software for the analysis of possible similarities and clusters. Recently, a website for MIRU-VNTR plus was introduced in 2009 (http://www.miru-vntrplus.org/MIRU/index.faces; jsessionid=FA7FA334240ACA82795C3CA2C6C0F48C) for comparing and analyzing data. We found that this website provides more flexible and reliable analysis in comparison to other software, otherwise this site provides facilities for 12, 15 and 24 locus analyzing possibilities, but because of the insertion of a small amount of data, it can not be used as a suitable data bank for MIRU-VNTR analysis.
Absents of Automatic Software for Calculating HGDI Index for Calculating Score of Clustering in Mycobacterium Tuberculosis Epidemiology Studies

Asgharzadeh M., Samadi Kafil M.

Department of Medical Bacteriology, School Of Medicine, Tarbiat Modares University, Tehran, Iran

Mycobacterium tuberculosis is one of the important causes of death worldwide; therefore, epidemiological studies are so important for controlling the spread of infection. There is no endpoint technique for typing isolates of M. tuberculosis. Current methods for these studies are IS6110-RFLP, MIRU-VNTR, ETR-VNTR. Reliability and controlling of the clustering power of these methods are calculated by the HGDI index. Higher HGDI index represents a more acceptable clustering power of the method. We carried out IS6110-RFLP and MIRU-VNTR our isolates. Several offline and online softwares were presented for interpreting the data. Unfortunately, none of the online and offline softwares could calculate the HGDI index of clustering. Hence, we think it is necessary to add these indices to the tools of the MIRU-VNTR plus web based software for TB data analysis.
Design of GMP Software for Analysis of Genome and Amino Acid Sequences

Babaabasi B. 1

Biology Department of Tonekabon Azad University of Iran

One of the main concerns of Bioinformatics scientists in the world is the design of powerful software in order to analyse of genome and amino acid sequences. Different programming languages are used worldwide to make of Bioinformatics software, such as Perl, Payton, java I tried in variety experiences that have careful utilize applicable function Visual basic (VB) to make an applicable software and user friendly software. My objective was to make Bioinformatics software by VB mathematics function and string .And all my activities successfully ended to design software that is called GMP. This software involves four portions: "GMP ALY", "GMP PRo", "GMP GEN ", "GMP PRO TO GEN" that allow analysis of genome and protein data. For More information regarding application and algorithm of these portions, are presented in the discussion section In GMP, a series of variables for the section of "convert protein to gen" was used ,which and suggested a new and simple algorithm in this partGMP software is being programmed by the "VB  programming language". Working by GMP is very easy and GMP takes a little time in for counting. The source of GMP is available in the internet and other users can complete it. GMP Can be downloaded from the following address: http://www.ibp.ir/html/gmp.htm This software is useful and applicable for researching in biotechnology.
An Online-Random Algorithm for Motif Finding Problem in Biological Networks

Masoodi M., Razzazi M. R.

Department of Computer Eng. And IT Amirkabir University of Technology

Identifying transcription factor binding sites (Motifs) in the promoter regions of genes is an important problem in computational biology, which has been under intensive research for a decade. Accurate identifying of these motifs facilitates the understanding of the biological mechanisms involved in the transcriptional process of the gene. To predict these motifs, many approximate and heuristic algorithms have been developed. However, the prediction accuracy is not satisfactory. In fact, motif finding problem is proven to be NPComplete. We present a method that uses online and random algorithms to achieve higher prediction accuracy. The performance of the proposed method is demonstrated by applying it to some biological networks.
Designing and bBlasting Primers for Amplification and Separation of Particular Sections of the Genome

Heidari S., Hhojati Z.

Molecular Genetic Department Isfahan State University, Isfahan, Iran

The continuously growing field of Genetics has so far demonstrated that DNA is the base of inheritance. Following genome extraction from cells, it is necessary in many cases, that a particular section of the under study DNA is separated from the rest of the genome, but the amount of genome extracted is usually low, it is thus necessary for such low quantities of DNA to be amplified. The PCR technique is used for DNA amplification and separation, for which specific primers are used. In this article, we present designing and blasting primers. For this purpose, we use an oligo and CLC software for designing primers and NCBI-BLAST for blasting primers. At first, we obtained the sequence of exon 10 of the CFTR gene that is involved in the fertility from Ensemble. We created a cutting site for the NdeI restriction enzyme in the forward primer by the CLC software and acquired suitable Revers primer by the Oligo software after blasting. These primers, these were then introduced to a relevant company for synthesizes.
Molecular Study on Breeding Plans of Iranian Native Golpaygani Cattle by Stochastic Simulation

Zakizadeh S. ¹, Nejati-Javaremi A.², Rahimi G.³, Reissmann M. ⁴, Miraie-Ashtiani S. R. ², Peter Reinecke⁴, Moradi-Shahrbabak M. ²

¹Hasheminejad High Education Center, Mashad, Iran
²University of Tehran, Animal Science Department of Agricultural Faculty, Iran
³Animal Science Department of Sari University, Iran
⁴Humboldt University of Berlin, Germany

Although the number of genes are unknown, but there are some candidate genes for against of quantitative traits. It is known that some genes have a greater influence on genetic variations. In the first step of this study, gene frequencies of growth hormone; Pit1, β-lactoglobulin, α-S1- and κ-caseins were determined, then the possibility of the use of molecular gene effects on breeding programs of the Golpaygani cattle were investigated. DNA was amplified by the PCR procedure with regard to exon 5 of the growth hormone, a part of intron 5 and exon 6 of Pit1, exon 4 of κ-casein and Beta-lactoglobulin. PCR products were then digested by endonuclease enzymes; Ddel for the growth hormone, Hinfl for Pit1 and κ-casein, HaeIII for β-lactoglobulin. α-S1-casein was analyzed by PCR-SSCP. Gene and genotypic frequencies were calculated by the counting method. The single trait model was used to estimate additive, dominance and gene substitution effects in order to simulate milk yield in Visual Basic 6 software. The base population included 4,950 females and 50 sires with 1,700 + 422 kg for mean and standard deviation and 0.27 for heritability. Generations were overlapped and females stayed in the herd at most for 5 lactations and males for 4 years. Evaluation of females was based on individual records of the first lactation. Two methods were considered to evaluate males; based on phenotypic dam records or individual Molecular Genetics. Data were simulated for 20 generations and 40 iterations. Results showed a higher total breeding value in the molecular based method and a higher polygenic breeding value in the phenotypic method. There was significance difference in the 3rd generations and later. Although the molecular based program showed faster response in early the generations, but it decreased and reached a plateau. After 20 generations, desirable allele frequencies of β-lactoglobulin, α-S1-casein, growth hormone, κ-caseins and Pit1 increased by 2-fold to, 33% 16% 10% and 2%, respectively. Inbreeding values reached 4.5% in both method and was not significant.
Protein-Protein Interactions Network Topology
Concept: Aim at Suitable Agriculture of Tomorrow

Alisoltani A., Ebrahimi A., Moradi A., Alizadeh H., Naghavi M. R.

Department of Horticultural Sciences, Faculty of Agriculture, University of Tabriz, 5166616471, Tabriz, Iran

The advent of proteomics has made it possible to identify a broad spectrum of proteins in living systems. This capability is especially useful for crops as it may give clues not only about nutritional value but also about yield and how these factors are affected by adverse conditions. The study of protein-protein interactions (PPI) is an increasingly important part of post-genomics strategies to understand protein function. Study the network of these interactions will provide valuable insight into the inner working of cells. The topology of a network refers to the relative connectivity of its nodes. Different topologies affect specific network properties. Here we define the most basic network measures that allow us to analyzing topological characterization of PPI networks, such as degree, degree distribution, power law distribution, network models and clustering coefficient.
Suggestion of a More Efficient Model of Artificial Protein According to Probable TIMP-4 Three Dimensional Structure

Sharifirad A., Ahadi A. M., Ayat H., Ilkhanizadeh Sh., Lohrasbi A.

Shahrekord university, Shahrekord, Iran

A programmed degradation of the extracellular matrix (ECM) is a necessary step in various physiological processes, such as embryogenesis, morphogenesis and wound healing. Any failure in this process results in many anomalies such as arthritis, autoimmune diseases and especially, metastasis in malignancies. The enzymes belong to the matrix metalloproteinase (MMP) family and have an important role in this pathway, which can destroy proteins of the extracellular matrix and basal membranes. Proteolytic activities of these enzymes are regulated by gene transcription, zymogen activation and specific inhibition by their physiological tissue inhibitors (TIMPs). The human TIMP-4 is a non-glycosylated, 195 amino acids polypeptide. Three-dimensional (3D) structure of TIMP-4 has not yet been determined, but a probable 3D structure is predictable due to the high sequence identity with the other TIMPs members. Crystallographic structural analysis of the TIMPs-MMPs has shown that the long edge of the wedged-shaped inhibitor occupies the entire length of the MMP active-site cleft. According to their conserved domain in the catalytic sites, using molecular dynamic softwares, such as HEX, CLC, etc., we predicted the TIMP-4 probable structure and designed a new artificial protein, with more affinity to catalytic sites of MMP-1, -2, -3, -7, -9, -26 associated with substrates of TIMP-4.
Study of Evolution Phylogeny of Sturgeon Fish Using Beluga (Huso huso) as Model

Gharaei A. 1, Mahboudi F. 2

1 Department of Fisheries, Faculty of Natural Resources, University of Zabol, P.O.Box: 86515-538 Zabol, Iran.

2 Department of Biotechnology, Institute of Pasteur, Tehran, Iran.

The Beluga (Huso huso) is the largest species of Acipenseriformes that belongs to the lower Jurassic. To elucidate identification of cGnRH-II, the phylogeny and evolution of the Beluga, we sequenced the cDNA of the cGnRH-II gene of the brain using 5'/3' RACE-PCR. The nucleotide sequence of the Beluga cGnRH-II was found to be 258 base pairs (bp) long; encoding peptides with 86 amino acids. The cGnRH-II precursor cDNA encodes a signal peptide (SP) composed of 24 amino acids, and a GnRH-associated peptide (GAP) composed of 49 amino acids, which is connected to GnRH by a Gly-Lys-Arg sequence. This is the first report of the cGnRH-II precursor cDNA from Sturgeons. Molecular phylogeny analysis combined with sequence comparison indicated that the Beluga is more similar to reptiles and amphibians than to fish with respect to the cGnRH-II precursor sequence. The phylogenetic tree of cGnRH-II precursors showed that among all vertebrates four subtypes of cGnRH-II exist: 1- birds, 2- mammals, 3- marsupial mammals, reptiles, amphibians and sturgeons, 4- bony fish. These results suggest that the sequence divergence seen in the Beluga may have appeared independently of that in the bony fish lineage.
Microsatellite Profiling Technique and Assignment Programs Serve to Detect Parentage in Caspian Salmon (*Salmo Trutta Caspius*)

Sourinejad I., Kalbassi M. R., Vera M., Martínez P.

1Tarbiat Modares University, Tehran, Iran.

2Departmento deGenetica, Santiago Universidad de Santiago de Compostela (USC) Faculted de Veteriravia Spain

Conservation of genetic diversity in the endangered Caspian salmon is necessary for appropriate restocking of natural populations when artificial breeding is conducted. One of the difficulties in maintaining genetic diversity in hatcheries is the lack of pedigree information. The advent of DNA microsatellite markers has simplified molecular based pedigree retention and genetic diversity quantification when mixed milt fertilization is performed in hatcheries to obtain progeny. Specialized assignment programs and other statistical analyses are used after microsatellite profiling to detect parentage of communally reared alevins from different breeders. In the present research, DNA of four wild male and four wild female Caspian salmon breeders were extracted and genotyped using 9 microsatellite loci with the ABI PRISM(R) 3730 automatic sequencer and GeneMapper 4.0 software. The 8 breeders were then crossed in 2 different combinations through mixed milt fertilization and the resultant progeny were also genotyped. Three loci (Str58, Str73 and Str591) were finally selected for parentage assignment of progenies with high combined probabilities of exclusion based on the primary screening of the 9 microsatellites. Exclusion-based parentage programs including CERVUS 3.0 and FAP v.3.5 unambiguously assigned 98.4% and 98.8 of progenies to a single pair of parents in 750 alevins analyzed in treatments 1 and 2. Effective population sizes of the breeders were 4.69 and 4.25 in comparison to the census size of 6 and 8 in treatments 1 and 2, respectively. Contribution of male and female breeders to the progeny were generally significantly different between females regardless of the mating pair (P<0.05), between males regardless of mating pair (P<0.05) and also regarding the mating pair in treatments (P<0.05). These results illustrate that the current hatchery breeding protocol causes some loss of genetic diversity in the produced progeny. Genetic contribution of breeders to the progeny should be more balanced by applying appropriate fertilization designs to the hatchery.
The Assessment of Biosafety in Cry Proteins by Using Allergenic Databases

Rashidi S. 1, Rezaei M. 2, Ebrahimi A. 3

Department of Horticultural Sciences, Faculty of Agriculture, University of Tabriz, 5166616471, Tabriz, Iran

The development of Bt plants has been investigated worldwide because of their high specificity and long-term resistance. Hence, a study of their biosafety is required with respect to their high importance. One of the methods for studying allergenic proteins is aligning them with allergenic proteins in databases, based on rules of FAO-WHO. In this study, 81 cry proteins were assessed by using three databases of allergen proteins, Allergenonline, Allermatch and SDAP. Based on the three methods of full alignment, 80 amino acids window and 6 or 8 continues identical amino acid. In the 80 amino acid method, was found any allergenic proteins in three databases. But in two other methods was found very allergenic proteins.
Bioinformatics Analysis for Study of PPAR Gamma Gene Expression during Neurogenesis

Shabani Kh.¹ Ghaedi K.¹,², Nasr-Esfahani M. H.², Tanhaei S.², Karbalaei Kh.², Vallian S.¹

¹Biology Department School of Science, University of Isfahan, Iran
²Department of Cell and Molecular Biology, Royan Institute for animal Biotechnology, ACECR, Isfahan, Iran

The PPARs (peroxisome proliferator-activated receptors) are members of the nuclear hormone receptor superfamily. Three major isoforms of PPARs (α, β and γ) have been identified that each with a possibility of different ligands, target genes and biological roles. The expression of PPARα, β and γ varies widely from tissue-to-tissue. PPAR γ has been demonstrated to play an important role in the regulation of cell differentiation. PPAR γ has 2 isoforms termed: γ1, γ2, which are arisen by differential transcription start sites and alternative splicing.

We have already demonstrated that the expression of PPAR γ1 is increased upon retinoic acid treatment of embryonic stem cells during the neurogenesis. In the present research, we have attempted to find a further look into the molecular mechanism of PPAR γ1 gene expression during of the differentiation of stem cells to neurons. Thus, coding sequence of PPAR γ1 (1857bp) was obtained from NCBI web site and using several bioinformatics soft wares especially for primer design, specific primers were designated. The aforementioned softwares were Beacon Designer and Perl primer. Beacon Designer is an ideal software for Real Time PCR primer design. This software classifies designed primers at 3 levels: best, good, poor. We chose primers with the optimized Tm. Data were evaluated by Perlprimer too. The designated primers were:

Sense: 5´-TGAGACCAACAGCCTGAC-3´  Tm: 52.9   GC%: 55
Antisense: 5´-GTTCACCGCTTCTTTCAAATC-3´  Tm: 53.4   GC%: 42.9

Using this primers we are going to chase the expression level of PPAR γ1, during the different steps of neurogenesis.
An Engineered Peptide for Sticky Hands of HIV Virus

Borna H.

Imam Hossein University, Tehran, Iran.

AIDS is caused by the HIV virus. This virus belongs to the RNA virus family that attacks the human immune system T-helper cells. HIV’s sticky hand for fusion with cells is the protein complex of Gp41-Gp120. Gp41 functions both as dock and zip for membrane fusions. There are lots of drugs inhibiting the entrance of the virus but virip, a little double chain, consisting of 32 amino acids and a natural blood protein, which inhibits fusion between the viral envelope and the cell membrane. In this study, the interactions of this peptide with the HIV fusion peptide is explored with the the HEX 4.5 docking software and the Patch Dock web server for design of appropriate peptidomimetics. Then analyzing the possible natural conformations, similarity search was accomplished by the Exposy server with regard to its secondary and tertiary structure. After recognition of impacting residues, good pharmacophores were selected in connection with their positions and interacting amino acids. In consideration the total construction of virip and the fusion peptide, another similarity search was accomplished by secondary alignment of the fusion peptide. This will help proper design of the complementary peptide. Consequently, some sample peptides were synthesized by Hyperchem 6, 2006 and folded by Gromacs 3.3.0, under both vacuum and soluble conditions for the purpose of better comparison. Finally, a raw sequence was achieved interacting with the fusion peptide.
A Similarity-based Selection Mechanism for Genetic Algorithms to Solve Assignment Problems

Rajabalipour Cheshne Gaz A. A., Rajabalipour Cheshne Gaz H.

Kerman Agricultural and Natural Resource of Kerman, Technical and Engineering Discipline

MIT University of Malaysia

In this paper, we illustrate a cellular structure mixed with Genetic Algorithms for solving assignment problems which have more than one feasible or optimum solution. Considering similarity among individuals in population, we use two dimensions Cellular Automata in order to place the individuals onto its cells to make the locality and neighborhood on Hamming distance basis. This new structure and using Genetic Algorithm on it, as 2D Cellular Automata Hamming GA, introduces locality for genetic selection and local knowledge for their selection process on cells of 2D Cellular Automata. The cellular selection based on the structure can ensure maintaining population diversity and fast convergence in the genetic search and improve the convergence performance during the genetic search. Keywords- Genetic algorithms; Assignment problems; Cellular automata; Optimization; NP-hard Multi solutions problems.
Bioinformatics Studies for Determination of Transcription Factor Binding Sites in Ppromoters of Mouse PPARγ Isoforms

Lachinani L. 1, Ghaedi K. 2,3, Nasr-Esfahani M. H. 2, Tanhaei S. 2, Yaghmaei P. 1

1Department of Biology, Science and Research Branch, Islamic Azad University, Tehran, Iran.

2Department of Cell and Molecular Biology, Royan institute for Animal Biotechnology, ACECR, Isfahan, Iran.

3Department of biology, School of Sciences, University of Isfahan, Isfahan, Iran.

Different transcription factors (TFs) implement their effects by binding to their specific target sites in the genome, and exist mainly at the site of regulatory structures like promoters. Thus, determination of TF binding sites would be very helpful for identification of promoter regions and comparison of promoter regions of homologous genes, determination of probable effective factors in regulation of the expression regulation of a specific promoter region, Peroxisome proliferator-activated receptor γ (PPARγ) is a nuclear hormone receptor and has two isoforms, PPARγ 1 and PPARγ 2, in human and mouse, which differ in their distribution in tissues and specificity for different ligands. These isoforms are transcribed from different promoters which contain several TF binding sites. Thus, identification of these TF binding sites at promoters, of the PPARγ isoforms is required to unravel the molecular mechanism of the regulation of expression of genes expression regulation in different tissues. We have used three different Bioinformatics web sites, Genomatix, TESS and GeneBuilder to find TF binding sites. On the basis of Genomatix predictions, we analyzed two putative regions, 870 bp and 689 bp, respectively as PPARγ 1 and PPARγ 2 promoter regions. These selective promoter regions contain consensus binding sequences of transcription factors AP1, AP2, C/EBP, and HNFs. They Digestive system and liver tissues are rich in these TFs, which are consistent with the expression of PPARγ transcripts in these organs. Transcriptional regulation by PPARs requires heterodimerization with the retinoid X receptor (RXR) and it is notable that these promoter regions contain consensus binding sequences of RXR, the PPAR/RXR and PPAR homodimers. Taken together, we used these high potential regions as promoters of mouse PPARγ isoforms.
Pathogenicity Analysis of G9055A Mutation in Mitochondrial ATPase 6 Gene in Iranian Patients with Familial Brugada Syndrome

Khatami M., Heidari M. M.

Department of Biology, Science school, University of Yazd, Yazd, Iran

Brugada syndrome is an autosomal disorder characterized by ST elevation in the right precordial leads, which can result in syncope or sudden death. Brugada syndrome typically predisposes men between 30-50 years to syncope, and mutations in the gene encoding the cardiac Na channel α subunit have been identified in 20-30% of the patients. Recently, interesting genetic data became available regarding the importance of mtDNA mutations, which might explain the influence of respiratory chain activities in these patients. In this study, we found G9055A mutation of the mitochondrial ATPase 6 gene in 4 Iranian patients. This mutation converts a conserved alanine, a hydrophobic AA into threonine, a neutral AA outside the transmembrane domain of F0 F1- ATP synthase. Interspecies conservation was assessed using BLAST and the Polymorphism phenotyping (PolyPhen) database. Secondary structure and physiochemical analysis were performed by (protein structure prediction and annotation Protean software, (sorting intolerant from tolerant (SIFT) program and the PolyPhen database. Our results showed that this mutation alters the hydropathy index from -1.16 to -1.03 and surface probability from 0.369 to 0.528. Also this change decreases the charge of the protein at pH=7 from 6.96 to 6.80 and probably alters the structure of the proton channel contained in complex V. Because it was not found in the controls (25 samples) and results predicted that it could possibly be damaging, it was considered potentially pathologic. However, further investigations are necessary to clarify this.
Alternative splicing AS is a process that causes production of multiple mRNA isoforms from single premRNA in different cells and tissues at the higher eukaryotes. Although there are a lot of information about AS in animal genomes, but little is known about the extent of this phenomenon in higher plants. In recent years some scientists have focused on the roles of AS in plant genomes and have found that this phenomenon has far more significant roles in biological functions of plants than was previously thought. Although AS occurs less frequently in plants than in animal systems but up to 35% of genes in Arabidopsis and rice genes are estimated to produce more than one transcript. A range of important plant functions including growth and development, signal transduction, disease resistance, biotic and abiotic stress responses, grain quality, flowering time and the circadian clock are regulated or effected by AS. Hence, AS should be consider as an important factor in plant breeding and trait selection programs. The increasing availability of plant genomic data will provide a great source for computational analyses of AS in plants and allow assessment of their importance in plant functions. In this paper, we reviewed biological functions of AS in plants and discussed the importance of doing more studies on AS in plants and also we suggested what directions to take regarding about future works.
Neutrality Test of Two PAH-BglII and PAH-EcoRI Polymorphisms in the Isfahan Population

Fazeli Z., Vallian S.

Division of Genetics, Department of Biology, Faculty of Science, University of Isfahan, Isfahan, Iran.

Two restriction fragment length polymorphic (RFLP) markers including BglII and EcoRI, were identified at intron 1 and intron 5 of the PAH gene. In order to test whether these polymorphisms are behaving as neutral alleles or are being subjected to selective pressure in the Isfahani population, 110 individuals were genotyped by PCR-RFLP. The Arlequin input file was prepared by use of the phase-known haplotype data and neutrality tests (Tajima D test and Fu's Fs test) were carried out using the Arlequin program. Forty two individuals were found heterozygous at both of these RFLPs with unknown haplotype phase. The BglII-EcoRI haplotype phase was only known in 68 individuals who were used for preparation of the input file. Tajima's D and Fs values in the t Isfahan population were 1.7 and 1.02, respectively. D>0 and positive value of Fs indicated that these polymorphisms are under selection pressure or a recent population bottleneck in the Isfahan population. Although these polymorphisms were in the non-coding region of this gene, but these were not neutral alleles and positive value of these tests provided evidence for balancing selection of these polymorphisms in the Isfahani population. The results of this study could improve our understanding of evolutionary history and structure of the Isfahanipopulation.
Introduction of Candidate Ssubstrates for Drosophila Gamma-Carboxylase

Vatandoost J., Zomorodipour A. R.

National Institute Of Genetics Engineering and Biotechnology Of Iran, Tehran, Iran.

Gamma-carboxylation is essential for biological activity of certain proteins, including a number of coagulation factors. This process occurs in the endoplasmic reticulum to modify certain glutamate residues by the action of gamma-carboxylase. Carboxylation of glutamate requires the γ-carboxylation recognition site (γ-CRS) that typically resides within a 12-28 residues (propeptide) adjacent to the signal peptide. This enzyme has been characterized in two invertebrates: Drosophila melanogaster and cone snails. Cone snail is the only invertebrate in which γ-carboxylase and its substrate have been identified. No γ-carboxylase substrate of Drosophila origin has been identified so far. While the propeptide regions of mammalian proteins bear considerable sequence, homology amongst themselves, the peptides obtained from cone snails exhibit γ-CRS and Gla domain sequences which are different from the mammalian. However, two of the mammalian γ-carboxylase substrates are recognized by the Drosophila enzyme. Screening of the Drosophila genome data-base, using a consensus sequence, representing a number of Gla-domain of human origin, was not able to identify any Gla protein in Drosophila. In this study, to identify a candidate substrate(s) for the Drosophila carboxylase, we used a sequence motif, representing a number of propeptides of mammalian and cone snail origin to screen the Drosophila genomic data-base. However, Gla domainssimilar to those found in mammalian Gla proteins.were However, screening of the Drosophila database using the propeptide sequences of individual Gla proteins in cone snail resulted in detection of some glutamate-rich proteins for which further Bioinformatics analysis of the candidate protein structure is required in order to be able to confirm their potentials as substrates for the Drosophila γ-carboxylase.
Annotation of Herceptin/Her2 Interactions Using Protein Databases and Molecular Structure

Shojaei S., Gardaneh M.

National Institute of Genetic Engineering and molecular biotechnology department, Iran, Tehran, Iran

The interface between antibody and antigen is often depicted as a lock and key, suggesting that an antibody surface can accommodate only one antigen. Trastuzumab (Herceptin) is a monoclonal antibody that interferes with the HER2/neu receptor. The HER receptors are proteins that are embedded in the cell membrane and communicate molecular signals from outside the cell to inside the cell, and turn genes on and off. HER2 is a ligand orphan receptor. It plays a critical role in the cellular effects of the ErbB receptor network by amplification the signal through heterodimerization with another ligand-activated ErbB receptor. Heterodimerization of HER2 and HER3 delivers the most potent and long lasting proliferative signal among the possible combinatorial pairs of ErbB members, and HER3/HER2 is the most representative heterodimer found in cancers The HER proteins regulate cell growth, survival, adhesion, migration, and differentiation—functions that are amplified or weakened in cancer cells. Signaling compounds called mitogens arrive at the cell membrane, and bind to the outside part of HER2. Evidence has accumulated that receptor dimerisation is essential for receptor activation. The HER2/HER3 dimer is the most active. In mammals, the HER2 receptor plays an important role in the development of cardiac and neural tissue. HER2 is activated, and sends a signal to the inside of the cell. The signal passes through different biochemical pathways. This includes the PI3K/Akt pathway and the MAPK pathway. However, in cancer, HER2 sends signals without being stimulated by mitogens first (that is, they are constitutively transmitted). These signals promote invasion, survival and growth of blood vessels (angiogenesis) of cells. Herceptin is the first FDA approved therapeutic antibody against HER2, has been actively used for the clinic treatment of women with HER2-overexpressing breast cancer. When it binds to defective HER2 proteins, the HER2 protein no longer causes cells in the breast to reproduce uncontrollably. This increases the survival of people with cancer. The crystal structure of the extracellular region of HER2 has been determined alone and in complex with the Fab fragment of Herceptin by X-ray method. The structure provides insight into HER2 activation in the absence of direct ligand binding. Herceptin is shown to bind to the juxtamembrane region of HER2, suggesting that antibodies raised specifically against
this region may have useful therapeutic properties in patient suffering breast cancer. The annotation of protein is essential in biology and for any systematic approach to the modeling of biological systems. Currently, functional annotation is essentially based on the expansion of the relatively small number of experimentally determined functions to large collections of proteins. We used protein sequence database that included Swiss-Prot, UniProtKB/Swiss-Prot and NCBI Protein Database. Then we used secondary Database for example ProSite, PRINT, Pfam and Interpro to complete information about our proteins (Herceptin and Her2). To analysis of 3D-structure of them we entered to PDB Database as a Molecular Structural Database and used 1n8z code for searching their x-ray structure and found The crystal structure of the extracellular region of HER2 that has been determined alone and in complex with the Fab fragment of Herceptin. To made a new annotation about these proteins we used other Molecular Structural Database that included RCSB, PDBE, SRS, MMDB, JenaLib, ProteoPedia, CATH, SCOP, FSSP, HSSP, PDBSWS, PQS, PROCOGNATE, ProSAT and Whatcheck. We classified all of these data and make a new annotation for Herceptin and Her2 proteins separately and Herceptin and Her2 interactions.
Bioinformatics Studies to Identify a Putative Region for Promoter of Mouse PEP

Seifi T., Ghaedi K., Nasr-Esfahani M. H., Tanhaei S., Hojati Z.

Department of Biology, School of Sciences, University of Isfahan, Isfahan, Iran

PeP is one of the Proxisomal matrix proteins that it was cloned by Ferrer Martinez et al. in 2002. However, there is no specific role has been identified yet for it. Our previous data were shown that this protein expression is increased during neurogenesis processes. Studies which was carried out revealed significant increasing in Proxisomal Protein (PeP) expression after Retinoic Acid (RA) treatment of stem cells during differentiation to neural cells. RA affects on genes expression via RXR (retinoeid X receptor) which is a nuclear receptor (NR). NRs are involved in diverse physiological functions such as metabolism, development and reproduction. Increasing evidence shows that certain NRs function in regulating stemness or differentiation of embryonic stem (ES) cells and tissue-specific adult stem cells. It is supposed that PeP protein has role in differentiation process to neuron. Thus we have used a bioinformatics search to identify PeP promoter and its response elements related to different signals.

At the first step, we used promoter prediction softwares such as Genomatix and Proscan to predict region having promoter potentiality. These softwares are designated based on promoter algorithms predict promoter. Results for PeP gene from blast result sequence proposed different elements in this predicted promoter. Genomatix predicts a putative promoter at -551/+101 position, RxR/VDR motif at -11/+14 position and TFII B motifs at +54/+66 position.

Taken together using this data we are going to clone this hypothetical region containing promoter activity for further analyses.
Survey of Data Mining Algorithms in the Protein-Protein Interactions Analysis

Ravaee H.

College of Engineering, Tehran University, Tehran, Iran

Interactions between proteins in biology in the developmental processes within the cell play a fundamental role. With recent advances in diagnostic techniques and tools to detect interactions between the proteins and also provide a large volume of information from the living organism, the researchers are thus provided with an opportunity is given. In this article, penetrable network interactions between proteins from two aspects will be studied. First, the algorithm will look into some seeking repeated subgraphs. These types of repeated subgraphs are called motifs. In other categories of search algorithms, clustering collection network vertices of the subsets are separated so that each subset most edge there is in each clusters and the number of edges between clusters is at a minimum. These algorithms search in the subgraphs with most density. This article, these two types of algorithms is expressed.
Bioinformatics Approach for Designation of the Most Appropriated shRNA to Knock Down Peroxisomal Protein Gene

Hahsemi M., Ghaedi K., Emadi M., Nasr-Esfahani M. H.

Department of Biology, School of Sciences, University of Isfahan Isfahan, Iran

PeP is one of the pProxisomal matrix proteins, which comprises 209 amino acid residues. There is a three peptides stretch, SKL, as a protein targeted signal type 1 at the c-terminus of PeP. PeP expression is high during neurogenesis and myogenesis. Our previous studies revealed a significant increase in PeP expression when mouse embryonic stem cells were differentiated into neural cells upon retinoic acid treatment. In order to a further look into the PeP gene function during differentiation of mouse embryonic stem cells into neurons, we designated appropriated shRNAs specific for knocking down PeP expression. In order, to design shRNAs, we utilized several siRNA designing websites such as siDirect, invitrogen, Ambion, Genescript, and Darmacon. Among the optimized designated siRNAs which were 19nt in length, more appropriate ones were checked for secondary structure formation using Geenbee and Mfold softwares. At the next step, the specificity of designated siRNAs against PeP was assessed using BLAST search. Finally three siRNAs were selected and modified to produce ShRNA by inserting a sequence loop between the sense and antisense sequences.
The Comparison Results of Two Secondary Protein Structure Prediction Methods on Mitochondrial A458T Missense Mutation in ND5 Protein

Heidari M. M., Khatami M.

Department of Biology, Science School, Yazd University of Yazd, Yazd, Iran

Friedreich’s ataxia (FRDA) is the most common autosomal recessively inherited ataxia. We found m.13708G>A mutation in the mitochondrial ND5 gene in 2 FRDA patients. This Mutation causes a change of Ala to Thr (A458T Missense Mutation). In this study, the effects of missense mutation upon protein structure were assayed by means of two secondary protein structure prediction methods (Chou-Fasman and Garnier-Osguthorpe-Robson). The secondary structure prediction results by Chou-Fasman method showed that the A458T mutation caused an increase in the extended strand from 74.3% to 75% in the transmembrane ND5 protein. The results of the secondary structure prediction by Garnier-Osguthorpe-Robson method revealed that the extended strand of ND5 protein by means of A458T mutation changed from 38.3% to 38.7% and the alpha-helix decreased from 29.1% to 28.8%. Although these two methods are incomplete methods for secondary protein structure prediction, but they are useful methods for comparison of the effects of mutation on protein structure.
Designing a Set of Universal Primers for Detection of β-Lactamase Super Family

Sabbaghi A., Fallah Mehrabadi J., Molla Agha Mirzaei H., Soltan Dallal M. M.,

Islamic Azad University, Branch of Zanjan, Zanjan, Iran

MARS Bioinformatics Institute, Tehran, Iran

Islamic Azad University, Branch of Zanjan, Zanjan, Iran

Tehran University of Medical Sciences, Tehran, Iran

Resistance to β-lactam antibiotics along with clinical isolates frequently result in the production of β-lactamase enzymes. Nowadays, the appearance of novel β-Lactamase enzymes with novel substrates such as extended spectrum β-lactamases (ESBLs) and AmpC β-lactamase among clinical isolate especially E. coli led to failure in treatment and diagnosis via recommended phenotypic tests by CLSI. Alternatively, β-lactamase super family like SHV, CITM and FOX genes have several subfamilies, so the purpose of this study was designing universal primers for each cluster to complete detection of these genes through PCR. Method: Submitted sequences related to SHV, CITM and FOX genes of E. coli were 50, 24 and 5 (respectively) in GenBank. The sequences of each cluster were aligned by MEGA 4 multiple-alignment program until analogous locals were identified. These locals were used for designing of primers by the Gene runner software. Subsequently, designed primers were tested on Submitted sequences by BLAST. Then, three sets of universal primers were evaluated using PCR in wet lab. Results: PCR was performed on 128 E. coli isolates and results showed 7 (5.5%) and 13 (10.2%) bla SHV and bla CITM, respectively. the Fox gene was not detected in any of the samples. Conclusion: the primers used, potentially covered all of the resistant isolates. This case is very significant because in recent years, the prevalence of infections caused by ESBLs and AmpC producers have increased and, which has led to a problem for diagnosis the disk diffusion method. Therefore, designing universal primers could be very practical for diagnostic laboratories to detect of subfamilies β-lactamas enzymes.
Use of Universal Primers to Targeting of 16S rRNA Gene for Identification of Gram-Positive Bacteria in Environmental Samples

Sabbaghi A., Najafi S.

Islamic Azad University, Branch of Zanjan, Zanjan, Iran

Broad-range PCR assay has been detected as one of the most important available methods for identification of the any kind of bacterial DNA in environmental samples through targeting conserved sequences. In this method, using universal primers and targeting 16s rDNA and followed by sequencing of the amplicons and comparing with submitted Sequences in data banks, could be valuable to detect all of the organisms. On the other hand, these conserved sequences are as high copy number available in the most organisms genome. Therefore, the aim of this investigation was designing a set of universal primer for direct detection of the gram-positive bacteria. Method: Submitted sequences related to 16S rRNA genes of the gram-positive bacteria were 300 in GenBank and subsequently these sequences through program of MEGA 4 multiple-alignment were aligned and detected of the similar regions. According to these regions, one set universal primer was designed and specialization evaluated for molecular assessing with BLAST program. Results: Designed primers were theoretically tested on all of the Submitted sequences in GenBank and results were showed that these primers could be using in other studies. Conclusion: Nowadays, detection of the organisms without using culture methods is very usual because a lot of important organisms are finding in unusual environment. Therefore, using culture methods could be result to missing of some unknown microorganisms. So, using of these primers was proposed for screening of bacteria by using molecular methods in environmental samples.
Evaluation a Set of Universal Primers for Molecular Detection of TEM and AmpC (Dha and Mox)-Type - Lactamase in Clinical Isolates of Escherichia Coli

Molla Agha Mirzaei H. 1, Sabbaghi A. 1, Fallah Mehrabadi J. 2, Soltan Dallal M. M. 3

Islamic Azad University, Branch of Zanjan, Zanjan, Iran
MARS Bioinformatics Institute, Tehran, Iran
Tehran University of Medical Sciences, Tehran, Iran

The fragments of genomic via mutation and transmit speared with high rate among of bacteria that coding a kind of enzymes, make them resistant to betalactam antibiotics. These enzymes called ESBLs. TEM is one of ESBLs subfamily which has sequence variation. So it is impossible to identify it with phenotypic methods. Designing a universal primer that detects all of the sequence of afailure of the TEM genes was the purpose of this study. Presence of AmpC that is co-produced with TEM causes afailure of the diagnostic method. We also designed 2 sets of primers of AmpC (Dha and MOX), so led to complete recognize during PCR. Materials and methods: The number of submitted sequences relating to TEM, Dha and MOX of E.coli were 120, 22 and 6 in GenBank. All of the sequences were aligned with the MEGA 4 program that was showed more than 90% homology. Then, conserved sequences with high homology were selected for designing universal primers by the Generunner software. Subsequently, the designed primers were tested on Submitted sequences by BLAST. So, three sets of universal primers were validated by PCR amplification and DNA sequencing. Result: Among 128 clinical samples that were used in the PCR test, 72 (62.2 %) and 3 (23.1%) indicated as TEM and Dha genes, respectively. The MOX gene was not detected in any samples. Conclusion: Designed primers covered most of the submitted sequences. Nowadays, the prevalence of beta-lactamase producing micro organisms has increased, which has it led to a problem for diagnosis. Therefore, using molecular methods in support of phenotypic tests are very essential for the detection of the beta-lactamases family.
Evaluation of SaCOL2291 and SaCOL2581 Genes Polymorphisms, New Candidate Genes for Preparation of S.aureus Vaccine

Rohani Nejad H., Pourmand M. R., Fallah Mehrabadi J., Zolfaghari M. R.

Islamic Azad University, Qom branch

Introduction: Staphylococcus aureus is one of the most important agents of nosocomial infection. The prevalence of resistant strain is high and it led to some complication for treatment of related staphylococcal infection. Vaccination strategy is suitable way against infectious diseases. Therefore, preparation an effective vaccine against S.aureus is a remarkable field. saCOL 2291 and saCOL 2581 genes are as a candidate for preparation of new vaccine against S.aureus and related sequence submitted in GenBank but just one sequence for each.

Materials and methods: In this research, polymorphism of these genes evaluate for detection of SNPs. So, genomic DNA was extracted from 30 S.aureus isolates and saCOL 2291 and saCOL 2581 genes amplified via PCR. Then, amplicons were sequenced and aligned with puls4 software and comparison with S.aureus COL strain genes.

Results: The results showed polymorphisms in DNA sequence, which could changes in some codons and it doesn't affect amino acid sequence.

Discussion: Despite of identification of polymorphism in saCOL 2291 and saCOL 2581 genes, most of these were transition mutation and did not influence on proteins ScaC and ScaD sequence. Hence, these changes have no imperative effect on their immunogenic function and it seems both of them could be used as new candidate immunogen proteins.
Informative SNP Selection by Particle Swarm Optimization

Ghasem Mahdevar, Javad Zahiri, Mehdi Sadeghi, Hayedeh Ahrabian, Abbas Nowzari-Dalini

Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran

National Institute of Genetic Engineering and Biotechnology, Tehran, Iran

Center of Excellence in Biomathematics, School of Mathematics, Statistics, and Computer Science, University of Tehran, Tehran, Iran

Currently of the most interesting areas in the field of genetics studies is to find variations in the genomes of population that cause several diseases; finding one or a set of differences in the constituent nucleotides of the genome that are significantly correlated with specific diseases. The abundance and high discriminatory potential of these single nucleotide polymorphisms (SNPs) have made them superior in comparison with other markers. Unfortunately, such as abundance that has made SNPs so valuable, has also led to the complexity inherent in these studies. Selection of informative SNPs in a way, which minimizes information loss, may obviously be a solution. The required time for finding the smallest statistically acceptable set, grows exponentially with the number of SNPs. Therefore providing an approximate solution in a feasible time is appropriate. In this paper a novel method based on particle swarm optimization for informative SNP selection approach has been proposed. This algorithm is implemented and compared with existing exact and approximate methods. Our running time is less than other methods and the achieved accuracy tends to be as that of the exact methods.
Analysis of Synonymous Codon Usage Bias and Base and Amino Acid Composition in 13 Species of Flaviviridae

Mohabatkar H., Moosawi F., Mohsenzadeh S.

Department of Biology, College of sciences, University of Shiraz, Shiraz, Iran

Flaviviridae are viruses that cause several diseases including dengue fever, Japanese encephalitis, Murray valley encephalitis, tick-borne encephalitis, West Nile encephalitis, yellow fever and hepatitis C virus infection. Here, we analyzed the codon usage of 13 species of this family by using the gene infinity package. Base and amino acid composition analyses also were also performed by the CAIcal and PseAA web servers, respectively. The results showed that the highest amount of A, G and C bases were in the RNA genome of the dengue virus 2, tick-borne encephalitis and hepatitis C viruses, respectively. The U bases were used in approximately equal numbers in this family, although the highest U nucleotide count was 23.77% in the Wesselsbron virus. The lowest amounts of C, G, U and A bases were seen in the bovine viral diarrhea virus, dengue virus 2, tick-borne encephalitis and hepatitis C viruses respectively. In our study, it was found that the complete genome of the classical swine fever virus had a lower average GC content and the genome of the tick-borne encephalitis hepatitis C and Powassan viruses had a higher average GC content than other species. We also classified the amino acids as rare (phenylalanine, cysteine, histidine, methionine, asparagine, glutamine, tryptophan and tyrosine), frequent (alanine, glutamic acid, glycine, leucine, valine and threonine), and intermediary (all others). Leucine was the most frequent and cysteine was the least frequent amino acid among the viruses in Flaviviridae. The highest and the least number of preferred codons existed in the Wesselsbron and West Nile viruses respectively.
Bioinformatic Investigation of Short Tandem Repeat Loci D6S2879 and D6S2880 Associated with HLA-DRB1

Sheikholeslam M., Vallian S.

Faculty of Science, Department of Biology, University of Isfahan, Isfahan, Iran.

The human leukocyte antigen (HLA) class II region exhibits a remarkable degree of genetic polymorphism and subdivides into four subregions; DP, DO, DQ, and DR. In the human genome, there are five different DR haplotype groups, all of which harbor one of the highly polymorphic alleles of the HLA-DRB1 locus. The association between HLA-DRB1 and large number of diseases has been confirmed. In the current study, two STR markers including D6S2879 and D6S2880, which were introduced in the MHC database (dbMHC), were subjected to Bioinformatics analysis to investigate whether these markers could be appropriate for HLA-DRB1 genotyping. The presence of these markers in three DR haplotype groups including DR51, DR52 and DR53 was investigated. Each DR sequence retrieved from the nucleotide database at NCBI was used to locate D6S2879 and D6S2880 markers. Additionally, the existence of these markers in six different MHC haplotype sequences containing HLA-DRB1 as well as the reference assembly sequence were investigated by the CLC main workbench software. DR52 sequence scan showed that D6S2879 and D6S2880 markers were located approximately 1.4 kb and 23 kb downstream of the HLA-DRB1, respectively. However, D6S2879 was also present downstream of the HLA-DRB3.

Furthermore, it was revealed that the primers introduced for the D6S2879 marker at the uniSTS and MHC databases could amplify this marker at both locations, but with different sizes of 283 and 411 base pairs (bp). The data indicated that D6S2879 and D6S2880 existed in DR51 and DR52 MHC haplotype sequences but not in DR53, which indicated that these markers could not be used for genotyping of the HLA-DRB1 gene.
Identification of Functional Primers for M2_3_22 STR Marker Next to the HLA-DRB1
Sheikholeslam M., Vallian S.
Faculty of Science, Department of Biology, University of Isfahan, Isfahan, Iran

The association between HLA-DRB1 and autoimmune diseases has been demonstrated. Polymorphic microsatellites can provide useful genetic markers in HLA-related research, such as disease mapping. Among the markers associated with HLA-DRB1, the M2_3_22 STR marker was subjected to Bioinformatics study in order to find whether this marker could be applied as a specific marker for HLA-DRB1 analysis. At first, this marker was searched in the MHC database and uniSTS resource at NCBI. Then, the DR haplotype groups and six different MHC haplotype sequences containing DRB1 were retrieved from the nucleotide database at NCBI and investigated by the CLC main workbench software to find M2_3_22. Our data showed that this marker existed in all MHC haplotype sequences and was located approximately 17 kb upstream of the HLA-DRB1. The introduced primer pairs for the marker in uniSTS and dbMHC was not compatible with some MHC haplotype sequences. Therefore, conserved regions flanking the M2_3_22 with maximum similarity with all these sequences were selected and used to design new primers by CLC and Oligo softwares. The Bioinformatics tools developed by the NCBI, called E-PCR and BLAST, were used to determine the physical mapping position(s) targeted by designed primer pairs and based on the obtained results; the best primer pair was selected.
Bioinformatics Comparison of Hemagglutinin in Different Genotypes of Measles Virus

Mohabatkar H., Mahjoub M., Mohsenzadeh S.

Department of Biology, College of Sciences, Shiraz University, Shiraz, Iran

Since there are 23 genotypes for the measles virus, to study the similarities and differences between the genotypes, we performed computational study of an important protein of the virus, hemagglutinin (MVH). MVH is important because it is a surface glycoprotein and is responsible for mediating virus attachment to the host cell. Prediction of amino acid composition, transmembrane domains, secondary structure and glycosylation sites were performed by the predictprotein server. Prediction of B-cell epitopes was carried out by the BCPREDS server. The most prevalent residue in all of the proteins was leucine (65-68 residues). By aligning transmembrane domains of the MVH protein, three patterns: 1: MVH-A, 2: MVH-G3 and MVH-H1, 3: other 20 genotypes were found. For glycosylation sites, 5 patterns: 1: 5 N and 0 O glycolysation (15 genotypes), 2: 6 N and 0 O glycosylation (5 genotypes), 3: 4 N and 0 O glycosylation (1 genotype), 4: 5 N and 1 O glycosylation (1 genotype), 5: 4 N and 1 O glycosylation (1 genotype) were predicted. The highest percentages of the secondary structures were predicted in the following viruses: α-helix: MVH-C2 and MVH-H2 (19.45%), beta-strand: MVH-D7 (35.82%), MVH-D10, MVH F and MVH-H1 (35.66%, 35.49% and 35.33%), random coil: MVH-B1 (48.62%), MVH-C1, MVH-D2 and MVH-D8 (48.14%), buried residues: MVH-D2 (56.08%) and exposed residues: MVH-D1 (46.19%). For B-cell epitopes, in general, ten epitopes were predicted. Epitopes in positions 3, 6 and 8 were present in all genotypes and epitopes 2, 4 and 5 were present in 22 genotypes, while epitopes in position 1 and 10 were involved in a few genotypes (16 genotypes). Computational studies of this kind can help in better understanding of viruses, particularly for those viruses with different genotypes or subtypes.
Genome-Wide Screening for TnrA Regulated Genes of *Bacillus Clausii* Associated with a TnrA Box

Farazmand A., Yakhchali B., Minuchehr Z., Shariati P.

*Iranian Research Organization for Science & Technology National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.*

*Bacillus clausii* TnrA is a member of the MerR family that regulates gene expression in response to changes to nitrogen availability. The general transcription factor TnrA of *B. clausii* activates and represses gene transcription when nitrogen is limited for growth. In order to obtain members of the TnrA regulon, responsible for assimilation of various nitrogen sources in *B. clausii*, we performed a genome-wide screening for TnrA regulated genes associated with a TnrA box.

Common regulatory signals with a conserved DNA motif, 5'-TGTNA-7N-TNAC-3', were used as our query sequences against the Nucleotide Basic Local Alignment Search Tool at NCBI in order to find the putative TnrA boxes within the whole-genome sequence. We identified TnrA-regulated genes of *Bacillus clausii* associated with a TnrA box. *Bacillus clausii* is a potential member of the Asp/Glu/Hydantoin racemase family with regard to the TnrA regulons.
Application of Nucleotide Substitution Models in Reconstructing Phylogenetic Tree

Nassaj Hoseini S. M., Shamsbakhsh M.

Tarbiat Modares University, Tehran, Iran

Use Molecular phylogenetics has become widespread and is useful in Comparative Genetics; however, estimating phylogenetic trees is not always easy. Phylogenetic trees are an intuitive way to infer relationships among copies of a gene or among loci of a multigene family. As such, phylogenetics facilitates analysis of gene duplications, rates of evolution, polymorphisms, recombination, divergence of lineages and population demographics. Accurate estimates of evolutionary parameters are often dependent on the validity of a single phylogenetic reconstruction upon which inference is based. Inaccurate estimation of trees may lead to biased results and erroneous inference of processes or mechanism of evolution. Some investigators use a model of nucleotide substitution model to estimate evolutionary parameters such as branch lengths and tree topology. There are many models to choose from, and use of the optimal model for a particular dataset is important to avoid a loss of power and accuracy in phylogenetic estimations. Here, we review some molecular evolutionary forces and parameters (such as transition/transversion ratio, nucleotide substitution rate, nucleotide frequency, rate heterogeneity over sites) included in some common models of evolution used to interpret resulting patterns of molecular variation. We describe some statistical methods of selecting a particular model of nucleotide substitution that help us to select the best-fit model for a particular dataset through likelihood ratio tests or information criteria.
Comparison of the Alpha-Amylase Proteins from Several Amylolytic Microbial Strains

Moradi M., Shariati P., Tabandeh F. Mirzaee Zadeh Z., Aliakbari F.

National Institute of Genetic Engineering and Biotechnology, Tehran, Iran

Three bacterial strains and a fungal strain with high amylolytic activities were isolated from three industrial sites in Iran. The Bacterial strains were identified as Bacillus amylo liquefaciens, Bacillus cereus and Bacillus licheniformis by means of morphological, biochemical and molecular (16s rRNA) studies. Also macroscopic, microscopic and molecular (18s rRNA) methods identified the fungal isolate as Aspergillus oryzae. All the above strains possess the important amylolytic enzyme Alpha-amylase. By means of different Bioinformatics methods, the sequences of alpha-amylase proteins belonging to similar strains were investigated and phylogenetic trees were constructed, which showed that among the bacteria, B.cereus had the most homology with A.oryzae when compared to the other strains. In this regard, most of the homology was observed in domain A of the alpha-amylases. In fact, the microbial alpha-amylase active site is generally situated in the domain A region (between amino acids 200-400).
Barcoding of Life: A Tool for Insect Identification

Karimi J. , Hassani M.

Department of plant protection, Faculty of agriculture, Ferdowsi university of Mashhad, Mshhad, Iran.

Thrips tabaci Lindeman is an extensively distributed pest insect in many countries such as Iran that affects plants through direct feeding on and thus damaging the leaves and flowers of crops. At the same time, it also causes damage as a vector of different viruses. As a basic first step to control pests is the process of authentic identification, which often is very specific and, needs the ability to correctly identify the species. However, the inability to determine morphological characters of the thrips species due to absence of expertise and limitation caused by polymorphism, sex, and stage of development, leads both to poor control in the field as well decrease in exports due to the presence of species of quarantine importance. Therefore, existence of a quick and developmental-stage, non-limiting method for the identification of this species is of vital importance in the study of vector transmission, insecticide resistance, biological control and quarantine. In this research, we tried to introduce a molecular technique to identify T. tabaci as an economically important species present in Iran. Four populations of T. tabaci were collected from Mashhad and the vicinity. The method was based on nucleotide sequencing analysis of the mitochondrial cytochrome oxidase I (COI) gene using Folmer's primer set. Phylogenetic analyses conducted by the neighbor-joining and distance matrix methods yielded almost identical phylogenetic reconstructions of trees that separated Thrips based on geographical origin. Molecular data indicated that different thrips species are located in distinct groups and T. tabaci of Palestine is a clade most closely related to variants from Iran. These results show that molecular keys can be a useful method to provide much-needed information on thrips identification for pest management officers and quarantine purposes.
Comparison of the IturinA Operon in Bacillus Species

Aliakbari F.1, Moradi M.1,2, Lotfinia M.3, Mazhar S. F.4, Shariati P.1, Azimzadeh Jamalkandi S.1,5

1Department of Industrial and Environmental Biotechnology. National Institute of Genetic Engineering and Biotechnology. Tehran. Iran.


3Department of Microbial Biotechnology, Esfahan University. Esfahan Iran.

4Department of Microbiology, Azad Islamic University, North Tehran branch. Tehran. Iran.


A bacterial strain producing antifungal agent was isolated from food stuff. This isolate was identified as *Bacillus subtilis* using morphological, biochemical and molecular (16s rRNA) methods. Iturin is a lipopeptide antifungal, which is usually produced by *Bacillus subtilis*. The *iturinA* is an operon in *B. subtilis*, which consists of the *ituD*, *ituA*, *ituB* and *ituC* genes containing 38845 nucleotides. Bioinformatics methods were employed to investigate and analyze the sequences of the *iturinA* operon and its proteins belonging to similar strains, which showed that this operon in *B. subtilis* had homology with *Bacillus amyloliquefaciens kdg* gene coding for the KdgK protein. In addition, the *iturinA* operon and some of the proteins that it codes for were homologous to the *bacillorin*, *mycosubtilin* and *bacillomycin D* operons and their respective proteins in another species of *Bacillus*. Multiple sequence alignment was carried out for proteins which were coded by these genes of the mentioned operons as well as the KdgK protein and a phylogenetic tree was constructed. These results show that these strains have interspecies affinity.
Microsatellite Diversity and Population Genetic Structure of Rutilus frisii kutum in Mazandaran coasts

Rezaei M., Shbani A., Shabanpour B., Kashiri H.

Department of Fisheries, Gorgan University of Agriculture Science and Natural Resources, Gorgan, Iran

Rutilus frisii kutum is a very valuable commercial fish found in the southern parts of the Caspian Sea and is in great demand, due to its taste. Investigations show that natural reproduction of this species has decreased during the last years. To restock this valuable species, about 200 million fry were produced by artificial breeding and released into the sea annually. Loss of genetic stocks and the gradual depletion of the gene bank are parameters of great concern in the long run. The present study employed ten microsatellite loci to investigate levels of genetic variations of the Rutilus frisii kutum in Tajan and Tonekabon region of the Mazandaran province. The results showed no genetic differentiation among regions by Fst, Rst and AMOVA, and a relatively high level of gene flow was found between populations. Genetic variation of the two regions (Tajan: mean number of alleles per locus, Na= 9, mean effective number of alleles, Ne= 6.13, observed heterozygosity, Ho= 0.79 and expected heterozygosity, He= 0.81) (Tonekabon: Na= 7.9, Ne= 5.54, Ho= 0.80 and He= 0.78) were not statistically different. There were evidences for genetic bottlenecks in the populations. Protection and restoration of habitat may help increase the population size and lower the risk of vulnerability of the species in the future.
Molecular Characterization and Functional Analyses of a Novel Cold Stress-Regulated and Chloroplast-Targeted Protein from Cereals

Valiellahi E., Farsi M., Niazi A.

Department of Biotechnology and Plant Breeding, College of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.

Cold acclimation is a multigenic trait that allows hardy plants to develop efficient tolerance mechanisms needed for winter survival. To determine the genetic nature of these mechanisms, several cold-responsive genes of unknown function were identified from cold-acclimated wheat (Triticum aestivum). To identify the putative functions and structural features of these new genes, integrated genomic approaches of data mining, expression profiling, and Bioinformatics predictions were used. We herein report the structural heterogeneity of cDNAs, distribution, low temperature-specificity and protein structure of the identified Wcor14 gene. Analyses of the cDNA and genomic DNA sequences by the Vector NTI 9.0 software, suggested that Wcor14 and its related sequences constitute a small multigene family with different intron sizes. The deduced WCOR14 polypeptide is a hydrophobic polypeptide with 140 amino acids (MW=13.5 kDa), showing high homology to the previously identified wheat and barley COR proteins. No homologous sequences were found in other organisms suggesting that this family is specific to the plant kingdom. The highly homologous signal peptides of WCOR14, BCOR14b and WCS19 contained one putative 14-3-3 protein recognition motif. In this motif, the S-residue was predicted as a phosphorylation site and besides this, four other putative phosphorylation sites in WCOR14 were predicted by the NetPhos version 2.0 software. Comparative analyses of gene expression profiling shows that the expression of this gene is correlated with the development of freezing tolerance in cereals. Protein structure was predicted by SCRATCH server.
The Use of DNA Barcoding for Identification and Classification of Cyst Forming Nematodes

Amid motlagh M. H. 1 Valiellahi E. 2, Shahabi M. 1

1 Department of Plant Protection, College of Agriculture, Shiraz University, Shiraz, Iran
2 Department of Plant Biotechnology, College of Agriculture, Ferdowsi University, Mashhad, Iran

DNA barcoding is a novel system designed to provide rapid, accurate and automatatable species identifications by using short, standardized gene regions as internal species tags. Plant parasitic nematodes are well-known pests in agroecosystems and are also thought to exert an important influence on the structure and stability of natural plant communities. Nematode species are very difficult to distinguish; over 80,000 have been described, of which over 15,000 are parasitic. It has been estimated that the total number of described and undescribed nematode species might be more than 500,000. Precise identification of the components of natural plant parasitic nematode communities is a prerequisite for other studies. Traditional morphology-based assessments are time-consuming and require specialists. In addition, many taxa can be diagnosed only from adult male or female-specific structures, or from population measures of relative morphological characters, but this process will be greatly facilitated by a DNA barcoding system. In this study, the 18S rRNA gene was taken as DNA barcode for phylogenetic analysis of the Heterodaridae family. The 18S rRNA gene has been obtained from the NCBI database and used for this analysis. The sequence alignments were performed by Vector NTI Advance v.11 and phylogenetic analysis was carried out by the MEGA4 software. Alignment was constructed containing all available full-lengths of the small subunit ribosomal RNA gene of cyst-forming nematodes. The object of this study was to evaluate genetic species variation and phylogentic relationships of the Heterodaridae family. Analysis of the available nematode full-length 18S rRNA sequences suggested that this gene is a good candidate and an effective marker for this purpose e as in many cases even closely related taxa were shown to have differences in their 18S rRNA sequence.
Understanding of the protein-protein interaction capacity improves our knowledge of how living systems function. Many critical processes are regulated by protein machines composed of more than one subunit. Among these processes are gene duplication, gene regulation, signal transduction, membrane fusion and so on. This indicates that protein-protein interactions very often rely on cooperative effects to keep proteins in the functional conformation. Here, we focus on the four-helix bundle machinery called the SNARE complex which mediates membrane fusion during neurotransmitter secretion in the nervous system. Since deposition of neurotransmitters into the synaptic space is the final step in the conversion of the membrane electrical signal to the chemical signal, neurotransmitter secretion must be synchronized. On the other hand, crowded intracellular vesicles involving neurotransmitters present a question of how secretory vesicles fuse onto the plasma membrane in a synchronized fashion. Complexin is one of the most experimentally studied proteins that regulate assembly of the fusogenic four-helix SNARE complex to synchronize with neurotransmitter secretion. For computational-based study of complexin interaction with the SNARE complex, we used MD simulation. Salt bridges and H-bonds between the complexin central helix and synaptobrevin stabilize the binding of complexin binding to the neural SNARE complex. Complexin alpha accessory helix is also necessary for tight binding of complexin to the SNARE bundle. Our results also showed that the complexin accessory helix and synaptobrevin interact with the same region of syntaxin. This suggests that the alpha accessory helix can compete with the docking of synaptobrevin and this, so, inhibit the formation of the fusogenic SNARE machinery which can clamp the neurotransmitter secretion.
Motif Search in DNA Sequences Using Fuzzy Logic
Mohammadi-Achajeloii S., Ahrabian H., Zare-Mirakabad F., Nowzari-Dalini A.

Center of Excellence in Biomathematics, School of Mathematics, Statistics, and Computer Science, University of Tehran, Tehran, Iran

Motif searches in DNA sequences is one the most challenging problems in bioinformatics. Computational prediction and analysis of transcription regulatory regions in DNA sequences have the potential to accelerate greatly our understanding of how cellular processes are controlled. Several algorithms for this purpose have been proposed, but still finding an algorithm for getting better results is a challenging problem. This paper presents a novel method for finding motifs in DNA sequences using the fuzzified hidden Markov model (HMM). Hidden markov models are known to be a powerful method for solving bioinformatics problems. Fuzzy hidden Markov models or generalized HMMs (GHMM) are a novel type of hidden Markov models based on fuzzy sets and fuzzy integrals which generalize the classical stochastic hidden Markov models. In the GHMM the proper definitions of the fuzzy forward and backward variables are utilized to define a fuzzy modification of the classical Viterbi algorithm to determine the fuzzy optimal state sequence. Using the fuzzy HMM instead of the classical HMM will enable us to get better results.
R Package, a Powerful Tool for Bioinformatics and Its Role in Statistical Analysis of Phylogenetic Trees

Adham Fani Maleki¹, Balal Sadeghi², Kobra Ahmadvour²

¹ Animal Science Department, Ferdowsi university of Mashhad, Mashhad, Iran
² Department of Animal Science, College of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

A phylogenetic tree shows the evolutionary relationships among various biological species or other entities that are known to have a common ancestor. Phylogenetic trees among a nontrivial number of input sequences are constructed using computational phylogenetics methods. Distance-matrix methods such as neighbor-joining, which calculate the genetic distance from multiple sequence alignments, are simplest to implement, but do not involve an evolutionary model. More advanced methods use the optimality criterion of maximum likelihood, often within a Bayesian framework, and apply an explicit model of evolution to phylogenetic tree estimation. The R programming language has been proved to be a powerful tool for bioinformatics. In this article, we describe the computer package apTreeshape that is dedicated to simulation and analysis of phylogenetic tree topologies using statistical indices. Beyond the software facilities for data analysis and graphical display offered by the R language, apTreeshape includes important corrections on classical shape statistics. One strength of the package is to present new tests based on the statistical theory of likelihoods, and therefore provide optimal power for testing null models of macroevolution. The apTreeshape package integrates recent development in the statistical theory of imbalance measures, which warrant the optimality of some testing procedures. The R package like the Tree-building methods can be accessed on the basis of several criteria, such as efficiency, power, consistency, robustness and falsifiability. In this comparison, apTreeshape benefits from the extended power of R for performing all types of data analyses (and its facilities for connecting to public databases). This should make this resource attractive to R users.
Analyzing ChIP-Seq Data

Mohammadian Mosamam A. 1, Bonifer C. 2, Westhead D. 2

1 Department of Mathematics, University of Zanjan, Iran- Institute of Molecular and Cellular Biology, Iran
2 Institute of Molecular Medicine, University of Leeds, United Kingdom

Chromatin Immunoprecipitation (ChIP) is a technique to detect interactions between proteins and DNA, which is based on the enrichment of DNA associated with a protein of interest. The ChIP-Seq technique combines ChIP with new massively parallel short-read sequencing and is proven to be a powerful method to identify genome-wide DNA binding sites for a protein of interest which has overcome several limitations inherent to microarray analysis of ChIP. Our preliminary work with ChIP-Seq has shown that the method can delineate specific marking of chromatin and binding of the cellular transcriptional machinery to precisely defined regions of chromosomes, improving significantly in resolution, sensitivity and precision by comparison with older techniques involving DNA microarrays. Here, we study the dynamic process of cellular differentiation in cells of the immune system. This brings significant space-time statistical challenges, in particular the need to quantify not only the genomic loci with specific markings and bound transcription factors but how these change in both position and amplitude with time. We outline preliminary studies for statistical analysis of such ChIP-seq data that are generated by the Illumina genome analyzer. The enrichment of the genome wide alignments is visualized by the Hilbert curve and also the UCSC genome browser as an exploratory analysis in order to discover meaningful dynamic patterns of the peaks.
Prediction of Three-Dimensional (3D) Structure and Function of ATAXIN-2 Protein in Rice

Heidari Sh. 1, Ebrahimie E. 1, Heidari P. 2, Assad M. T. 1

1College of Agriculture, University of Shiraz, Shiraz, Iran.

2College of Agriculture, University of Zanjan, Zanjan, Iran.

The major goal of this research was to predict the tertiary structure of ATAXIN-2 as well as its function in plants by Bioinformatics tools. Protein has a specific chemical function and each protein has an unique three-dimensional (3D) structure. We observed up-regulation of ATAXIN-2 EST in somatic embryogenesis in callus. ATAXIN-2 protein is a protein of unknown function in rice. This protein was widely expressed in animals in early embryonic development. However, there is no report on the role of ATAXIN-2 in plant embryogenesis. We suggested that the ATAXIN-2 protein possibly plays a key role in the embryogenic process in plants. Since, the function and structure of ATAXIN-2 in rice plants is unknown, in this research, we investigated probable function and tertiary structure of ATAXIN-2 protein using data banks and softwares. The human ATAXIN-2 protein structure (PDB ID: 2ERR) with 397 amino acids was estimated by the Rasmol software. The analysis shows that this protein has 2 alpha-helix and 5 beta-sheets. There is no PDB file for the rice ATAXIN-2, so we studied tertiary structure with the phyre server. The Rice ATAXIN-2 protein structure with 513 amino acids matched with the c2d9iA protein using the structural classification of proteins (SCOP) with an e-value 2.3×10⁻⁹. The results showed that the rice ATAXIN-2 protein has 2 alpha-helices and 4 beta-sheets, very similar to the human one. For prediction of the rice ATAXIN-2 function, we used the string server. The human ATAXIN-2 function is RNA binding, playing a key role in regulating the splicing network and survival of the human embryonic stem cells. In contrast, we predicted rice ATAXIN-2 function as a NEDD4-binding protein. The Nedd4 protein is highly expressed in many mouse embryonic tissues and regulated gene in the mouse. Nedd4 has a fundamental role to play in embryonic processes. However, whether Nedd4 has such a function is currently unknown. Finally, a comparison between motif, secondary structure and amino acids of rice and the human ATAXIN-2 protein was performed.
Application of GenAlEx and TFPGA Softwares for Identifying Fall and Spring Migrating Caspian Salmon (Salmo trutta caspius Kessler, 1877) by Using AFLP Molecular Marker

Habibi E.

Department of Fisheries, Faculty of Marine sciences, Tarbiat Modares University, Tehran, Iran.

Application of GeneticS and Biotechnology based on analytical softwares currently has a major role in accurate data analysis such as Bioinformatics tools in Biological Sciences. In this regard, GenAlEx and TFPGA softwares were applied for evaluating the possibility of identification of fall and spring migrating Caspian salmon using AFLP molecular markers. Salmo trutta caspius is a valuable endemic species in the Caspian Sea that is indanger of extinction. There are two migrating forms of spring and fall run with different appearance characteristics. Twelve different EcoRI /MseI primer combinations were used from which a total of 807 bands were produced, 227 of which were polymorphic loci. In the GenAlEx software, the calculated mean of polymorphic loci was 80.62%. Also, Nei’s genetic distance was 0.034% and the variance between and within the two populations were 6% and 94%, respectively. The UPGMA dendrogram based on population genetic similarity was drawn by the TFPGA software. The results showed that there were no significant genetic variations between fall and spring migrating Caspian salmon. Therefore, spring forms are probably a part of Fall-run population and the difference could be attributed to physiology or other related factors.

Keywords: GenAlEx, TFPGA, Softwares, AFLP, Salmo trutta caspius
In Silico Expression Analysis of a Protein Phosphatase 2C Subfamily in Rice

Aghagolzadeh P. 1, Ebrahimie E. 2, Shobbar Z. S. 3, Kashfi R. 1

1Department of Crop Production and Plant Breeding, Shiraz University, Shiraz, Iran
2Department of Crop Production and Plant Breeding, Shiraz University, Shiraz, Iran
3Department of Bioinformatics, Agricultural Biotechnology Research Institute of Iran

Protein phosphatase 2C family consists of a group of evolutionary conserved serine/threonine phosphatases which play a role in stress signal transduction. Two of which in Arabidopsis, including ABI1 and ABI2, are known as components of ABA signal transduction pathway. Their mutants are hypersensitive to ABA showing increased expression during seed dormancy and adaptive responses to drought. Considering sensitivity of rice to abiotic stresses, particularly drought, study of this gene family in rice and their role in response to stress would be beneficial. Eight OsPP2C proteins were found in rice (OsPP2C1 to OsPP2C8), carrying all the conserved motifs of this subfamily. In this research, the expression patterns of these genes were analyzed in response to drought and salt stresses and in different tissues. The related row microarray data were collected from “Affymetrix” and “PLEXdb” websites. The statistical analysis revealed that the expression of OsPP2C5 has upregulated considerably under drought and salinity condition (~2.7 and 2.2 fold respectively) while the rest of the genes are slightly drought and salt responsive (~1.5 fold). As well, the expression level of OsPP2C5 was varied among different tissues while no significant changes were observed for the rest. Developing seeds (during day 21-29) contained the highest amounts of OsPP2C5 transcripts amongst the 30 studied tissues, when ABA content is increased during seed desiccation and dormancy acquisition. The promoter analysis also confirmed the existence of ABA responsive cis-elements in OsPP2C5 promoter. Then, involvement of this gene in the abiotic stresses tolerance operating through ABA is probable.
A Novel View of Nano-Molecular Modeling Study of the Synchronize Interaction between Tamoxifen and Aspirin with Human Serum Albumin

Pourgonabadi S.¹, Saberi M. R. ², Chamani J.¹

1. Department of Biology, Faculty of Science, Islamic Azad University-Mashhad Branch, Mashhad, Iran,
2. Medical Chemistry Division, School of Pharmacy, University Complex, Mashhad, Iran

Molecular modeling method has been employed to understanding the interaction of drugs and proteins [1]. The investigation of 3-D structure of crystalline albumin showed that HSA contains three homologous domains (Ι,Π and Ω), Ι (residues 1-195), Π (196-383), and Ω (384-585) and each domain can be divided into two sub-domains (A and B) [2]. The crystallographic analysis has revealed that HSA has binding sites of compounds within hydrophobic cavities in sub-domains ΠA and ΩA, which are corresponding to site Ι and site Π, respectively and role tryptophan residue (Trp 214) of HSA is in sub-domains ΠA [3]. There is a large hydrophobic cavity present in sub-domains ΠA that many drugs can bind to that. The autodock 4.0 program was used to calculate the interaction modes between aspirin and tamoxifen with HSA. The Lamarckian genetic algorithm (LGA) implemented in autodock was applied to calculate the possible conformation of the drug that bind to protein. During the docking process, a maximum of 10 conformers was considered for the drugs. The conformer with the lowest binding free energy was used for further analysis. Aspirin is the best located within the binding pocket of subdomain. Furthermore, there is one hydrogen bond between aspirin and HSA. There is interaction between Lys199 and aspirin with the length of the hydrogen bond 2.68Å. There is no hydrogen bond between tamoxifen and HSA, therefore exhibit the optimal energy ranked result of tamoxifen interaction with the residues of HSA and the tamoxifen-HSA space-fill conformation.
Probing the Binding Cavity of Lomefloxacin on HSA Based on its Comparative Docking with Ciprofloxacin

Vahedian-Movahed H. ¹, Saberi M. R. ², Chamani J.¹

¹. Department of Biology, Faculty of Sciences, Islamic Azad University-Mashhad Branch, Mashhad, Iran

². Medicinal Chemistry Department, School of Pharmacy, Mashhad University of Medical Sciences, Vakilabad Boulevard, Mashhad, Iran

Due to the importance of lomefloxacin (LMF), a fluoroquinolones antibiotic, there are numerous experimental works suggesting LMF as a drug carrying by Human Serum Albumin (HSA). Also, we have determined the binding constant and the number of binding site based on fluorescence data. However, the binding cavity of the LMF on the HSA is not known yet. Since the identification of the binding cavity of a drug on HSA can help us to predict drugs-drugs competition, in this study we probe the binding site of LMF by a comparative docking with ciprofloxacin (CIP), in which there is some evidence on its binding cavity. The principal ligand binding regions on HSA are located within domains II and III. Moreover, it is known that CIP site is located somewhere on domain II. So, we applied autodock4 program to calculate the possible conformations of LMF and CIP on HSA based on Lamarckian genetic algorithm by adopting three different grid box sizes, in such a way that it encompasses the whole proteins domains, domain II and domain III separately. Finally, the best docking energy results were assumed as the possible candidates for ligand-protein interaction. This conclusion is supported by experimental data of fluorescence spectroscopy which will be presented soon.
Searching Genomes of Retroviridae Family for Riboswitches
Lotfinia M., Mohamad Beigi M.

Microbial biotechnology, Department of Biotechnology, Isfahan University, Iran

New regulatory RNAs with complex structures have recently been discovered. Novel types of these RNAs are riboswitches. Riboswitches are folded RNA domains that serve as receptors for low molecular weight metabolites such as guanine, glycine, coenzyme B12. These domains are usually located in the 5' untranslated regions of such mRNAs. They bind to small metabolites and upon binding, conformational changes occur that trigger the expression of a certain gene. Novel validated cellular targets for antibiotic therapy are riboswitches as unique RNAs that can be used as selective receptors for small drug-like metabolites. There have been fewer studies on finding metabolic riboswitches of viruses than those of prokaryotes and eukaryotes. In this study, we surveyed genomes of the Retroviridae family including 58 members, by RNA motif search web server RIBOSWITCH FINDER (http://www.biozentrum.uni-wuerzburg.de/bioinformatik/Riboswitch) in order to identify riboswitches. This software predicts RNA secondary structures based on the calculation of minimum free energy. We used the FASTA format of Retroviridae from GenBank (http://www.ncbi.nlm.nih.gov/genomes/genlist.cgi?taxid=10239&type=5&name=Viruses) and transferred it to the RIBOSWITCH FINDER. We identified two guanine-like riboswitches in the Fline immunodeficiency and Mason-pfizer monkey viruses. This study provides evidences that there are metabolic riboswitches in viruses and also presents the following idea; examination of guanine effects on these viruses in the laboratory in order to obtain information about new features of virus therapy.
Expressed Sequences Tags (ESTs) Analysis of Vernalization Process in Shoot Apical Meristem of Triticum Monococcum

Shamloo Dashtpagerdi R., Ebrahimie E., Razi H., Masoudi-Nejad A., Niazi A. 4

1 Department of Plant Production and Breeding, Faculty of Agriculture, Shiraz University, Shiraz, Iran

2 Department of Crop Production and Breeding, Faculty of Agriculture, Shiraz, Shiraz University, Iran.

3Department of Crop Production and Breeding, Faculty of Agriculture, Shiraz, Shiraz University, Iran.

4Institute of Biochemistry and Biophysics, University of Tehran, Iran

Prolonged exposure to low winter temperatures (vernalization) accelerates the progression from vegetative to reproductive growth in many plant species including temperate cereals. Triticum monococcum (2n=2x=14), generally known as an ancient diploid relative of wheat, is a member of temperate cereals which has gradually been recognized as an attractive diploid model for exploitation of useful traits, discovery of novel genes and variant alleles, and functional genomics. In this study, Bioinformatics analysis of expressed sequences tags (ESTs) was performed on 3363 EST sequences of shoot apical meristems of T. monococcum spp. Aegilopoides during vernalization process. A total number of 430 contigs and 2031 singletons were formed after assembly of the ESTs. Blastx search revealed that 359 contigs and 1395 singletons have distinct hits with uttermost e-value= 1x10^-5. Functional analysis of ESTs showed that sequences categorized in biological process, cellular component and molecular function groups which 511, 121 and 211 sub categories, respectively. Analysis of sub-categories demonstrated that 48 genes in cold responsive category were expressed significantly during the vernalization process. In addition, some of the genes responsible for to regulation of flower development category, such as the MADS-Box (AGL20 and AGL22) group, VIL1 and Ubiquitine conjugating enzyme1, genes represented significantly with a p-value= 0.00139. Interestingly, the VIL1 gene had two transcripts in both, cold responsive and regulation of flower development categories. Promoter analysis of this gene revealed that some transcription factors involved in cold stress, light and hormone responses can bind to this region. As a result, VIL1 has this ability to be expressed during both, stress response and developmental processes.
Differences of $ftSA$ Gene Sequence of Two Different Intracellular Bacteria: Wolbachia

Baradaran M.

Razi vaccine & serum research institute, Ahvaz, Iran

Wolbachia and Neorickettsia are two genera of intracellular bacteria from the order Rickettsiales. Wolbachia was first reported within reproductive tissues of the mosquito Culex pipens by Wolbach in 1924 and was initially classified as Rickettsia sp. But now it is classified in family Anaplasmataceae with the genera Ehrlichia, Neorickettsia and Anaplasma. Neorickettsia is a helminth-borne bacterial genus. Wolbachia occupies a position intermediate between the two tick-transmitted genera (Ehrlichia and Anaplasma) and Neorickettsia in the phylogenetic tree. However, wolbachia is not recognized as a vertebrate pathogen, since mammalian infections have never been documented. Wolbachia is cytoplasmically inherited and is found in reproductive tissues of a wide range of arthropods and nematods. The relationship between Wolbachia and its arthropod hosts ranges from mutualistic to parasitic. Wolbachia isolates have been found in numerous disease carrying insects and in parasitic nematods. Wolbachia causes a number of reproductive alterations in its arthropod hosts. In nematodes and some arthropod residences wolbachia is necessary for host reproduction. There is growing interest in the potential uses of Wolbachia as tools for biological control and genetic manipulation of pests and disease vectors. In this study, we alligned sequences of $ftSA$ (encoding cell division protein) gene from endosymbiont wolbachia of Drosophila and Neorickettsia with EMBOSS-Pairwise - Allignment and CLC Main Workbench softwares.
Phylogenetic and Functional Analysis of Zn/Cd/Pb-ATPase Pump Using Bioinformatics

Ashrafi E. ¹, Alemzadeh A. ², Ebrahimie E. ²

¹Department of Plant Production and Breeding, Faculty of Agriculture, Shiraz University, Shiraz, Iran.

²Department of Crop Production and Breeding, Faculty of Agriculture, Shiraz University, Shiraz, Iran.

In recent years, heavy metals have increased in natural ecosystems and make serious global problems for the environment. A number of plants, such as Thlaspi caerulescens and Arabidopsis halleri can naturally cleanup contaminated soils and groundwater by hyperaccumulating metals in aerial organs in a process called phytoremediation. The phytoremediation is an environmentally friendly, green technology that is cost effective and energetically inexpensive. One of the most important elements providing this outstanding attitude for these special hyperaccumulator plants is related to the P1B-ATPase transporter pumps. P1B-type ATPases belong to the large P-type ATPase family of membrane proteins that transport a variety of metals (Cd²⁺, Zn²⁺, Pb²⁺, Co²⁺, Cu²⁺, Ag⁺ and Cu⁺). against their concentration gradients, using the energy provided by ATP hydrolysis. P1B-type ATPases, unitage kind of metal transport, is organized into five subfamilies (I, II, III, IV and V). Uncovering and determining the protein characteristics of P1B-ATPases in hyperaccumulator plants can provide a new vista for engineering proteins for phytoremediation. In this study, using bioinformatics tools, a comprehensive set of protein features (2644 features) of first, secondary and tertiary structures of p1B-ATPases Subgroup II (Zn/Cd/Pb) in hyperaccumulator and non-hyperaccumulator were extracted and compared. The results revealed that frequency of negatively charged, positively charged, cysteine, histidine, glutamate, lysine Basic, acidic residues and frequency of cys-cys dipeptid, his-his dipeptid in hyperaccumulator plants are more than non-hyperaccumulator. Frequency of non-polar residues, psi angle means, count of strand, beta bulges, sheet in hyperaccumulator plants are less than other groups. In phylogenetic tree revealed, that Arabidopsis thaliana and poaceae family are closed to the hyperaccumulators.
A Molecular Modeling Study of Interaction between Amlodipine and Human Serum albumin Based on Lamarckian Genetic Algorithm by Adopting Three Different Grid Box sizes

Abdollapour N. ¹, Saberi M. R. ², Asoode A. ³, Chamani J. ¹

¹Department of Biology, Faculty of Sciences, Islamic Azad University-Mashhad Branch, Mashhad, Iran.

²Medicinal Chemistry Department, School of Pharmacy, Mashhad University of Medical Science, University Complex, Vakilabad, Mashhad, Iran.

³Department of Chemistry, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran

It is necessary to study the interaction between drugs and proteins in order to understand the mechanism of drug action at a molecular level. Parameters calculated from molecular modeling can be used here as input to our simulations in order to calculate bigger systems. Crystal structure analysis has revealed that HSA consists of a single polypeptide chain of 585 amino acid residues with. Three structurally homologous domains (I-III): I (residues 1-195), II (196-383), and III (384-585) that assemble to form a heart-shaped molecule. Six binding sites for various ligands were found in the albumin structure. Two important sites site I and II and the drug usually binds with high affinity to one of these site. We applied Autodock4 program to calculate the possible conformations of amlodipine on HSA. Study of molecular modeling also indicated that aspirin could bind to site II of HAS and amlodipine binds to site II of HSA mainly by a hydrophobic and hydrogen bond interactions between HSA and the drugs. The conformation information between amlodipine and HAS consist of inhibition constant = 12.41uM and intermol-energy = -6.57, binding-energy = -6.69 and ligand efficiency = -0.24 and electrostatic energy = -1.44 and 1hydrogen bonds formed 48:H at 2.108 Å of HSA. Our results have been confirmed by spectroscopic analysis.
Bioinformatics Analysis for Design and Construction of a Plant/Bacteria Hybrid SigmaFactor, for Specific Gene Expression in the Plant Plastids

Mohsenpour M., Tohidfar M., Babaeian Jelodar N.

Agricultural Biotechnology Research Institute of Iran (ABRII), University of Agricultural and Natural Resources, Sari, Iran

Application of plastid transformation technology is limited by the difficulty of obtaining regulated, selective expression of the transgenes. Constitutive expression of transgene products can be deleterious to the plants’ health because of toxicity or interference with metabolism. Interactions between heterologous products and metabolism in different growth stages reduce potential plant productivity or even inhibit selection of primary transformants. In addition, constitutive expression may cause different problems of even greater relevance if the products are harmful, for example by unintended human consumption of pharmaceutical transgene products. Thus, it is required to express a transgene at a particular developmental stage, or to harmonize the timing of expression with ecological concerns. There are no suitable specific inducible endogenous promoters available in plastoms; hence transgenes are constitutively expressed throughout plant development. As a consequence, plant growth is often heavily disturbed by the noxious effect of overproduction of the transgene product. In this study for resolving these problems, a system was designed by using bioinformatics analysis that based on construction of a hybrid plant/bacteria hybrid sigma factor. The principle of this system designing was this fact that the specificity of promoter recognition in bacteria is accomplished by transcription initiation factors, so-called sigma factors. RNA polymerase recruits alternative sigma factors as a means of switching on specific regulons. In this study a plastid vector was designed and construction containing desired gene under control of E. coli groE promoter. Chloroplast RNA polymerase unable to recognition of this heat shock promoter. Transcription from this promoter becomes possible only when a suitable sigma factor be targeted to plastid. This sigma factor must be able recognize either E. coli heat shock promoter or plastid RNA polymerase (PEP). By bioinformatics analysis, it is certain that C-terminal motifs of sigma factors involve in promoter recognition and N-terminal motifs interact with RNA polymerase. The chloroplast RNA polymerase likes prokaryote ones uses sigma factors that code in nucleus and target to plastid to active transcription.
of plastid genes. In this study we obtained the sequence of plant sigma factors from
Gene Bank and determined the regions of their active domains. Furthermore the
sequence of sigma32 also was received and analyzed. We succeed to construct hybrid
sigma factors by composition of the N-terminal motifs of tobacco sigma factors that
contains the chloroplast signal peptide and RNA polymerase interaction domains, with
C-terminal motifs of E. coli sigma32 that able to recognizes and binds to groE
promoters. These motifs was amplified by primer designing from plant and bacteria
sigma factor sequences and joined together by SOEing PCR technique by retaining the
main open reading frame (ORF) of hybrid sigma factor. The amplified sequences that
called Hsig was cloned in pTZ57R/T vector and then isolated by BamHI and Sacl
enzymes and ligated in pClB vector under PEPC green tissue specific promoter and 35S
terminator. In the next step the cassette of Hsig was digested from pCIB-HSig vector by
HindIII and EcoRI enzymes and recloned in T-DNA regions of pBI(-k) binary vector.
Finally the resulting vector called pBI(-k)-PEPC-HSig was used for plant transformation
mediated Agrobacterium for targeting the hybrid sigma factor to plastids.
Haplotype Inference with Greedy Parsimonious Tree Grow Method

Taheri M. ¹, Sadeghi M.², Aghaee Naser Gh. ³

¹Shikh bahaee university of Isfahan Iran, Isfahan
²National Center of Genetic and bio technology Tehran, Iran
³Faculty of computer Engineering-Isfahan Isfahan Iran

Haplotype inference is one of the most attractive challenge in Bioinformatics research area due to analysis of fine scale genetic cells and create a simple method for mapping inheritance disease to some pattern in SNP sequence from DNA of one person. But since direct gathering of haplotype information from biological data from time and cost is too costly. So that computational method for gathering this precious data is popular. One of these heuristic methods that suggest recently is PTG that try to create a parsimony tree and solve haplotype inference problem by padding this tree. We try to analysis this method and express some weak aspect of PTG in this paper and suggest a greedy method for modify original PTG algorithm to decrease random behavior of PTG and named this new version as greedy PTG .we implement greedy PTG and evaluate it with simulate and real data set. For estimate accuracy of result, we compare result of greedy PTG with 2SNP, GERBIL and PHASE as three most tools for Haplotype inference. Result of evaluation represents validity of approach.
Structural Bioinformatics Approaches to the Discovery of New Antimycobacterial Drugs

Momtazi A., Beigi M. M.

Department of Biotechnology, University of Isfahan, Isfahan, Iran.

Bioinformatics methods have the potential to speed up the drug discovery process and represent new targets for therapy. In this study, structural bioinformatics is used to accelerate and rationalize the process of antimycobacterial drug discovery and design, with the immediate goals to identify viable drug targets and produce a set of the critically evaluated protein target models and corresponding sets of probable lead compounds. To develop drugs to cure tuberculosis: we focus on two new methods that have been applied to discover new drug targets: homologous protein mapping (HPM)[1] and protein-ligand interactions network. The first method offers homologous protein mapping via constructing the protein-protein interaction (PPI) network of Mycobacterium tuberculosis H37RV using data in the DIP database (http://dip.doe-mbi.ucla.edu/dip). Proteins with high interactions in the network are selected as the best candidates for inhibition. It has been shown that molecular chaperones, ribosomal proteins and ABC transporters are highly interconnected proteins. But this method does not include some of the important targets in Mycobacterium tuberculosis H37RV just because of lower interconnectivity, for example proteins of the anabolic pathway of mycolic acid such as the enoyl-acyl carrier protein reductase (InhA), Ketoacyl-ACP synthase (kasA/kasB), also some drug resistance-related proteins such as cytochromes, antibiotic degrading enzymes and target-modifying enzymes, which can be selected as suitable targets for inhibition. The second method predicts a network of proteins that have similar ligand-binding sites and uses this network in order to find the one with higher connectivity and thereby discover drug targets. We suggest that a combination of both methods can be used for drug discovery. Using the second method we propose entacapone and tolcapone to inhibit a protein (with low connectivity in the protein-protein interactions network of the first method) in the pathway of mycolic acid synthesis that is known to be essential for mycobacterium growth.
Optimization of Primer Designing for TAIL-PCR Based Promoter Isolation

Ebrahimi A., Mahna N.

Department of Horticultural Sciences, Faculty of Agriculture, University of Tabriz, 5166616471, Tabriz, Iran

The most important factor for polymerase chain reaction is how to design the primers. An inappropriate primer design could lead to a less PCR product or the production of unwanted bands/fragments. At a worse condition, it could result even to a complete failure of PCR reaction. To avoid these undesirable consequences and to obtain a successful PCR product, it is necessary to take into the account some crucial points when attempting to design a correct primer or a set of primers, including primer length, primer specificity, GC%, ΔG, melting temperature, etc. TAIL-PCR is a method for identification and isolation of flanking sequences of a known sequence which is carried out using complex PCR cycles. Primer designing in this sophisticated method, is perhaps the most fundamental and important step. In this research, the aim has been to optimally design primers for TAIL-PCR based isolation of the promoter region of one of apple genes in which the work steps, bioinformatic methods used, and the features of the designed primers have been discussed. In conclusion, five internal specific primers, sixteen arbitrary ten-mer primers, and two degenerate primers were designed for this promoter isolation endeavor.
Wednesday Posters
In Silico Elucidation of Iron Acquisition Genetics in Rosaceaous Fruit Trees

Ebrahimi A., Mahna N.

Department of Horticultural Sciences, Faculty of Agriculture, University of Tabriz, 5166616471, Tabriz, Iran.

There are two strategies for iron acquisition in plants. The plants expressing each strategy, express the other activity to some extent. The genetics of both strategies are going to be more investigated and several genes with a probable role in this process have already been identified. This study examines the existence and expression of these genes in the Rosaceaous fruit trees such as apples and some Prunus trees which have rich EST databases. To achieve this goal, an in silico approach were exploited and accordingly, all EST’s available were translated and compared to the proteins involved in each of strategies. The results showed that there are some genes in fruit trees with a significant similarity to their counterpart in other plants. Similarity data, E-values, accession numbers and so forth have been illustrated and the phylogenetic relationship of the proteins has been presented.
Codon Optimization of cry6 in Order to Increase its Expression Level in Potato with Using Bioinformatics Analysis

Abedini R., Mohsenpoor M., Tohidfar M.

Agriculture Biotechnology Research Institute of Iran (ABRII)

Codon usage optimization that is adaptation of the codon usage of the transcript gene to the typical codon usage of host, caused to increase the foreign gene stability and expression levels in target host. In this study we optimize cry6 gene codons for high level expression in Solanum tubersonum (potato) plants. For this purpose the sequence of cry6 gene from Bacillus thuringiensis was obtained from Gene Bank, and it ’s main open reading frame (ORF) was recognized. Subsequently the number of each amino acids in this genes was calculated. Further more the percentage of potato codon usage was obtained from codone usage database (http://Genomes.urv./OPTIMIZER). We tried to close the percentage of each codons of cry6 gene to the same codons percentage in potato genome, by using bioinformatics software. It is expected that cry6 codon usage optimization increase the stability, and expression of this genes in transgenic potato, and increase the resistance of plant against pests.
Structureal and Kinetic Studies of native 5-enolpyrovil shikimate -3-phosphate Synthase with Mutant and CP4 EPSPS in order to obtain Glyphosate Tolerance

Ghavami M., Amani J., Salmanian A. H., Mousavi A.

National Institute of Genetic Engineering and Biotechnology, Iran

5-enolpyrovil shikimate -3-phosphate synthase (AroA Ec.2.5.1.19) is an essential enzyme of the shikimate pathway and is the target for the glyphosate based herbicides such as Round up. Glyphosate tolerance activity of the enzyme could be obtained by native occurrence such as the EPSPS of Agrobacterium strain CP4 or by site directed mutagenesis of E.coli EPSPS on some critical residues(Gly96/Ala,Ala183/Thr).

In the present study function and structure of native E.coli EPSPS were compared with 2mut EPSPS (Gly96/Ala,Ala183/Thr) and CP4 EPSPS by nucleotide and polypeptide back bone chain alignment. Then we investigated on their kinetics and crystallography structures. Further analysis is under progress.
The Study of 3D-Structure of SUMO4 Based on Homology Modeling

Hosseini E. S., Erfanitabar F., Zarrabi M.

Biotechnology Department, Alzahra University, Tehran, Iran

Four small ubiquitin-related modifier (SUMO) genes termed as SUMO1, SUMO2, SUMO3 and SUMO4 have been identified in humans. SUMO1, 2 and 3 have a wide tissue distribution while SUMO4 distribution is tissue or organ dependent with high-level expression in immune tissue. The characterization of SUMO4 expression in immune cells and pancreatic islets provide a foundation for demonstration of its role in pathogenesis of Type I Diabetes (T1D).

In this paper, by focusing on the significant role of SUMO4 in T1D pathogenesis, we will discuss the importance of detecting a 3D-structure for this protein. The applied method is based on comparative modeling specially comparison scores of E.value and the amount of identity between relevant proteins that have similar sequence of amino acids with SUMO4. By aligning sequences in CLUSTALW database, we try to find the similarity and differences between members of SUMO family. It is discovered that SUMO2 must have the closest structure to SUMO4. So the based on x-ray structure of SUMO2, the three dimensional structure of SUMO4 is constructed by comparative homology modeling using MODELLER. The refined model is then investigated in terms of the structural features to improve our understanding of the SUMO4 function. The results of homology modeling which are in good agreement with the experimental data, indicates that our 3D model structure of the SUMO4 is reasonable and therefore, can be used for more understanding its role in T1D pathogenesis.
Homology Modeling Of Functional Domain of PyocinS2 from the Crystal Structure of ColicinE7 and Studying Of PyocinS2-Immunity Protein Interactions

Zarrabi M., Dalir Fardooi R., Tavajjohi Z.

Alzahra University, Faculty of Science, Department of Biology, Tehran, Iran

Pyocins are bacteriocins produced by Pseudomonas aeruginosa to kill other bacteria in the same species. PyocinS2, a type of S pyocin, is a protein which is soluble and sensitive to protease and has a potential of DNase activity and also inhibits lipid synthesis. PyocinS2 makes up a complex of two proteins in the producing cells: Killer protein and Immunity protein. Killer protein shows DNase activity and this activity is inhibited by the immunity protein and by producing this protein, bacterial cell survives, consequently. From the point of medicine and pharmaceutics, pyocinS2 is important because it is proved that it possesses antimicrobial and anticancer effects. Therefore determining secondary and tertiary (3D) structure of pyocinS2, can be useful to study and investigate manner of its function. Because of the pyocinS2 biochemical properties mentioned above, its purification and experimental determination of 3D structure are very difficult. In this paper, we have proposed a 3D structure of pyocinS2 by homology modeling. The crystal structure of colicinE7 was chosen as the template based on high sequence identity with the protein. The refined model was then characterized in terms of the structural and the interactional features to improve our understanding of the pyocin function. In order to investigate binding mode between the pyosinS2 and immunity protein, protein-protein docking simulation was carried out after generating 3D structure of the pyosin using homology modeling method. Based on these predicted structural features of pyocinS2-immunity protein Complex, further theoretical and experimental studies are proposed to continue to elucidate the structure and function of pyosinS2.
Phylogenic Investigation of Plants Based on Alignment of p5CS Gene Product Sequences

Sheikhehpour S., Kahrizi D.

Agronomy and Plant Breeding Department, Razi University, Kermanshah, Iran.

The pyrroline is an osmolyte and stress amino acid. This amino acid regulates osmosis pressure, inhibits minerals ions destroy and suppresses protein degradation processes. The delta 1-pyrroline-5-carboxylate synthetase (p5CS) is a key enzyme in the pyrroline biosynthetic pathway. Expression of this enzyme is associated with pyrroline synthesis. Hence, studying this enzyme is highly necessary for plant abiotic stresses researches. For this aim 32 p5CS sequences from 16 plant species were derived from Gene bank and analyzed with DNASTAR Bioinformatics package. The phylogenic studies results showed that there is high similarity for this gene product among plants intraspecies. Based on this result the Fabaceae, Brasicaceae and Poaceae species were assiged to the separate clusters. These results confirm the closed genetic relationship of species within a family. The alignment results showed that the 532-573 and 697-745 regions were detected as more conserved residues, which may have important biological function, such as the active site area.
Phylogenic Study and Alignment of p5CS Gene Product in Poaceae Species

Sheikhehpour S., Kahrizi D.

Agronomy and Plant Breeding Department, Razi University, Kermanshah, Iran.

The plants during drought and salinity stresses induce osmolyte substrate in their cells for against this stresses. Pyrroline is an amino acid that is induced during osmolyte stresses. In osmolyte stresses, mineral ions cause degradation of proteins and cell membranes. Pyrroline acts as an inhibitor for this degradation process. Several enzymes catalyze the pyrroline biosynthesis pathway. The first enzyme in this pathway is delta 1-pyrroline-5-carboxylate synthetase (p5CS) Biosynthesis of this enzyme is correlated with increasing pyrroline levels in plant tissues. For this purpose 10 sequences of this gene from 6 Poaceae species were analyzed. These sequences were collected from Gene Bank and were studied with the DNASTAR package. The alignment results showed that 87-126, 166-199, 421-462, 421-462, 666-622 and 668-709 were detected as the conserved regions of this gene and the 557-572 residues represented highly conserved region. This showed that the conservation of mentioned regions is very important for enzyme activity With regard to its evolution. Phylogenic analysis divided the 6 Poaceae species into 3 clusters. The Triticume Aestivume and Oryza sativa, a Sorghum bicolor cultivar Saccharum officinarum, and Zea mays and another Sorghum bicolor cultivar species were in the same cluster, respectively. The plants that were assigned to the same cluster are related morphologically.
Application of QTL Databases to Agriculture Biotechnology and Molecular Plant Breeding with Emphasis on Rice (Oryza sativa L)

Sabouri A., Sabouri H., Dadras A. R.

Department of Agronomy & Plant Breeding, Faculty of Agriculture, University of Guilan, Iran.

Gonbad High Education Center, Gonbad, Golestan, Iran. Islamic Azad University of Lahijan Unit, Guilan, Iran.

Nowadays, there is an enormous amount of scientific information available on plant genetics that certainly researchers don't enable to retrieve data quickly and efficiently without using databases. In agricultural biotechnology and the area of plant breeding, there are many databases that greatly facilitate the potential for comparative analysis among researches and lab based results with different plants. In present study, we used QTL databases and carried out analysis of detected QTLs of our own QTL mapping project and compared with other curated and submitted QTLs from different populations in databases. In this study, eight QTLs were detected for amylose content (AC) in rice (Oryza sativa L.) using a mapping population consisting of 236 F2:3 families derived from the cross between indica/indica rice varieties (Gharib with good eating quality as female parent and Sepidroud with poor eating quality). More analysis continued with QTL databases and results revealed one of these QTLs (qAC-6a) was in agreement with results of some other QTL analyses using different mapping populations. This QTL was mapped at the interval RM586-RM190 on the short arm of chromosome 6 using composite interval mapping and QTL Cartographer software in the studied population. To assess the value of qAC-6a by using of QTL databases especially Gramene database (http://www.gramene.org) showed that this genome region of rice is related to other 16 traits such as panicle length, potassium uptake, sodium concentration, in different mapping populations. Based on the results, we suggest this population could be investigated with respect to the relation between the region and these 16 traits. Obviously the identified QTL on specific chromosome region, explaining high phenotypic variance of some important traits could be considered for the use in marker assisted selection (MAS), fine mapping, map based cloning and QTL pyramiding programs.
Modeling the Structure of a Carrier Protein: Human Diferric Transferrin

Vahedian Movahed H., Saberi M. R., Chamani J.

Department of Biology, Faculty of Sciences, Islamic Azad University of Mashhad, Mashhad, Iran.

Availability of the three dimensional structure of carrier proteins is an important issue in the study of protein-drug interactions, which is one of the important research areas of drug administrations. Human diferric transferrin is one of these proteins that has an important task in binding and transporting intrinsic and extrinsic ligands such as drugs, vitamins and minerals. Despite the iron free form of transferrin, apo form, which has been resolved and is available in the Protein Data Bank (PDB), up until now, the holo form of this protein has not been resolved. So, the determination of the structure of holo-transferrin is of great importance. To do achieve this, the amino acid sequence of the protein was retrieved from UniProt. Then, a similarity search using, Psi-Blast, was carried out in order to find proteins with resolved structure and high identity as candidate templates. In the first stage, the N-lobe of the human holo transferrin was chosen as template and accordingly the the N-lobe of the target protein was modeled by modeler 9v7. The resultant structure representing an initial template and accompanied by the rabbit holo-transferrin as the second template were used for modeling the remaining C-lobe of the target protein. Moreover, the protein’s model were tested by different softwares, namely ERRAT, VERIFY 3D, WHAT IF and the Ramachandran plots. The protein models were refined and accordingly tested again until the best structure was achieved. Currently, we are using the obtained structure for docking with our drugs which in together with our experimental data will be presented soon.
A Novel View of the Interaction between Cyclophosphamide and Human Serum Albumin: A Molecular Modeling Approach

Akbari Tusi Sh H., Saberi M. R., Chamani J.

Medicinal Chemistry Department, School of Pharmacy, Mashhad University of Medical Sciences, Mashhad, Iran.

Department of Biology, Faculty of Sciences, Islamic Azad University-Mashhad Branch, Mashhad, Iran.

The experimental observations were followed up with docking studies where cyclophosphamide was docked to HSA to determine the preferred binding site on the protein. The crystal structure of HSA taken from the Protein Data Bank (entry PDB code 1AO6) was used to find the binding site of cyclophosphamide. Crystal structure of HSA is a heart-shape helical monomer of 66 kDa composed of three similar homologous domains named I (residues 1-195), II (196-383), III (384-585) and each domain include two sub-domain called A and B to form a cylinder and each sub-domains involves 6 and 4 α-helix, respectively. The only Trp residue (Trp214) located in sub-domain IIA. Almost all hydrophobic amide acids are embedded in the cylinder and role on absorption, metabolism, and transportation of biomolecule. Through molecular modeling by Autodock 4.0, the optimum binding mode and site was displayed. It can be seen that the entrance of site I, a small cavity with positive charge residues formed by interaction between sub-domain IIA and IB is most possible binding site. In the interaction between cyclophosphamide and HSA the binding energy obtain -3.34 kcal/mol and ligand efficiency is -0.24 kcal/mol. Otherwise the association constant is 3.55 and association constant unit is mM that indicate the binding of cyclophosphamide to HSA is not strong. The binding energy is -4.32 kcal/mol and electrostatic energy is -0.1 kcal/mol. Torsional energy is 1.37 and there is no hydrogen bond in this interaction. The results indicated that the formation of hydrogen bond decreased the hydrophilicity and increased the hydrophobicity to stability the cyclophosphamide-HSA system.
Codon Usage Optimization of cry1Ab Gene to Enhance Expression Brassica Napus L. Plant Using Bioinformatics Analysis

Sheykhhassan M. 1,2, Motallebi R. 1,3, Rahnama H. 1, Kazemitabar S. K. 2

1- Agriculture Biotechnology Research Institute of Iran (ABRII)
2- Sari Agricultural Sciences and Natural Resources University (SANRU)
3- Agriculture college of Zabol University

Codon usage phenomenon is a problem in gene cloning process, which may usually caused inappropriate expression of one specific organism’s genes in the host organism. Therefore, to overcome this problem, it is possible to optimize the codon gene target based on the common codon of the host organism. In this study, cry1Ab gene codons were optimized to enhance expression level of this gene in the Brassica napus L. plant by bioinformatics analysis. To obtain this goal, the target gene sequence and brassica napus genome sequence as host plant were obtained from the NCBI database. Then, the open reading frame (ORF) of each sequence recognized by the Vector NTI software was distinguished and the numbers of amino acids resulting from this gene were counted by EDITSEQ. Next, the percentage of each codon of the cry1Ab gene to the percentage of same codon in the Brassica napus L. genome was closed by bioinformatics analysis. To enhance expression and stability percentage of the cry1Ab gene in the Brassica napus L. also caused pest resistance in this plant.
A Bioinformatics Study of Human Glutathione S-Transferases

Zamansani F. 1, Ebrahime E. 1, Tale A. M. 2, Ebrahimie M. 4

1 Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran

2 Department of Crop Production & Plant Breeding, International Division, Shiraz University, Shiraz, Iran

4 Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran

Glutathione S-transferases (GTS) are multifunctional proteins involved in diverse intracellular events, such as primary, secondary, signaling and stress metabolism. In addition, GSTs are important detoxification enzymes. In this investigation, 35 human GST protein sequences were extracted and categorized according to their subcellular location in two groups. The first group included the cytoplasmic GSTs and the second one was the membrane single-pass type II membrane protein. No conservation region was observed when 35 GSTs were aligned together, however, within cytoplasmic and membrane groups many high conservation regions were determined. For each protein sequence, different features were extracted. In membrane GSTs, 3 amino acids had the highest frequency; leucine (11.5%) alanine(11.48%), and glycine(8.73%) (alanine and leucine) were hydrophobic and the other one (glycine) was hydrophilic. In the contrast, in cytoplasmic group, the frequencies of the two hydrophilic amino acids (lysine (8.04%) and glutamine(7.81%) and one hydrophobic one (leucine (12.87%)) were higher. These results clearly show that cytoplasm in GSTs are mostly hydrophilic while membrane GSTs are mostly hydrophobic. According to this classification, the length and the weight of these proteins were than statistically analyzed. Statistical feature comparisons showed that there was significant difference between membrane and cytoplasm is GSTs regarding length and weight (P=0.001). The length and weight of membrane GSTs were almost 3 times more than those of the cytoplasmic GSTs. In other words, the GSTs proteins lengths and widths were highly dependent on their location in the cell.
Codon Optimization of mtld in Order to Increase Its Expression Level in Brassica napus L. with Using Bioinformatics Analysis

Motallebi Chaleshtori R. 1,2, Siyahsar B. 1, Rahnana H. 2, Roodbar Shojaei T. 2

1 University of Zabol, Zabol, Iran.
2 Agricultural Biotechnology Research Institute of Iran, Karaj, Iran.

An important characteristic of exons is codon usage. Codon usage optimization basically involves altering the rare codons in the target gene; so that, they more closely reflect the codon usage of the host without modifying the amino acid sequence of the encoded protein. It has predicted that optimal codons will help to achieve faster translation rates and high accuracy. In this study we try to optimize mtld gene codons to enhance gene expression level in Brassica napus L. (Canola). For this purpose the mtld gene sequence from Escherichia coli was obtained from Gene Bank and its major Open Reading Frame (ORF) was distinguished. The target gene sequence and plant genome sequence was obtained from NCBI database and open reading frame (ORF) of both mentioned sequences were separated. Next ORF sequences were analyzed by EDITSEQ software. Afterward, codons percentage of each amino acid in bacteria was matched to similar codons in plant. We are expected that optimization of mtld bacteria gene codons, causing plant abiotic stress resistance, will increase expression level of mtld gene in plant.
Application of Haplotype Block Partitioning in Genome-Wide Case-Control Association Study

Katanforoush A., Sadeghi M., Pezeshk H., Elahi E.

Department of Computer Sciences, Shahid Beheshti University, G.C., Tehran, Iran.

A Genome-wide Association Study (GWAS) entails evaluation of statistical tests on Millions of loci over the genome. The conventional way to extend performance of the study over whole genome is the technique of sliding window. Though several efforts have been made to improve efficiency of haplotype-based methods in GWAS, the application of haplotype blocks, as firm genomic regions for the association test has been fairly ignored. Given a block partitioning, we performed association tests over each haplotype block so that an association map for of the entire genome was achieved. To do this, we applied the method of hierarchical clustering on haplotypes of case and control samples and used the chi-square ($\chi^2$) test to identify possible association between trait and haplotype clusters in each block. Based on simulated haplotypes of the case and control samples, we assessed the results with respect to the block structure defined by various methods of haplotype block partitioning, different models of disease risk contribution and the effect of low marker density on mapping accuracy. In 80 percent of tests, the trait-associated locus was found exactly, while the rate of false discovery was 10 percent. The results indicate that block-based approaches outperform methods of sliding window, with better accuracy and less computation. Moreover, results demonstrate that in case a sparse distribution of markers is available, the use of the optimal blocks defined by global partitioning methods can obtain association maps with the same accuracy as that generated by a high density set of markers generates.
Promoter Analysis of AtMYC2 and AtMYB2 Genes: Cis-acting Regulatory Elements and Their Possible Roles in Controlling Gene Expression Levels

Aliakbari M., Shamloo Dashtpagerdi R., Razi H., Ebrahimie E.

Department of Plant Production and Breeding, Faculty of Agriculture, Shiraz University, Shiraz, Iran

In Arabidopsis, many abscisic acid (ABA) inducible genes contain a conserved ABA-responsive cis-acting element named ABRE (ABA responsive element; PyACGTGGC) in their promoter regions. It has been reported that following accumulation of endogenous ABA, the transcription of AtMYC2 and AtMYB2 increases. These transcription factors bind to cis-elements in the promoter region of a dehydration-responsive gene, RD22, causing it to . Previous studies have reported that the expression level of AtMYC2 is higher than AtMYB2 under drought stress conditions. In this study, in order to detect cis-regulatory elements in the promoter region and delineate their possible role in the expression levels of Atmyc2 under drought stress conditions, promoter analysis was performed using Bioinfarmatics tools. Several regulatory elements of different elicitors were found in promoter regions of both AtMYC2, and AtMYB2 explaining involvement of these transcription factors in many biological processes. In addition, our results revealed that the number of ABRE (ABA response element) in the AtMYC2 promoter were higher than those in AtMYB2. Moreover, the distribution of ABRE elements was different. Interestingly, some MYB binding site elements were identified in the AtMYC2 promoter. These findings suggest that the number of ABRE cis-regulatory elements and their locations in the promoter regions of AtMYC2 and AtMYB2 may play important role in regulation of their transcription levels through interaction with the ABA signaling pathway. In addition, we demonstrated that the expression of AtMYB2 has a positive regulatory effect on expression of AtMYC2 through binding to MYB binding sites in the promoter region of AtMYC2. This is the first report on comparative promoter regions of AtMYC2 and AtMYB2 with regard their expressions levels.
Microarray Data Analysis of Two Transcription Factors "ERF4" and "TFIIB" in Soybean (Glycine Max) Somatic Embryogenesis

Kashfi R., Ebrahimie E., Poorkhorshid Z.

Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran.

In plant tissue culture, development of embryo from somatic cells is called somatic embryogenesis. The stages of in globular, heart, and torpedo shapes are the main events of somatic embryogenesis. In plants, somatic embryos are a model system to study the basic aspects of embryogenesis and are also a tool for transformation. The goal of this study was to shed light on the role of transcription factors in somatic embryogenesis. Transcription factors (T.F.) are a group of proteins that attach to some specific sequences of DNA and control the genetic information transfer from DNA to mRNA. Many of transcription factors play basic roles in the development of multi-cellular organisms. In this research, 2 transcription factors involved in direct somatic embryogenesis in soybean, "ERF4" and "TFIIB", have been studied by microarray data analysis. In this experiment, expression modulation of two T.F. in three developmental stages (1. globular-stage embryo 2. heart-stage embryo 3. cotyledon-stage embryo) and in 10 different tissues (whole seed - young trifolate leaf - embryo proper - endosperm - endothelium - epidermis - hilum - inner integument - outer integument - suspensor) were statistically analyzed. The results indicated ERF4 and TFIIB were jointly expressed in Embryo proper and Endosperm tissues. In the next step, the involvement of ERF4 in the embryo proper and TFIIB in the ndosperm were separately indicated. ERF4 acts as a novel negative regulator of JA-responsive defense gene expression against necrotrophic fungal pathogens such as Fusarium oxysporum and antagonizes JA inhibition of root elongation.
The arsR Gene: A Comparative Phylogenetic and Bioinformatic Analysis to Study Reasons for Different Sensitivity to Arsenic Compounds

Daneshpour M. , Saeedi Jam M.

Medical Biotechnology, School of Medicine, Hamadan University of Medical sciences, Hamadan, Iran.

Arsenic is a known toxic metalloid, whose trivalent and pentavalent ions can inhibit many biochemical processes. The ars operons which encode. Arsenic resistance has been found in multicopy plasmids of both Gram -negative and gram positive bacteria. This operon contains three important genes consisting of arsR, arsB and arsC. Several researches based on molecular techniques showed obvious homology between this operon and special sequences of chromosomal DNA from a number of bacterial species. These results suggest that the chromosomal ars operon may be the evolutionary precursor of the plasmid-borne operon. Moreover, the experimental studies indicated that the ars operon from plasmid R773 is more sensitive to arsenic compounds compared to chromosomal ars operon. One of the probable factors for this phenomenon may be the arsR gene (the first cistron of the ars operon) and its product, ArsR protein, which is a trans-acting repressor that regulates expression of ars operon. The overall goal of our study was to carry out a set of comparative analyses of arsR gene and the ArsR protein in plasmid R773 and bacterial chromosome from Escherichia coli BL-21(DE3). The PDB and NCBI databases and Chimera, Mega4, and CLC main workbench softwares 3D-jigsaw and Prosite servers were mad use to in this study. By using these softwares and servers, multiple analyses including determination of residue composition, secondary structure and motifs, 3D structure, conserved regions, were carried out. Our results suggest that such high sensitivity to arsenic compounds in the ars-containing plasmid R773 may be due to the related ArsR protein characteristics such as amino acids composition, secondary and tertiary structure, hydrophobicity and level of interaction with DNA. In addition, our phylogenetic analyses suggest that there is an interesting possible evolutionary pathway from prokaryotes to eukaryotes based on the arsR gene.
Breast Carcinoma, Intratumoural Heterogeneity and Cancer Grading

Sharifi-Salamatian V.

World Health Organisation (WHO).

The aim of this work is to investigate intratumoral heterogeneity in breast cancer and its relationships with breast cancer grading. A novel methodological approach to measure heterogeneity is used here and an estimation of the three histological grades of Scarff, Bloom and Richardson are proposed. The geostatistical method used here lies upon notions of asymptotic behaviour, dispersion variance, ergodicity and integral range. After computing the integral range, the estimation of the convergence "speed" of the spatial average to the statistical mean of the spatial point process alpha is obtained. The more the process is heterogeneous, the slower the convergence of the dispersion variance to zero and the smaller the value of (alpha). Very recently, we realized that the use of the asymptotic slope of the Hurst (H) parameter (related to fractals) bears a relationship with dispersion variance and procures more robustness to the estimations. Twenty tumours were obtained and paraffin sections were stained by MIB-1 (Ki-67). The results were expressed with respect to alpha as an index of heterogeneity. There were significantly different grades for 1 and 2 than for 3. Similar discrimination was obtained with the same sections using the Hurst parameter (H). Thus heterogeneity, as measured by alpha or H, clearly bears a relationship with grading. We use an index of heterogeneity which might have better individual prognostic values than the current grading system. The final proof of an improved grading using these parameters will of course require a confrontation with the results of survival studies, which are currently in progress.
Comparison of Root Genes Expression Profiling of Wheat under Drought Stress by Expressed Sequence Tags (EST) Analysis

Ravash R. 1, Shiran B. 1, Ebrahimie E. 2, Shamloo Dashtpagerdi R. 2

1 Shahrekord university, College of Agriculture, Department of Crop Production & Plant Breeding, Shahrekord, Iran
2 Shiraz University, College of Agriculture, Department of Crop Production & Plant Breeding, Shiraz, Iran

Drought is one of the most important environmental constraints limiting plant growth and agricultural productivity. Water deficit can induce several genes and cause different changes in plant expression pattern. In this research, two EST libraries wheat root in normal and stress conditions were downloaded from Harvard University data bank (http://combio.dfci.harvard.edu/). Normal library contained 1024 ESTs, and 1306 ESTs were deposited in stress library. ESTs sequences from two libraries were clustered and assembled together using EGassembler, an online bioinformatics service (http://egassembler.hgc.jp/). All clusters (contigs and singletons) of these libraries were compared with Oryza sativa proteins by Blastx (E-value ≤ 1*10^-5) using CLC Protein Workbench software (www.clcbio.com). Assembling of 2330 ESTs of two wheat cDNA libraries resulted in 184 high represented genes (contigs) that each of them contained ESTs from drought stress and unstressed libraries. Differentially expressed genes between drought and normal libraries were found; 69 up-regulated genes and 56 down-regulated. Accessions like (NP_001059473, NP_001060897, NP_001053892, EAZ17794, EAY94718, EAY80813) had significantly over expression under stress condition and refer to hypothetical proteins, genes encoding late embryogenesis abundant protein, Alpha-1 tubulin and S-adenosylmethionine decarboxylase, had significantly up-regulated too. Cases that showed significant down-regulation under drought stress are exemplified by Genes encoding Cytoplasmic and cytosolic Malate dehydrogenases, interleukin-1 beta converting enzyme and glycine-rich protein.
Phylogenetic Analysis of the Enterobacteriaceae Family Based on the Nucleotide Sequences of rpsL Gene

Dehghani A. ¹, Daneshpour M. ²

¹Genetic, Bioscience school, Shahid Beheshti University, Tehran, Iran
²Medical Biotechnology, Medical College, Hamedan University of Medical sciences, Hamadan, Iran.

The S12 is a ribosomal bacterial protein coded by the rpsL gene in E.coli. The function of the S12 protein is to ensure that the mRNA properly aligns with the ribosome prior to the start of translation. Streptomycin, an aminoglycoside, inhibits protein synthesis in E.coli by binding specifically to the S12 protein. A missense mutation in the rpsL gene results in the insertion of a different amino acid in the synthesis of a polypeptide chain, and has the potential to render the S12 protein nonfunctional or alter the structure of a functioning S12 protein. Reptomycin resistant and dependent mutants are the consequence of a missense mutation in the rpsL gene, such that a structurally altered S12 protein is still functional. Phylogenetic analysis is widely used to determine the evolution and function of genes, to identify the uses of bacteria based on gene sequences. The use of protein-coding genes that are known to evolve much faster than rRNAs seems to be more appropriate for the phylogenetic analysis of closely related bacteria. The enteric bacteria belonging to the Enterobacteriaceae family are of special microbiological interest because of their pathogenic and non-pathogenic relationships with the human gastrointestinal tract.

Since it is expected that the rpsL gene has significant conserved regions, in this study, we carried out a phylogenetic analysis of Enterobacteriaceae based upon the rpsL gene. In addition, our results and similar analysis of Enterobacteriaceae based upon gyrB and 16s rRNA genes were compared. We used NCBI databases, Mega4 and CLC main workbench softwares. Evolutionary trees based on Nucleotide sequences and translated amino acid sequences of rpsL were constructed. Consequently 16S rRNA gene sequences were useful in describing phylogenetic relationships between distantly related Enterobacteriaceae, whereas gyrB and rpsL sequence comparisons were useful for inferring intra- and some intergeneric relationships.
A Report about Choosing the Appropriate Gene Source for Cloning Based on Bioinformatics Studies

Rikhtegaran Tehrani Z. , Daneshpour M. , Saeedi Jam M.

Department of Genetic & Molecular Biology, University of Medical Sciences, Hamadan, Iran

Streptavidin is a protein that binds to the vitamin, D-biotin (vitamin H), with remarkably high affinity constants. This unique feature is the basis of avidin-biotin technology, which has evolved into a universal tool in various fields of the biological sciences. Streptavidin cloning, as a good way to produce this valuable protein, is usually performed using treptomyces avidinii as a gene source. In order to use other microorganisms containing the streptavidin gene instead of Streptomyces avidinii, the most suitable initially involved bioinformatics studies. To achieve this aim, as a first step, we searched the NCBI database to find out microorganisms containing the streptavidin gene. Then, they were aligned and their phylogenetic trees were reconstructed by the Mega4 software. It was observed that there was a higher relationship between certain organisms and avidinii with respect to the streptavidin gene. Then, we determined the conserved motif in the sequences using the EXPASY server and acquired definite results, which indicated homology between streptavidin of some organisms and avidinii. We then tried to compare each sequence with the target gene by basic local alignment tool at NCBI, and realized that there are two organisms containing genes very similar to the streptavidin of avidinii. Next we obtained their 3-dimensional structures using the by 3JIGSAW server and compared them with that of avidinii. Finally, the ClC main workbench software was implemented in order to predict physical features of the selected proteins, comparing then with streptavidin of avidinii. Following and as a result of these bioinformatics studies, we could increase the probability of achieving more appropriate results in the laboratory and thus save a lot of time.
A Study on Different Analysis Methods of Real-Time PCR Data

Tahmasebi A., Ebrahimie M., Ebrahimie E., Mohammadi-Dehcheshmeh M.

Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran
Plant Research Centre, School of Agriculture Food and Wine, University of Adelaide, Australia
Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran

Gene expression analysis is increasingly important in many fields of biological research. Real-time PCR is a powerful tool to quantify gene expression, because of its high sensitivity, good reproducibility and quantification range. The output of the Real-time PCR is summarized in a single value called the threshold cycle (CT). Generally, two different methods of analyzing data from real-time, quantitative PCR experiments exist: absolute quantification and relative quantification. An absolute quantification provides the exact copy number following transformation of the data via a standard curve. A relative quantification is based on the relative expression of a targeting gene versus a reference gene. This paper compares the differences, superiorities, and applications of these two methods.
Introduction of Empirical Bayes Approach, as a Novel Reliable Method in Microarray Data Analysis

Baseri S., Ebrahimie E., Ebrahimie M.

Department of Statistics University of Shiraz, Shiraz, Iran

Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran.

Department of Crop Production & Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran.

Microarray is a novel technology facilitating simultaneous measurement of thousands of gene expression levels. This technology has helped gene expression investigations to progress significantly. One of the major problems, which scientists have to deal with, is to find an appropriate and reliable statistical way to determine the significantly expressed gene during the appearance of the phenomenon. Three main approaches, including classic, Bayes and empirical Bayes have been compared in this paper. With respect to limited numbers of replications (observations) and large numbers of variables (genes), the classical statistical methods are not able to analyze microarray data precisely. Moreover, Bayes method needs prior distribution data, which are not known in most cases; hence, it is not possible to use Bayes methods. In contrast, empirical Bayes is a practical strategy for estimation and inference. Empirical Bayes can provide practical strategy for selection of effective genes during occurrence of a phenomenon, the most serious problem in microarray analysis, since this approach can work with a smaller number of observations (replications).
Designing a Bioinformatic-Based System for Identification of Self-Incompatibility Alleles in the Pyrus Species

Babaei F.¹, Abdollahi H.², Amin Afshar M.

¹Science and Research Unit, Azad University of Tehran, Iran
²Seed & Plant Improvement Institute, Tehran, Iran

Self incompatibility in some conditions reduce fruit set and yield the Pyrus species. Selection of compatible pollinators are important both to overcome self-incompatibility as well as inter-cultivars incompatibility. Recently more than 30 S-RNase alleles recognized in the Pyrus species were needed to complete the categorization and identification system. This research was aimed at homology analysis, primer designing and exact identification of S-RNase alleles in the three Pyrus species, P. communis L., P. pyrifolia and P. ussuriensis. All sequenced and identified pear s-alleles were aligned (clustalW2 software) and cluster were analyzed with to a 172 bp conserved region (NTSYS software), selected for degenerating primer design (OLIGO software) able to anneal with all s-alleles. The primer pair amplified a 108 bp region in all identified s-alleles. By using restriction enzymes a unique restriction map analysis was obtained for the entire alleles. 32 restriction enzymes were able to cleave this conserved region in all identified alleles. 10 enzymes out of 32 were not able to identify polymorphism in the s-alleles. 9 alleles out of 39 were identifiable with just one restriction enzyme and 5 alleles with two restriction enzymes. A multi step digestion system was designed to identify a total 19 S-RNase alleles in the Pyrus species.
The Study on Conserved Regions of Pyrroline-5-Carboxylate Reductase Gene (p5CR) in the Important Plant Species

Motamedi M. J., Zebarjadi A., Kahrizi D. 1

1Biotechnology for Drought Tolerance Department, Razi University, Kermanshah, Iran

Proline accumulation is one of the major responses to water stress in most plants. It has been suggested that the accumulated proline plays a role as an osmoregulator and an osmoprotectant. The pyrroline-5-carboxylate reductase catalyzes the final step in proline biosynthesis. Moreover, it is interesting to note that increased p5CR activities and gene expression have been shown in a number of plants in which proline accumulation participates in the process of osmotic adjustment. The stimulation of proline synthesis in stressed plants is correlated with increased p5CR activity. In this research, we p5CR gene sequences of 23 important plant species including Brassicaceae, Poaceae and Fabaceae from the Gene Bank database. These sequences were then analyzed using the DNASTAR package. The alignment results showed that the aminoacids Gly 63, Arg 71, Lys 233, Val 258, Gly 263, Leu 273, Leu 427, Pro 459, Pro 581 and Val 612 were highly conserved residues. Furthermore, the regions 337-351, 431-434 and 481-493 had the lowest variations in amino acids and were identified as relatively conserved regions. In this regards, it can be said that above residues and regions are very important for structure and function stability of this enzyme and have become conserved as a result of the evolution process. According to phylogenetic tree results, the mentioned 23 species were divided into 4 clusters. The species in the same cluster confirmed the closely genetic relationship of species within a family. In addition, the species numbered 14, 15 and 16 had showed the most conserved sequences among the 23 species that belonged to Fabaceae family.
Positive Selection in Molecular Evolution of the HSP70 Gene Family

Ahmadpanah J., Moradi M., Masoudi A., Vaez Torshizi R.

Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran

Heat shock proteins (HSPs) are present in all cells of the body and function as molecular chaperons to protect the other proteins from degradation and malfunction. HSPs are distributed across the genome and divided in to various groups according to the molecular weights. One of the members of HSPs, HSP70, shows differentiated expression against unusual environmental changes such as stress, feed restriction and high temperature. Therefore, the genes encoding these proteins have been identified as candidate genes in view of animal science. In animals, occurrence of some SNPs in the structure of these genes was associated with resistance to mastitis and infectious bursal disease, conservation of homeostasis, adaptation to hot regions and increasing the immunes systems. Current study was conducted to identify nucleotide substitutions in the evolutionary path of the HSP70 gene family in cattle and comparing to other animal species. Genomic sequences of the HSP70 gene family were identified by in silico screening of the animals' genomic information at NCBI. Multi- alignment analysis, estimation of the pattern of nucleotide substitution and selection analysis of the sequences were carried out using GENETYX and MEGA4 softwares. The overall frequencies of pyrimidine and purine nucleotides in HSP70 gene family were 0.466 and 0.534, respectively. However the highest substitution rate occurred between thymine and cytosine bases (16.63), but this rate was 1.645 for purines. Furthermore, evidence from the comparisons of the nonsynonymous to synonymous substitutions and distribution of the mutation sites suggested that HSP70 gene family evolved by positive selection. In conclusion, it can be suggested that positive selection plays an important role in the evolution of this gene family which maybe interpreted as development of various biological functions of these proteins during evolution.
Bioinformatics Analysis of Sequence tag Sites (STSs) in SMA Locus

Sedghi M., Vallian S.

Department of Biology, University of Isfahan, Isfahan, Iran.

Spinal muscular atrophy (SMA) is a common autosomal recessive disorder characterized by degeneration of alpha-motor neuron in the spinal cord. The SMA determining gene, termed survival motor neuron1 (SMN1) is present on 5q13 within a region characterized by a large 500 kb inverted duplication containing several repeated genes. In view of the complexity of the SMA locus, analysis of genetic markers in this locus could provide valuable tools for carrier detection and prenatal diagnosis of the SMA disease. A forward electronic PCR was performed on the genomic sequence of the SMA locus to investigate sequence tag sites (STSs). A large number of STSs was detected. However, those STSs that were mapped to only one location and were close to the SMN1 gene were selected. Three of these STSs including D5S1414, D5S351 and D5S1408 were used in the laboratory to determine whether these markers were informative in the Iranian population.
Bioinformatics Study of Δ 12-Desaturase Gene in Six Plant Species

Mohammadi Z. ¹, Kahrizi D. ²

¹ Agronomy and Plant Breeding Department, Ilam University, Ilam, Iran
² Biotechnology for Drought Tolerance Department, Razi University, Kermanshah, Iran

Fatty acids are carboxylic acids with long hydrocarbon chains that play an important role in biological processes. They are one of the sources for energy and form a part of the cell membrane. The enzymes that transform the saturated of fatty acids to the unsaturated form are called desaturases, which are complex enzymes. The Δ 12-Desaturase enzyme desaturates the binary bond in the 12th position of the carboxyl group. This action causes promotion of fatty acid quality. In this research, alignment of Δ 12-Desaturase gene in six plant species was carried out with bioinformatics software. The purpose of this study was the phylogenetic investigation of species based on mentioned the gene. Alignment results showed that five species including Arabidopsis thaliana, Eruca sativa, Brassica juncea and Lepidium virginicum were assigned to the same cluster. These results confirm the close genetic relationship between these species. Finally, the Linum usitatissimum species was sited in a separate cluster that indicates the high genetic distance between this species and others.
Acanivorax Borkumensis Whole Genome Analysis for P450 Like Proteins

Nouri A. 1, Minuchehr Z. 2, Madadkar-Sobhani A. 3, Bambaei B. 2

1Chemistry Engineering Department University of Tehran, Tehran, Iran
2National Institute for Genetic Engineering and Biotechnology, Tehran, I.R. Iran.
3Chemistry Engineering Department University of Tehran, Tehran, Iran

Oil pollution is caused by natural phenomena or human activities like oil tanker disasters or under water pipeline interruptions. It is estimated than annually more that 5 million tons of crude or refined oil enters the marine ecosystems. Currently, there is no efficient chemical or physical method for removing these pollutions. Bioremediation, on the other hand, is one the most friendly techniques in removal of oil pollution. Acanivorax borkumensis, a hydrocarbon degrading bacterium has the strongest role among the microbial systems in bioremediation of marine environments. To catabolize released oil, this bacterium is equipped with different enzyme systems for hydroxylating linear and cyclic hydrocarbons. Cytochrome P450s, are one of the most versatile enzymes in oxidation of hydrocarbons. In the sequenced genome there are 3 annotated P450s. Here we searched the genome for possible new P450s that are not recognized in the first draft of the genome, to accomplish this task we developed new approaches for scanning the genome for recognizing new P450 or P450-like sequences. Our results revealed that there are at least three other ORFs in the genome of A. borkumensis that might demonstrate P450 activity. Our next step would be to express theses ORFs to assess their activity and their role in hydrocarbon degradation. These putative new enzymes will have huge potentials in biotransformation or biocatalytic reactions in industrial biotechnology.
System Genetics Approach for Analysing Rat Muscle Expression Data

Ghaderi-Zefrehei M.

Yasouj University, Yasouj, Iran

Ensembles of different methods (expression QTL mapping and weighted co-expression networking) were used to get insight into rat’s hindlimb muscle gene expression data. The data included three experiments. Extra source of information was used to mediate eQTL toward system genetic analysis. We identified following genes: M54987, M59742, X01454, M17523, X03639, U62779, M59786, X51707, U37026, D90219, M35862 as hub acting genes and sought to be associated with rat hind limb muscle. Increasing muscle loading duration; gave arisen that a hub acting gene M35862 to be identified. Expression QTL mapping for interested in genes ended up with Trans-acting regulatory region, addressing complex regulation for aforementioned dataset. This study showed that differently expressed gene is slightly likely to act as hub genes across dataset. This study indicated that soundly using different sources of information most likely can be insightful in system genetics (system biology + genetics) analysis or in post-era genomic paradigm, especially when the richness of genomic data is at premium.
Molecular Characterization of the Iranian Isolate of Barley Dwarf Virus Rep Protein and Its Similarity with Other Mastreviruses

Amid Motlagh M. H., Behjatnia S. A. A., Afsharifar A. R., Izadpanah K.

Department of Plant Protection, College of Agriculture, Shiraz University, Shiraz, Iran

The complete genome of an Iranian isolate of Barley dwarf virus (BDV-Ir) consisted of 2733 nucleotides. It encodes four proteins: The movement protein and the coat protein on the virion-sense strand and the replication associated proteins (RepA and Rep) on the complementary-sense strand. Rep is the only protein required for geminivirus DNA replication and catalyzes multiple reactions during the reproductive cycle of the virus in a rolling circle replication mechanism model. Phylogenetic analyses conducted using MEGA4, Vector NTI Advance v.11 and ClustalX programs indicated that BDV isolates share 89.8-99.6% sequence identity among themselves. The Rep gene of BDV-Ir is composed of 1044 nucleotides coding for 347 amino acids. An alignment of Rep protein sequences of BDV-Ir and other BDV and Wheat dwarf virus (WDV) sequences available in GenBank indicated a mean pairwise percentage sequence identity of 92.8%-96.3%. Analysis of amino acid sequences indicated that the Rep protein was highly conserved between the BDV and the WDV, except for a stretch of twelve nucleotides in the C-terminus, where BDV isolates lacked four corresponding amino acids while completely conserved in the wheat isolates. This informative insertion/deletion of four amino acids allows for a qualitative differentiation between WDV and BDV isolates. Phylogenetic analysis based on Rep nucleotide and amino acid sequences showed that BDV-Ir was completely distinct from other BDV isolates. This may indicate intergenic recombination in these viruses.
Potential of MS Excel in Developing Bioinformatics Softwares

Kadkhodai S., Khayyam Nikoii M., Ghaderi M.

Department of Agriculture, Yasouj University, Iran

Because of increasingly generation of Omics information and data in daily base from one side and immediate necessary of analysis and processing Omics data, developing and using software helping users, is becoming very crucial. In investigation of large scale genetics variations, in general, we are dealing with manipulation of large datasets and many parameters inflicted from it, these both shall increase the probability of committing different errors. Nowadays, applying different computer softwares has brought about of rapid and accurate of genomics data analysis. However, singling out software which will be able to do all kind of analysis in genomics data requires lot of money or is rarely available. As such, the researcher is to use different computer software to get his research objectives. Current software which is developed and programmed in Microsoft Excel 2007 is quite simple to be used in research investigation. In current version of ABRISTAT29.2 which many Excel computation features have used e.g. mathematical functions, search task, conditional logic search, contains the following sections: A: the section of importing of raw data which include the size of genotypic bands in different genotypes of under investigation. B: The section of 0 and 1, which converts data in the aforementioned section (A) into 0 and 1. C: the section of Allelic Genotype, which make shows the allele of each genotype as English Alphabet. This accomplishment is mainly used in co-dominant genetic markers (SSRs). D: the section of Alleles, which make shows alleles size for each genotype. E: The section of parameter estimation, which some parameters like: \( H_0, H_e \) (\( \text{PIC}=1-\sum f_i^2 \)), Discrimination power (\( 1-\sum p_i^2 \)), Index Genotypes, etc. are calculated using aforementioned sections.
Genetic Diseases and Bioinformatics

Esmaeilzadeh A., Keshtkar S.

Department of microbiology faculty of biological science, shahid Beheshti University, Iran

One of the main goals of systems biology and modern genetics is the identification of pathways and genes underlying complex traits such as common human diseases. Many of these traits involve both genetic and environmental factors. Recently, there has been some effort to integrate DNA variation, expression data, and genotypic data as well as analysis on tissue-to-tissue co-expression data to enhance identification of the associations between DNA variation and disease and characterize those parts of the molecular networks that drive disease. Other studies focus on new methodologies on different datasets to infer genetic factors behind these complex traits. Another area of research focuses on looking at the disease at a broad spectrum to highlight genetic roots and pathways behind different diseases and the way these might be linked. Here we review different approaches for integrating expression quantitative trait loci (eQTLs), expression, and clinical data to infer causal relationships among gene expression traits and between expression and disease traits. We further review methods to integrate these data in a more comprehensive manner by constructing co-expression gene networks that utilise pair-wise gene interaction data to represent more general relationships. A study which utilizes an integrative genomics approach for identification of genes associated with a particular disease is. This describes a multistep process for extraction of causal information from gene-expression data related to a complex trait. The case study is obesity on a segregating mouse population. The dataset used is the BXD dataset of a mouse population and associated liver gene-expression data which has been also used by other similar genomic studies. The main focus and contribution of this work is to identify genes that cause a particular disease rather than the ones which respond to a disease state.
Structural Studies of Therapeutic Monoclonal Antibodies Using IMGT/3D Structure; Tools and Database

Akbarzadeh Sharbaf S. 1, Zeinali S. 2, Minuchehr Z. 1, Shokrgozar M A 2, Yakhchali B. 1

1National Institute for Genetic Engineering and Biotechnology, Tehran, I.R. Iran.
2Pasteur Institute of Iran.

Nowadays, there is an increase pressure to develop of new antibodies with increased specificity and efficacy. IMGT-3D structures-DB is a novel 3D immunological proteins structure database and has some powerful tools for variable domain analysis. The defined structure of the fragment variable (Fv) module of four recombinant mAbs approved for cancer therapy namely Trastuzumab (1N8Z), Pertuzumab (1S78), Bevacizumab (1B1) and Cetuximab (1YY9), and three new developed mAbs namely Dual specific bH1 (3BE1, 3BDY), Matuzumab(3C09) and 2H10(2VWE) are presented and analyzed with IMGT/3D structure tools and database. Amino acids of rmAbs involved in hydrogen bonds with the antigens in light and heavy chains are obtained from IMGT/3Dstructure-DB, http://imgt.cines.fr, and compared. The genetic origin of these recombinant monoclonal antibodies, determined through the IMGT/3Dstructure-DB database and IMGT/V-QUEST. Moreover IMGT-3Dstructures-DB provides 2D aphical representations (Collier de Perles) and result of contact analysis. IMGT/Domain Superimpose tool is used for structural alignment and superimpose two IMGT domain 3D structures. The similarity and difference between these structures are discussed.
The Nano-Molecular Modeling of the Interaction between Ropinirole Hydrochloride and Human Holo-Transferrin Based on Lamarckian Genetic Algorithm

Kabiri M.¹, Bozorgmehr M. R. ², Chamani J.¹

¹Department of Biology, Faculty of Sciences, Islamic Azad University-Mashhad Branch, Mashhad, Iran

²Department of Chemistry, Faculty of Science, Islamic Azad University-Mashhad Branch, Mashhad, Iran

The application of molecular modeling has been employed to study the interaction between Human holo-transferrin (hTf) and Ropinirole hydrochloride. The hTf is divided into two evolutionary related lobes, designated the N-lobe (336 amino acids) and C-lobes (343 amino acids), which are linked by a short spacer sequence. Each lobe contains two domains comprising a series of α-helixes, which overlay a central β-sheet backbone. The domains interact to form a deep, hydrophilic metal ion-binding site. The binding site in both the N and C terminal lobes has four conserved amino acids including two tyrosines, one aspartic acid and one histidin (N-terminal lobe- Asp-63, Tyr-95, Tyr-188 and His-249). The potential of the 3-D structure of hTf was assigned according to the Amber 4.0 force field with Kollman-all-atom charges. The initial structure of the Ropinirole hydrochloride was generated by molecular modeling software Vegga zz. The geometrics of the molecules were subsequently optimized to minimal energy using the Tripos force field with Gasteiger-Marsili charges. At last, Autodock 4 program was used to establish the interaction modes between Ropinirole hydrochloride and hTf. Lamarckian Genetic Algorithm (LGA) program was used to calculate conformational possibility between Ropinirole hydrochloride and hTf and exhibit the optimal energy ranked result of Ropinirole hydrochloride interaction with the residues of hTf and the Ropinirole hydrochloride-hTf conformation. Protein-drug docking was performed and the binding constant of Human holo-transferrin with Ropinirole hydrochloride for 9 complexes with at least ΔG and high affinity were calculated. The region with highest affinity binding was located between N-lobe and C-lobe cleft in near of α-helix position.
A Novel View of the Interaction between Tamoxifen and Human Holo-Transferrin Modeled: A Molecular Modeling Approach

Sarzehi S.¹, Bozorgmehr M. R.², Chamani J.¹

¹ Department of Biology, Faculty of Sciences, Islamic Azad University-Mashhad Branch, Mashhad, Iran (Email: Sareh_sarzehi@yahoo.com) Tel: +98511 6040642.

² Department of Chemistry, Faculty of Sciences, Islamic Azad University-Mashhad Branch, Mashhad, Iran

The application of molecular modeling has been employed to study the interaction between Tamoxifen (TMX) and human serum Transferrin (hTf). The potential of the 3-D structure of hTf was assigned according to the Amber 4.0 force field with Kollman-all-atom charges. The initial structure of the TMX was generated by molecular modeling software Vega zz. The geometrics of the molecules were subsequently optimized to minimal energy using the Tripos force field with Gasteiger-Marsili charges. At last, Autodock 4 program was used to establish the interaction modes between TMX and hTf. Lamarckian genetic algorithm (LGA) program was used to calculate conformational possibility between TMX and hTf and exhibit the optimal energy ranked result of TMX interaction with the residues of hTf and the TMX-hTf conformation. Transferrin is a large non-hem iron-binding glycoprotein that consists of about 679 amino acid residues. Crystal structure analysis has revealed that hTf comprises of two homologous lobes: (C and N) that first half of the polypeptide chain (residues 1 to 330) as N-lobe, and the terminal half of the chain (residues 350 to 679) as the C-lobe, each lobe comprising of two similarly sized domains. Tamoxifen belongs to the chemical class of riphenylethylenes (M⁰ = 371.5). Docking program obtain the binding constant of TMX with hTf for nine complex with minimum ∆G° (High affinity) that position with high affinity is within region between two cleft of N and C lobes close to α-helix location. Variable radius of ∆G° report medium inhibition constant for different complexes that is partly correspondent for Tamoxifen drug. These results may be of significance in pharmacology and clinical medicine.
Primer Designing and Restriction Enzyme Detection for Exon 14 of ABCG2 Gene in Iranian Holstein Bulls

Mousavizadeh A. ¹, Salehi A. ¹, Amin Afshar M. ²

¹ University of Tehran, Aboureyhan Campus, Animal Science Department
² Faculty of Agriculture & Natural Resource, Science & Research Branch, Islamic Azad University

ABCG2 (ATP binding cassette subfamily G member 2) gene have mapped in chromosome 6 which is encoded ABCG2 protein that transports various xenobiotics, cytostatic drugs across the plasma membrane and cholesterol into milk. A single nucleotide change (A/C) in base 86 of exon 14 is capable of encoding a substitution of Tyrosine to Serine (Y581S) in the ABCG2 gene and increases milk yield and decreases fat and protein concentration. The aim of this research was to design primer for ABCG2 gene and detect restriction enzyme for Mutant loci of the gene in Iranian Holstein bulls. Genomic DNA was extracted from semen samples using high pure PCR template preparation kit. Primers were designed with OLIGO software and checked in NCBI. Designed primers were utilized in PCR and after amplification, the PCR products were sequenced. Melting point for both primers was about 64 °C and no hairpin was observed. The amplified fragments length were equal to 240 base pairs. The amplified fragment include the end of intron 13 (Base number 4105 to base number 4141), entire of exon 14 and initial parts of intron 14 (68 base) of ABCG2 gene. The sequence of forward primer was 5’-GTATTCACGAGACTGTCAGGG-3’ and it was 5’-GGCTTTATTCTGGCTGTTTCC-3’ for the reverse. We want to use PBR technique for genotyping the animals for the SNP 86 of exon14 of ABCG2 gene, Therefore the StyLT1 is a restriction enzyme that was found for A/C mutation in exon14 of ABCG2 gene by MAPDRAW software for the first time. Recognition site of StyLT1 is 5’-CAGAG-3’, 3’-GTCTC-5’ according to mutant loci of ABCG2 gene.
Comparison of Conserve 16s Ribosomal RNA Sequence for Designing Universal Primer to Detect Obligate Anaerobic Bacteria

Khorasani M. ¹, Fallah Mehrabadi J. ²

¹ Islamic Azad University, Branch of Zanjan, Zanjan, Iran
² MARS Bioinformatics Institute, Tehran, Iran

Identification of bacterial population from unknown environmental samples is difficult and time consuming work. Culturing and isolating obligate anaerobic bacteria, some of which may be pathogenic to humans, need special conditions that some times can be problematic. For identification of microbial populations, detection of all bacteria is essential. Molecular methods, an alternative to culturing procedures are essential and have more accuracy and are less time consuming. Design of universal primer based on conserved 16s ribosomal sequences for polymerase chain reaction (PCR) by Bioinformatics tools is highly valuable in applied research.

M AND METHODS: Comparison of conserved 16s ribosomal RNA and design of a universal primer for obligate anaerobic bacteria were aim of this study. At first, all species of obligate bacteria were collected and then the sequence of the 16s ribosomal RNA sequence longer than 1200bp was obtained from NCBI and using the CLUSTALW2 software alignment were carried out then based on conservative regions of the 16s ribosomal RNA. Universal primer is designed by gene runner.

RESULTS: Obligate anaerobic bacterial genera consisting of different species, some of which were pathogenic to humans and others are including obligate anaerobic photosynthesing, methanogenic, sulfide reducing bacteria were identified by comparison of conservative 16s ribosomal RNA gene sequences at the species level.

CONCLUSION: The design of universal primers for conservative regions of the 16s ribosomal RNA in obligate bacteria assists with identification of these bacteria at the species level by polymerase chain reaction (PCR) of unknown samples.
The Use of Bioinformatics in DNA Fingerprinting of Pistachio Cultivars

Tagizad A. ¹, Ahmadi J. ², Haddad R.², Zarrabi M. ².

¹ Imam Khomeini International University
² Department of Agricultural Biotechnology, Imam Khomeini International University

Pistachio is one of the most important horticultural product of Iran. The best way to get the maximum yields is related to have genetically pure and monotonous gardens. Thereby, study of genetic variation and providing genetic identifications, make this possible to have homogenous gardens with high performance genotypes. In the present study, genetic relationships and diversity among nineteen cultivars (five cultivars of them in separate samples both five male and five female samples were supplied) were assessed using inter-simple sequence repeats (ISSR) primers. The survey was done by 21 ISSR primers in agarose gel, through them 10 primers with clear polymorphic bands selected for genotyping. using 10 ISSR primers at all 114 loci were produced that 73 loci of them (64.03%) were polymorphic. The mean of polymorphism information content (PIC) for the primers was high and was ranged from 85%(minimum) to 91%(maximum). The genetic similarity matrices were constructed using Jaccard coefficient. Clustering dendrogram were constructed by unweighted pair group method using arithmetic average (UPGMA) method. The efficiency of clustering algorithms and their goodness of fit were determined based on the cophenetic correlation coefficient. Cluster analysis revealed three main groups with two, five and 17 pistachio varieties in each group. The present study showed that use of ISSR markers for investigation on genetic diversity of different pistachio cultivars could be useful and informatic.
Analysis of Alternative Splicing Mediated by the PASA Software

Valiellahi E. 1, Kavousi H. R. 2, Mirshamsi A. 1, Farsi M. 1, Valiellahi M. 3

1Department of Biotechnology and Plant Breeding, College of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.

2Department of Agriculture, Payam Noor University of Mashhad, Mashhad, Iran

3Department of Computer Science, Imam Reza University of Mashhad, Mashhad

Alternative splicing (AS) of mRNA transcripts provides a mechanism by which a single gene can express transcripts that encode proteins with altered functions, and physiological properties. The mechanism of AS is not completely understood but does appear to involve regulatory motifs in the transcript sequence coupled with regulatory proteins that together influence the pattern of splicing. The identification of alternatively spliced genes most often involves examining expressed transcript sequences in the form of expressed sequence tags (ESTs) or, more recently, full-length cDNAs (FL-cDNAs). Variations in mRNA processing including those derived from AS are evident from pairwise comparisons of individual fully processed transcript sequences derived from the same gene, or from examining alignments of the mRNA sequences to their cognate DNA sequence. The PASA (Program to Assemble Spliced Alignments) software includes a set of tools to leverage ESTs and FL-cDNA sequences for eukaryotic gene structure annotation and for studying AS. PASA executes a series of steps to reconstruct transcript isoforms from transcript alignments and then to identify all splicing variations evidenced by differences in gene structures between individual isoforms. Although PASA was originally developed as a genome annotation tool, the above functions of the software provide a standard framework for analyzing AS for any eukaryote provided a complete genome sequence and database of expressed transcript sequences.
Application of Bioinformatics in Genetic Comparison of Caspian Salmon (Salmo Trutta Caspius) Immigrant Forms

Shirangi S. A. ¹, Kalbassi M. R. ¹, ²

¹ Department of Fisheries Faculty of Marine sciences, Tarbiat Modares University, Noor, Mazandaran, Iran.

² Research and Development Center for Biotechnology (RDCB), Tarbiat Modares University, Tehran, Iran.

The Bioinformatics developed during eighties and defined as “the scientific discipline that encompasses all aspects of biological information acquisition, processing, storage, distribution, analysis and interpretation”. It combines the tools of Biology, Chemistry, Mathematics, Statistics, and Computer Sciences to understand life and its processes. The term Marine Bioinformatics refers to the use of computer and networking technologies to gather, store, integrate, analyse, interpret and disseminate the marine organism’s data such as organism’s distribution, description, systematic classification, phylogeny, and their bimolecular structure and sequence data along with the functional aspects using marine proteomics and genomics. The marine bioinformatics includes marine genomics and marine proteomics. Marine genomics refers to any attempt to analyze or compare the entire genome complement of a species of marine origin. This article discusses the genetic variation and genetic differentiation of Caspian salmon immigrant forms by using computer based program, namely GenAlex. Salmo trutta caspius in respect to its reproductive life cycle has two immigrant forms namely fall-run and spring-run. Regarding to lack of information on genetically differences and structure between two groups, in this study, genetic structure of these two immigrant forms were analyzed by using 5 microsatellite loci. As GeneAlex is the main software associated with microsatellite marker, all genetic analysis was done by using this software such as allelic frequencies, observed and expected heterozygosity, deviation from Hardy-Weinberg equilibrium, genetic distance and population distinctive index (Fst). Final results revealed that, 5 microsatellite loci used in this study show low genetic differentiation between fall-runs and spring-runs and the most allelic frequencies as well as more genetic variation were observed in fall-runs.
Analysis of Codon Usage Bias of Kinase Genes for Different Species

Chitsazian F. 1, Ahmadi M. 2, Masoudi-Nejad A. 2

1Institute of Biochemistry and Biophysics and Center of Excellence in Biomathematics, University of Tehran, Iran

2Laboratory of Systems Biology and Bioinformatics (LBB), Institute of Biochemistry and Biophysics and Center of Excellence in Biomathematics, University of Tehran, Iran

Codon usage has direct uses in molecular characterization of species and is also a marker for molecular evolution which has been used in this work for clustering different species using kinase genes. Species were selected from Fungi, Chromalveolata, Excavata, Protista, Animalia and Plantae kingdoms of the Eukarya and Bacterial domain. To understand codon usage bias within kinase genes, a total of 737700 codons with 1414 coding sequences (CDSs) from 47 species were analyzed. A comprehensive codon usage table for selected species was generated, which is the first one for kinase genes. The clustering was performed by the K-mean method and single linkage method using Minitab. Taxonomical classification with a subfamily level resolution was usually seen in higher organisms but in more primitive ones it was not discriminated even at the phyla level. Animalia were classified into two groups, "vertebrates" and "invertebrates" and the clustering pattern remained loyal to taxonomical grouping down the Subfamily level. In Plantae, although monocotyledons and eudicots were grouped in two discrete clusters, there was some incompatibility compared to the finer phylogenetic tree. General conformity between clusters and phylogeny was observed in Fungi despite few discrepancies. The major factor affecting codon usage bias is coding sequence content, which varies in different species. Eukaryotes have stronger repair systems to correct errors in DNA than prokaryotes which may be the reason why phylogeny-consistent clustering based on codon usage is seen in the former but not the latter domain.
Missing Value Imputation in DNA Microarrays Based on Conjugate Gradient Method

Dorri F., Azmi P.

Electrical Engineering Department, Tarbiat Modares University

Microarray technology is a popular gene expression profiling procedure and many tools have been developed for data interpretation. In addition to providing information on specific genes, its additional notable quality is to allow a global overview of gene expression profiles especially tissues under particular conditions. The usual purpose of genome wide microarray gene expression experiments is identification of genes or biological pathways relevant to most common disease phenotypes. Data sets derived from microarray usually have missing values. The missing data clearly reduces optimal recovery of information and, therefore, power of the protocol. Furthermore, implementations of some algorithms applicable to microarray data require complete datasets as input. Various linear and nonlinear imputation methods have been suggested for estimation of missing microarray data. In the present study, a nonlinear algorithm based on the conjugate gradient method (CGBA) was developed and applied to various publicly available datasets. For implementation of CGBA, a subset of genes is identified such that expressions of the genes are most closely related to expression of genes for which data is missing. Optimized conjugate gradient method then estimates the missing values using available information on the subset of genes. Other protocols for estimation of missing values include sequential local least square imputation (SLLSimpute), K-nearest neighbors (KNNimpute) analysis and row average estimation. These protocols and CGBA were applied to some databases and the results were compared. The conjugated gradient method CGBA appeared to be a more robust and sensitive method for missing value estimation in different data sets.
Identification of Alzheimer Related Genes Using Supervised Independent Component

Dorri F. 1, Elahi E. 2, Paylakhi S. H. 3

Electrical Engineering Department, Tarbiat Modares University

Alzheimer’s disease (AD) is the most common neurodegenerative disorder among people aged 65 and older. One of the objectives of the current research is identification of genes and molecular pathways involved in the etiology of AD. One of the approaches implemented for this purpose is analysis of whole gene expression patterns using microarrays. Microarray data is analyzed with various outlooks, and one of these involves application of mathematical tools to best retrieve and reveal information hidden in the vast amount of data produced. In the present study, a supervised independent component analysis (SICA) is applied to two publicly available gene expression microarray data sets with the objective of identifying the most informative genes with regard to AD. ICA is a computational method for decomposing a multi-variant signal into additive components that assumes the source signals are non-Gaussian and mutually independent. In a supervised ICA, information about classes to which samples belong is also provided. We believe the SICA protocol used in this study for analysis of AD data was fruitful. One of its desirable qualities is that it ultimately identified a fairly small number of genes (100 of the 20,000 originally analyzed for expression level). Most of the genes repeatedly identified in multiple runs of the analysis showed biological relevance to the AD phenotype. The algorithm used needs to be applied to microarray gene expression data available for other complex diseases in order to assess its broader value as a general approach for identification of genes relevant to diseases.
Molecular Modeling Investigation of The Interaction between Amlodipine and Human Serum Albumin at Three Different pH: A Comparative with In Vivo like Model

Atarodi P. 1, Saberi M. R. 2, Chamani J. 1

1 Department of Biology, Faculty of Sciences, Islamic Azad University-Mashhad Branch, Mashhad, Iran
2 Medicinal Chemistry Department, School of Pharmacy, Mashhad University of Medical Sciences, Mashhad, Iran

The investigation OF 3-D structure of crystalline albumin showed that HSA contains three homologous domain (I,II,III) and each domain can be divided in to two sub-domain (A and B). The crystallographic analysis has revealed that HSA binding sites of compounds within hydrophobic cavities in sub-domains II A and IIIB, which are corresponding to site I and site II respectively and many drugs can bind to sub-domain IIA. Amlodipine is docked to HSA in different pH by used of moe 2008 program which pH affect on binding energy and inhibit constant HSA, energy minimization is done by use of amber 99 and in order to investigation of interaction between amlodipine with HSA is used of autodock4 program. We applied autodock4 program to calculate the possible conformations of amlodipine on HSA based on Lamarckian genetic algorithm by adopting three different grid box sizes. Modeling studies showed binding energy and inhibit constant in different pH and domains and contain: pH = 6.4, pH = 7.4, , pH= 9.4 in domain I, II, III. Furthermore, Docking data revealed that the stable form is in domain II in each pH and pH=9.4 is the best condition for interaction between amlodipine with HAS with binding energy= -6.39 and inhibit constant=20.88µm
A New Algorithm for Protein Pattern Search
Engareh M., Shiri M. E., Ghaeminejad S., Mohajerani H.

AmirKabir University of Tehran, Tehran, Iran
Tafresh University, Tafresh, Iran
Azad University of Arak, Arak, Iran

Numerous algorithms have been developed for protein pattern search, which use various methodologies and protocols for searching protein pattern. In this paper, we propose a new different form of Boyer-Moore algorithm, which acts well for protein sequences because the shift quantity increases with the pattern length. The two important ideas behind our improvement were to keep a record of all previously matched characters within the current alignment and not move the reading position completely to the end of pattern when a mismatch occurs. The new algorithm reads the characters of the sequence at most once and the preprocessing step that builds the comparable automation is polynomial in size of the pattern. We also develop and test different intend by practical implementations.
Phylogenetic Study on 6 Plants Species Based on Chitinase Gene Alignment

Obeidi L., Kahrizi D.

Agronomy and Plant Breeding Department, Ilam University, Ilam, Iran

Biotechnology for Drought Tolerance Department, Razi University, Kermanshah, Iran

Fungi are one of the most important causes of diseases in plants. One of the ways to control fungal diseases is to limit fungal growth in the plant rhizosphere. Moreover, we can consider existing chitinase genes in plants as one of the most effective ways to produce chitinase enzymes. These enzymes often come to act against the cause of disease, via breaking apart the main structure of fungi. Among the structural parts, the cell wall is the first and the most primary line of defense against enzymes. For this reason, many hydrolytic enzymes, especially chitinases play a major role in destroying the cell wall. Considering the particular importance of these chitinases, in this study we carried out alignment of encoding genes in 6 plants species (Brassica napus, Brassica oleracea, Nicotiana tabacum, Oryza sativa, Capsicum annuum, Zea mays). The results indicate that B. napus, B. oleracea and N. tabacum are in the same cluster and O. sativa and Zea mays are in another cluster, whereas C. annuum is sited in a separate cluster. The species the in same cluster are closely related based on morphological aspects.
Molecular Epidemiology Application to a Mycobacterium Tuberculosis Transmission by MIRU-VNTR plus web site

Jafarian M., Farnia P., Karegar M., Aghalimerza M., Kazempour M., Masjedi M. R., Velayati A. A.

Mycobacterium Research Center (MRC) National Research Institute Of Tuberculosis and Lung Disease (NRILTD), Shahid Beheshti University Medical Camus, Tehran, Iran

Introduction and objective: Molecular Epidemiological analyses are frequently used in epidemiologic testing and only occasionally in forensics. Although methods of molecular epidemiological studies to examine other samples, but there are some limitations such as time of the technique needed to have high values. Recently, mycobacterium interspersed repetitive unit (MIRU) typing has become an important method, as it allows high-throughput, discriminatory and reproducible analysis of clinical isolate. The aim of this study was to describe the genotypes of our clinical isolates by using MIRU-VNTR plus database and to demonstrate easy feasibility of the database in term of application, time effectiveness, and results of analysis. Material and method: The 12 standard MIRU loci were performed on 94 drug resistant, sensitive strain of M.tuberculoses. Results of MIRU-VNTR plus website. Accordingly pattern, of genetic similarity, allelic diversity were identified. Result: we found that the most predominate genetic patterns were Haarlem and LAM family, whereas Cameroon and X were the least prevalence families. Based on the segregation index, discriminatory locus of MIRU (10, 16, 26, 31 and 40) discriminatory power (HGI >= 6), locus of MIRU (23, 27 and 39) as a locus of donor differentiation medium (HGI >= 0.4 - 0.6) and subject to other locus of differentiation is weak donor. Conclusion: MIRU-VNTR plus is a freely accessible website that compares clinical strains with its reference strains. Comparisons can be based on single MIRU-VNTR, spoligotyping, RD, SNP and susceptibility - data, or by a combination of different data sets, which allows a polyphonic typing approach. Also the MIRU-VNTR plus database offers three main functions: Identification of strains, analysis of strains and nomenclature.
Prediction the Tertiary Structure of Extracellular Domains of ROR1, an Approach for Preparation Cancer Vaccines

Shahnazari P.

Avicenna Research Institute, Shahid beheshti University, Tehran, Iran

RECEPTOR TYROSINE KINASE-LIKE ORPHAN RECEPTORs (RORs) are a family of orphan receptors that are related to muscle specific kinase (MuSK) and Trk neurotrophin receptors. ROR1 and ROR2 define as a family of receptors whose extracellular domains contain Ig-like, FZ (frizzled domain) with cysteine-rich region, and kringle domains. The tyrosine kinase domains are followed by serine/threonine- and proline-rich motifs. These receptors were originally identified on the basis of similarity of their Tyrosine Kinase domains to the Trk family of neurotrophin receptors. Recent data has indicated that over-expression ROR1 will result in various forms of human cancer, including those which stem from CNS or PNS neuroectoderm. In the other words, ROR1, is a constitutively expressed tumor-specific cell surface antigen and an ideal target for therapeutic antibodies. Even though, current studies on anti-ROR1 monoclonal antibodies in some cancers such as breast cancer and B-cell chronic lymphocytic leukemia (B-CLL) have begun to shed light on cancer vaccines, making specific monoclonal antibodies that suppress adequately the over-expression of ROR1 have been unsuccessful. Using bioinformatics methods for prediction suitable peptides for preparation anti-human ROR1 can be effective method; however because of the lack of crystallography data, and suitable templates, the prediction tertiary structure of whole ROR1 using modeling and simulation tools is insignificant. A good strategy for the prediction tertiary structure of ROR1 is modeling the tertiary structure of extracellular domains individually. In this study the tertiary structure of Ig-like, FZ and Kringle domains were detected using treading methods (LOMETS and I-TASSER) then they were refined by modeller software for preparation valid models. On the other hand, using epitope prediction software and protein solvent accessibility detection lead to estimate and illustrate some locations on extracellular domains of ROR1 for vaccine preparation. Moreover, the phylogenetic tree, sequence annotation structure (SAS) and the amino acid compositions of particular tyrosine kinases with ROR1 and overlapping the tertiary structure of ROR1, ROR2, TIE1 (tyrosine kinase with immunoglobulin-like and EGF-like domains 1), and MuSK (MUSK_MOUSE Muscle, skeletal receptor tyrosine protein) cause precisely ROR1 structure has analyzed.
Homology Modeling and Docking Studies of Thyroid-Stimulating Hormone (TSH)

Eslami H., Rahimi H. 2, Madadkar-Sobhani A. 3

1. Hurmozgan University of Medical Sciences, Hurmozgan, Iran.
2. Pasteur Institute of Iran, Tehran, Iran
3. Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran

Thyroid-stimulating hormone (TSH; thyrotropin), a 30-kDa glycoprotein synthesized and secreted from the anterior pituitary gland, is one of the key proteins that controls thyroid function. Thyroid-stimulating hormone stimulates the thyroid gland to secrete the hormones thyroxine (T4) and triiodothyronine (T3). The tertiary structure based on NMR or x-ray crystallography has not been reported for this protein so far. A structural model of thyrotropin would be instrumental in gaining understanding of hormone structure-function relationships and can provide valuable information regarding the hormone interaction with its receptor (TSHR). In this study, we present a model of TSH using homology modeling, based on the human follicle stimulating hormone (PDB ID: 1XWD) as template. Template was selected using multiple sequence alignment and identity score. A preliminary model was made with the MODELLER package. The final model was further assessed by PROCHECK, which showed that the model had good geometrical properties. Molecular dynamics simulation of the resulting model is in progress using GROMACS version 4.
Prediction of Three-Dimensional Structure of Flagellin from *E. coli*

Rahimi H.\(^1\), Kashef B. \(^1\), Madadkar-Sobhani A.\(^2\)

\(^1\) Pasteur Institute of Iran, Tehran, Iran  
\(^2\) Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran

Most Gram-negative bacteria express flagella surface structures that confer motility. Flagella are composed of a basal body that serves as a rotatory motor, a filament that extends into the space around the bacterium to provide motive force, and a hook that connects the two. The filament consists of a long homopolymer of a single protein, flagellin, with a small cap protein at the end. Flagella are unusually conserved among diverse bacterial species. Over 50 genes are involved in the synthesis and function of flagella, suggesting that their preservation and role in chemotaxis and motility are important in the survival of many organisms. Recent reports indicate that the component of flagella responsible for eliciting host immune responses is the filament protein flagellin. The host response to flagellins from *E. coli* and several other bacteria is mediated by Toll-like receptor 5 (TLR5), which signals through nuclear factor kappa B (NF-KB) to induce transcription of pro-inflammatory cytokines such as interleukin (IL)-8. The aim of this study is the prediction of the three-dimensional structure of fliC encoded protein, flagellin, in *E. coli* using homology modeling method. Using sequence alignment against PDB, Salmonella typhimurium flagellin (PDB ID: 1IO1) was selected as the template with 50% identity. Then 10000 models generated using MODELLER version 9.7 package and the model corresponding to the lowest DOPE score was selected for the next step. The stability of the proposed model investigated using PROCHECK. At the present time, molecular dynamics simulation for 5 ns using GROMACS package is in progress.
The Suitable Binding Sites Finding of the Interaction between Human Serum Albumin and Fluoxymestrone as a Model of Drug Binding Protein

Sabaghian M. 1, Saberi M. R. 2, Housaindokht M. R. 3, Chamani J.1

1. Department of Biology, Faculty of Sciences, Islamic Azad University-Mashhad Branch, Mashhad, Iran

2. Medicinal Chemistry Department, School of Pharmacy, Mashhad University of Medical Sciences, Mashhad, Iran

3. Department of Chemistry, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran

Serum proteins have many important functions that are of great interest in pharmaceutical science, medical science, biology, and chemistry; some of them are related to the fixation and transport of metabolites, hormones, and exogenous, substances like pharmaceutical drugs. Molecular modeling method has been employed to promote the understanding of the interaction of drugs and HSA. The 3D structure of crystalline albumin has revealed that HSA comprises of three homologous domains that assemble to form a heart-shaped molecule. HSA is monomeric but contains the three structurally similar helical domains (I-III); each domain has two sub-domains (A and B), which have six (A) and four (B) α-helices, respectively. The principal ligand-binding regions of HSA are located in the hydrophobic cavities of sub-domains IIA and IIIA, called site I and site II, respectively, and one tryptophan residue (Trp214) of HSA exists in sub-domain IIA. The 3D structure of HSA was obtained from the Protein Data Bank database. The autodock4 program was chosen to examine the binding mode of fluoxymestrone at the active site of HSA. During docking process, a maximum of 20 conformers were considered for this compound. The best binding energy result is seen in site II and its value is -6.09 kcal/mol. At this site, are calculated values of electrostatic energy, torsional free energy, and association constant 0.05 kcal/mol, +0.55 kcal/mol, 34.39 μM, respectively. Moreover, there is not hydrogen binding between fluoxymestrone and HSA. The results indicated that the hydrophobic bond increased the hydrophobicity and decreased the hydrophilicity to stability the fluoxymestrone-HSA system. The calculated binding Gibbs free energy was not very close to the experimental data in some degree. A possible explanation may be that the X-ray structure of the protein from crystals differs from that of the aqueous system used in this study.
Is *acdS* Gene Inherited by Horizontal Gene Transfer?
Farajzadeh D., Yakhchali B., Aliasgharzad N., Sokhandan Bashir N.

*National Institute of Genetic Engineering and Biotechnology, Tehran, Iran*

We present evidence, based on in silico experiments, supporting the notion that the use of complete nucleotide sequences of the *acdS* gene to assess horizontal gene transfer (HGT) leads to more reliable results than when partial sequences are used. A phylogenetic tree based on the coding region of the complete sequences of all known ACC deaminase genes deposited in NCBI revealed that, the studied strains are grouped with their nearest phylogenetic relatives and no distribution of these strains is observed in the tree. The 16S rRNA tree of some studied strains also showed the same pattern with regard to their *acdS* tree. Therefore, our results somehow substantiate earlier reports suggesting that the *acdS* genes may have undergone HGT but in contrast with previous findings, it is revealed that HGT, in the case of the *acdS* gene, is limited to very closely related bacteria.
Sequencing and Comparison of the Nuclear Ribosomal Internal Transcribed Spacer Region for Iranian Black Caraway Populations

Azimzadeh M.1, Amiri R. 1, Assareh M. H.2, Forootan M.2, Bihamta M. R.3

1 Department of Agronomy and Plant Breeding Sciences, College of Abouraihan, University of Tehran, Iran
2 Research Institute of Forests and Rangelands, Tehran, Iran
3 Department of Agricultural Biotechnology, College of Agriculture and Natural Resources, University of Tehran, Iran

Black caraway (Bunium persicum Boiss.) belonging to the family Apiaceae is an economically important medicinal and aromatic plant found naturally growing in dry temperate regions of Iran. The present investigation assessed the difference of the sequences of ITS2 region between nine populations of black caraway from Iran by multiple sequence alignment. The sequencing was performed on an Applied Biosystems 3130xl genetic analyzer, using the BigDye terminator cycle sequencing kit (Ver. 3.1) following the manufacturer’s instructions. Multiple sequence alignment was done by MegAlign (DNASTAR Lasergene software, version 7.1) using ClustalW method. This region showed some mutations and was different between the compared populations. The multiple sequence alignment grouped the populations into 4 main clusters. Minimum dissimilarity was between Mashhad and Yazd ecotypes (2.8 %) and maximum was between Asgari and Sirch ecotypes (14.4%).
Introducing 1-methyl Malate as a New and Potent Inhibitor of Class A β-Lactamase by Computational Studies

Mirzaie S., Shahverdi A. R.

Department of biochemistry, Islamic azad university, Sanandaj branch, Sanandaj, Iran

1-methyl malate is one of the important organic compounds in Berberis integerrima fruits that enhance the antibacterial activity of ampicillin against Staphylococcus aureus. The crystal structure of Staphylococcus aureus β-lactamase (PDB ID: 3BLM) was selected for prediction of binding mode of 1-methyl malate to class A lactamase. In this study AutoDock 3.0.5 along with its LGA algorithm were used for automated flexible ligand docking and docking accuracy and reliability of the estimated inhibition constants were evaluated. In the active site of β-lactamase, binding and catalytic residues like Ser70 and Ser130 are directly involved in the catalytic mechanism. Arg244 interact with the substrate carboxylate. So, its role is positioning of the β-lactame substrate in the active site. Structural analogs of 1-methyl malate like Citrate, isocitrate and amino analog of citric acid have been recently shown, by X-ray diffraction analysis, to perfectly fit into the active site of Bacillus licheniformis β-lactamases with the same inhibition mechanism, and to behave as a modest inhibitor of this serine enzyme. Based on our simulation results, and also with regard to prior studies, we propose that 1-methyl malate could be act as inactivator of class A β-lactamases in Staphylococcus aureus and could be suitable target for further studies. The results have good agreement compare to the experimental data.
Sarcin/ricin rRNA Motif Detection in PDF Files 3G4S, 3G6E, and 3CXC using FR3D Program

Tolunia B., Sepahi S., Jalal R.

Department of Chemistry, Ferdowsi University of Mashhad, Mashhad, Iran

Research and Technology Center of Biomolecules, Ferdowsi University of Mashhad, Mashhad, Iran

Institute of Biotechnology, Ferdowsi University of Mashhad, Mashhad, Iran

National Institute of Genetic Engineering and Biotechnology, Tehran, Iran

Systematic and exhaustive RNA motif identification and classification is crucial for integration of RNA structural and sequence data. New methods are described for finding recurrent three-dimensional (3D) motifs in RNA atomic-resolution structures. FindRNA 3D (FR3D) is a suite of Matlab programs that implement geometric, symbolic, or mixed search. The goal of RNA 3D motif searching is to find and rank candidate motifs according to how closely they resemble the structure of the query motif. The inputs include a query motif in the form of a list of m nucleotides from a particular RNA 3D structure file, a set of RNA 3D structure files to search, and a cutoff discrepancy, D0. The output is a list of candidate motifs from these structure files, sorted according to the geometric discrepancy between the candidate and query motifs. We used FR3D for annotations of the sarcin/ricin motif in the 50S ribosomal subunit in Haloarcula Marismortui, Protein Data Bank PDB files 3G4S, 3G6E, and 3CXC [4]. Geometric, symbolic, or mixed representations of RNA structure have been implemented in FR3D. The complete sarcin/ricin motif has nine nucleotides, and we used a six-nucleotide core sarcin/ricin submotif as the query motif. The geometric, symbolic, and mixed searches were conducted with the discrepancy cutoff D0 set to 0.5. We found specific the relationships between these six bases in FR3D as follows that G2701/A2694 - trans Hoogsteen/suger-edage, A2702/U2693 - trans Watson-crick/Hoogsteen and A2703/A2691 - trans Hoogsteen/Hoogsteen in the query motif.
Computational Simulation of Gramicidin like Channel as A Peptide Nanotube

Saeedi A, Kyani A., Madadkar-Sobhani A. , Goliaei B.

Department of Biophysics, Institute of Biochemistry and Biphysics, University of Tehran, Iran

The possible conformations of integral membrane proteins are restricted by the nature of their environment. As possible models of the intramembranous regions of integral membrane proteins, three types of regular structures are more common, of which the alpha helix and the beta-pleated sheet, are regularly occurring structural features of soluble proteins. The third is a class of conformations called beta helices. This later kind of helices has unique features which make them particularly well-suited to the lipid bilayer environment. Such structures could function as transmembrane ion channels. Recent studies show that a short beta sheet tetraicosapeptide VSLGLSIAFVVSIAWFSARSG, where all As are D-alanine), can form gramicidin-like beta12-helical channels in membranes. We have performed molecular dynamics simulations of the interactions of this synthetic peptide with palmitoylphosphatidylcholine (POPC) lipid bilayers. We used various initial positions and orientations of the peptide with respect to the lipid bilayer, including a surface-bound state parallel to the interface and a trans-membrane state. Our simulations show that the peptide is stable in lipid environment, but it is unstable in aqueous environment.
Prediction of the Mode of Interaction between Monoterpenes and the Nitroreductase from Enterobacter Cloacae by Docking Simulation

Mirzaee S. 1, Shahverdi A.R. 2, Madadkar-Sobhani A.3

1Department of Biochemistry, Islamic Azad University, Sanandaj Branch, Sanandaj, Iran
2Department of Pharmaceutical Biotechnology and Biotechnology Research Center, Faculty of Pharmacy, Tehran University of Medical Sciences, Tehran, Iran
3Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran

Nitrofurantoin is used for the treatment of urinary tract infections and recently nitrofurantoin-resistant strains of various clinical bacteria have been found. Flavin-containing nitroreductases catalyze the reductive activation of nitrofurantoin and other nitro drugs to bactericidal compounds. Since, monoterpenes could reduce the antibiotic resistance in some bacteria; we simulated the theoretical mode of binding of five chemicals to the Enterobacter cloacae nitroreductase. In this study the crystal structure of Enterobacter cloacae nitroreductase with PDB code 1KQB was selected for prediction of binding mode of five monoterpenes to nitroreductase. Docking simulations were done by AutoDock 3.0.5. Coenzyme FMN and tight bounded water molecules to nitroreductase active site were retained in docking calculations. Theoretical data showed monoterpenes interact with Phe 124 and Gly 166 by hydrophobic interactions. Also monoterpenes with carbonyl group interacts with NH group of residue Thr 44 by hydrogen bond. The pearson correlation between monoterpenes estimated free energy of binding and inhibitory activity (MIC) was done and R2 was 0.79. Based on our results, enhancement of nitro drug potency in the presence of monoterpenes may be the result of modulation of nitroreductase activity. Binding of the nitroreductase to monoterpenes may decrease the efficient conversion of toxic reactive intermediates to final products lacking bactericidal activity.
Quantitative Structure Activity Relationship Analysis of Some Isocoumarins as Pancreatic Cholesterol Esterase Inhibitors

Navidpour L.¹, Ebrahim-Habibi A.²

¹. Department of Medicinal Chemistry, Faculty of Pharmacy, Tehran University of Medical Sciences, Tehran 14174, Iran.

². Endocrinology and Metabolism Research Center, Tehran University of Medical Sciences, 5th floor Dr. Shariati Hospital, North Kargar Avenue, 14114 Tehran, Iran.

A Quantitative Structure Activity Relationship (QSAR) was performed on some isocoumarins regarding their inhibitory activity against pancreatic cholesterol esterase. The compounds in the selected series were studied using QuaSAR module of Molecular Operating Environment (MOE). Significant equations that were derived from regression analysis showed the importance of surface-charge related descriptors contribution towards the activity. The relation of positively charged surface area (ASA+, CASA+) contribution suggests the significance of an interaction with the nucleophilic activated serine of the active site that would be essential for cholesterol esterase inhibitory activity. Addition of a third parameter, the polar surface area (TPSA) which is negatively related to activity and includes the contribution of polar atoms, results in a model with correlation coefficient ($r^2$) of 0.91 and a cross validation coefficient of $q^2=0.80$ using the leave one out cross validation. The physicochemical interpretation of the descriptors and the models will be discussed.
Modeling the Epithelial-Mesenchymal Transition and its Effect on Tumor Invasiveness and Metastasis

Rouhani M.

Bioinformatics Department, institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran

The genetic factors that enable primary tumor cells to invade and metastasize are to a great extent unknown. Based on analogies with the embryonic process, it has been proposed that carcinoma cells may complete invasive and metastatic steps by the activation of the so-called epithelial-mesenchymal transition (EMT), characterized by the loss of epithelial phenotypes and the acquisition of mesenchymal properties. Although more pre-clinical and clinical studies are required to elucidate the process of EMT and the molecular pathways involved in carcinoma cells, the development of mathematical models grounded on available experimental data can contribute to the generation of rational hypotheses that can be validated through future experimental work. Hence, the known molecules involved in the EMT and their interactions were scrutinized and a complicated network of receptors, signal transducers and transcription factors were drawn. Three key regulator proteins involved in cell-cell adhesion were chosen and a simple mathematical model was proposed for the EMT process. The model was based on ordinary differential equations and contained rates of synthesis and degradation of molecules and the effect they have on the transcription rates of each other. We utilized MATLAB codes and referred to an empirical study in order to validate our model. The model was consistent with the available experimental data and can be extended from subcellular state to cellular and tissue levels and may also be used to predict possible treatment strategies.
A Novel View of the Interaction between Human Serum Albumin and Estradiol: A Molecular Docking Approach

Amani Malkeshi N. ¹, Chamani J.¹, Saberi M. R. ²

¹ Department of Biology, Faculty of Science, Islamic Azad University, Mashhad Branch, Mashhad Iran

² Medical Chemistry Division, Scholar Pharmacy, University Complex, Mashhad, Iran

HSA is a globular protein consisting of a single peptide chain of 585 amino acids. It is composed of three homologous domains, III and III, which have similar three-dimensional structure. Each domain can be divided into two subdomains A and B, which have six (A) and four (B) α-helix respectively. The principal ligand-binding regions of HSA are located in the hydrophobic cavities of subdomains IIA and IIIA, called site I and site II, respectively. The inside wall of the pocket of subdomain IIA is formed by hydrophobic side chains, whereas the entrance of the pocket is surrounded by positively charged residue such as ARG257, ARG222, LYS199, HIS242, ARG218 and LYS195. The crystal structure of human serum albumin was taken from the Brookhaven Protein Data Bank (entry code 1A06). The most possible interactions between HSA and estradiol ranked with minimize energy by inhibit constants that are taken of Autodock4 program. The best result of inhibit constants ($K_I=6.4$ micro molar) show that region interaction between human serum albumin and estradiol was located in domain I, that confirmed by fluorescence spectroscopy.
TITAN Protein Secondary Structure Assignment by Emphasis on Tight Turns

Afzalifar R. ¹,², Madadkar-Sobhani A.¹, Minuchehr Z.³, Goliaei B. ¹

¹Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran

¹Dental Research Center, Tehran University of Medical Sciences, Tehran, Iran

³National Institute of Genetic Engineering and Biotechnology

Turns are irregular secondary structure elements that provide a direction change of the polypeptide chain. A turn formed by six or less amino acids is called a tight turn to be distinguished from the loose turn. Tight turns play an important role, especially in globular proteins, both functionally and structurally. They are generally categorized as delta, gama, beta, alpha and pi turns, based on having two to six amino acid residues, respectively and each has several subtypes according to the backbone dihedral angles in the inner residues. Despite the fact that there are many secondary structure assignment methods (SSAMs), to our best knowledge, none of them can differentiate between various types of tight turns. For instance, two mostly used SSAMs, DSSP and STRIDE, only detect beta turns and their corresponding subtypes. Here, we report a new program called tight turn analyzer, which in addition to detection of regular secondary structure elements, i.e, alpha helices and beta sheets, can also detect and report all kinds of tight turns and their subtypes.
Molecular Modeling Studies on Cholesterol Esterase Inhibitors: Proposition of New Ligands, and Insight into Possible Additional Effects

Ebrahim-Habibi A., Mohammadi B., Haghpanah V., Larijani B.

Endocrinology and Metabolism Research Center, Tehran University of Medical Sciences, 5th floor Dr. Shariati Hospital, Tehran, Iran

Cholesterol esterase (EC 3.1.1.13) is a potential target in the treatment of obesity and hypercholesterolemia. From the available reports on its inhibitors, a series of isocoumarins with maximal and minimal inhibitory activities were selected as a primary set of ligands for the modeling study. The human enzyme structure (1FW6.pdb) was first protonated and the active site residues repositioned with the use of the rotamer library implemented in MOE 2008.10. Selected ligands were docked onto this structure with the use of Autodock Vina. Important interactions were then summarized in a pharmacophore and screened against the MOE database of 653000 small molecules. The 152 obtained hits were docked onto the enzyme’s active site using the docking module of MOE and the 10 best scoring ligands were selected, alongside six other compounds that showed potential of inducing variation in the structure based on visual inspection of the structures. These ligands were then redocked with Auto dock Vina, and the ones showing additional interactions as compared with the original set were retained. As a side experiment, enzymes possessing similar active sites, but no significant overall similarity of structure were found using the PDBSite database (including hydroxyl nitrile lyase and protease) and the best ligands obtained from the above modeling studies were docked onto these structures in order to get an insight about possible undirect effects of these compounds. In conclusion, new ligands are proposed for cholesterol esterase.
**Tertiary Structure (3D) Prediction of Extracellular Domains of Ror1**

Shahnazari P., Jedi-Tehrani M., Rabbani H.

*Avicenna institute, Monoclonal Antibody Research Center, Avicenna Research Institute, ACECR, Tehran, Iran*

Recently, we and others have reported the expression of the receptor tyrosine kinase Ror1 in patients with acute and chronic lymphocytic leukemia. Ror1 belongs to a family of orphan receptors that are related to the muscle specific kinase (MuSK) and Trk neurotrophin receptors. The extracellular part of Ror1 contains three major Ig-like (Ig), cysteine rich (CRD) and kringle (Kr) domains. The CRD and Kr domains of the Ror family proteins are known as potential ligand binding sites and suitable for targeting with monoclonal antibodies. No functional antibody has yet been generated against Ror1 capable of inducing apoptosis in Ror1 expressing cancer cells. One approach for generating functional antibodies is to use synthetic peptides as immunogens from the ligand binding region. However, due to the lack of any crystallographic data on Ror1, the only option is the prediction of the 3D structure of Ror1 by Bioinformatic approaches. In this study the tertiary structure of Ig-like, CRD (Frizzled) and Kringle domains were predicted using treading methods (LOMETS and I-TASSER). Furthermore, using epitope prediction software and protein solvent accessibility detection, some locations on extracellular domains of Ror1 were determined as potential target sites. Moreover, the phylogenetic tree, sequence annotation structure (SAS) and the amino acid compositions and overlapping of the tertiary structures of Ror1, Ror2, Tie1, and MuSK are discussed.
Cholesterol Oxidase Gene Analysis Derived from a Native *Rhodococcus* sp. Strain 501

Lashkarian H. E., Raheb J, Shahzamani K, Minuchehr Z, Bardania H

*National Institute of Genetic Engineering and Biotechnology, Faculty of Science, Tehran, Iran.*

*Lorestan University, Lorestan, Iran.*

Cholesterol oxidase (Cho) is an enzyme that catalyzes the oxidation of cholesterol and converts 5-cholesten-3β-ol into 4-cholesten-3-one. Due to extensive applications of Cho to medicine, food and agricultural industries; isolation of bacterial strains producing Cho is of great importance. Recently, we have isolated a Cho producing bacteria from soil. Using phylogenic tests, the isolated bacterium was identified as *Rhodococcus* sp. Strain 501 and was deposited in Genebank with the Accession number: FN298676. The Cho gene was amplified by PCR, cloned into the STV28 cloning vector and then sequenced. Methods: In the present study, nucleotide sequence of Cho gene (Accession number: FN421337) was aligned and analyzed using sequence analyzing programs such as Chromas and Mega 4 softwares. Using Blast P, the amino acid sequence was compared with Cho proteins of other bacteria. A phylogenic tree was also constructed by Mega 4 and multiple alignments were carried out using Clustal W2 method. Results: The nucleotide sequence of the fragment containing 1.6 Kb Cho gene, which encodes an extracellular Cho enzyme was determined. Nucleotide sequences of the Cho gene was aligned and showed a 99% homology with the other bacterial Cho genes. A single open reading frame encodes cholesterol oxidase containing 533 amino acids. Phylogenetic analysis of amino acid sequences showed a 100% homology with Cho proteins produced by other bacteria. Moreover, multiple alignments of amino acid sequences and their comparison with the other Cho proteins derived from bacteria showed a very close similarity in amino acid sequences. Discussion: These results should provide a new opportunity for finding correlation between the Cho gene and other related genes. Our results can be useful for the prediction of the secondary and tertiary structure of the Cho protein.
Modeling of Mutation in Phenylalanine Hydroxylase S231F Using Computational Methods

Afzalifar R. 1, Madadkar-Sobhani A. 1, Mahmoudian M. 2

1 Bioinformatics Department, Institute of Biochemistry and Biophysics, University of ehran, Iran

2 Department of Pharmacology, School of Medicine, Iran University of Medical Sciences, Iran

Phenylketonuria (PKU), the most frequent disorder of amino-acid metabolism, is caused by mutations in human phenylalanine hydroxylase gene (PAH), leading to deficient enzyme activity. Previously reported PAH gene mutation was detected in Serbian patients with PKU. Analysis of this mutation shows mutation of Ser-231 to phenylalanine. This enzyme is not active (less than 1% of intact enzyme). We used molecular modeling techniques to study this mutation. We used MODELLER software to build the mutated phenylalanine hydroxylase, using 1J8T, 1PHZ and 2PAH as templates. Then 10 ns molecular dynamics simulation was carried out using GROMACS version 4. The final stable structure shows the active site of enzyme. The active site shows difference with that of active enzyme. A picture of active site of both active and none active phenotypes and various residues will be presented. We conclude that change in the active site will result in the inactivity of the enzyme and as a result PKU.
Modeling Three Dimensional Structures of WW Domain of WWOX Using Homology and Molecular Dynamics Simulation

Jamshidiha M. , Madadkar-Sobhani A. , Ghahremani M. H.

Department Pharmacology and Toxicology, Faculty of Pharmacy, Tehran University of Medical Sciences, Tehran, Iran.

Bioinformatics Department, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran.

Department Pharmacology and Toxicology, Faculty of Pharmacy, Tehran University of Medical Sciences, Tehran, Iran

WWOX is a tumor suppressor, a defect in which has been found in many cancers such as breast, prostate and colon carcinoma. WWOX is activated through phosphorylation and mediates its effect through interactions with different proteins like p53 and JNK. Due to the importance of phosphorylation in its activation, we evaluated proteins that maybe effective in this pathway. We evaluated WWOX interaction with p38, a MAPKinase protein using immuno precipitation. Also, WWOX interaction with c-ABL, a putative tyrosine kinase, was examined using in silico studies. The WW domain of WWOX was modeled using homology modeling. Using sequence alignment against PDB, an appropriate model with 65% identity was selected. Then 10000 models were generated using MODELLER version 9.7 package and the model corresponding to the lowest DOPE score was selected for the next step. Molecular dynamic simulations for 30 ns were carried out using the GROMACS package. The average structure was used for docking against Rosseta which is currently in progress.
Comparison of Dock Techniques in Structural Based Inhibitor Design of HIV-gp41

Gomari H., Sardari S., Azadmanesh K., Mahboudi F.

Bioinformatics and Drug Design Unit, Pasteur Institute of Iran, Tehran, Iran

Designing new drugs is necessary in cases like HIV and other infections due to drug resistance, high cost of production, low stability and difficult usage. In our study, we have focused on gp41 of HIV which mediates fusion between viral and cellular membranes. Two lead molecules namely NB-2 and NB-64, which decrease HIV infection as fusion inhibitors targeting gp41, were considered. Accordingly, we started to look for similar molecules through database mining, like ZINC and Pubchem. Argus lab and FlexX \textit{in silico} application tools had been used. Data base docking was used in both and ionic interactions in one of them. The grid-box interaction site was that of lead structures. To find the best condition for docking, different conformations and surrounding conditions were used for gp41 docking. Molecular docking analysis had calculated minimum energy for each molecule, which was determined by similarity search. Dock projects and their results were compared with each other and experimental data. In this study, we have found these applications appropriate for revealing the best protein inhibitor because our experimental results had validated the computational outcomes. This correlation allows for continuation of our study to find the best molecules which suppress the HIV fusion process.
Molecular Dynamics Study of Human Proinsulin-Protein A Fusion System

Dehghani Amirabad A., Sardari S., Mahboudi F.

Bioinformatics and Drug Design Unit, Recombinant Protein Unit, Pasteur Institute of Iran, Tehran, Iran.

Production of native proinsulin in Escherichia coli is still a challenging task in biotechnology and belongs to the class of the difficult-to-express proteins in E. coli. Problems mainly arise due to its small size, a high proteolytic decay and necessity to native disulfide patterns. To circumvent these problems, fusion techniques have been widely used. One of the promising fusion systems for the production of native proinsulin in the periplasm is single or double Z domains of staphylococcus protein A that is fused at the N-terminus of proinsulin, at this state the oxidative environment of periplasm helps to form native disulfide patterns. It was suggested that acidic region of the Z-domain of protein A might interact with the basic residues of proinsulin via electrostatic interactions, thus protecting the fusion protein from proteolysis. In this study, we investigated hypothesis of proteolytic protection of protein A through molecular dynamics simulation with Gromacs 4.0.3. The electrostatic interactions were calculated using the Particle-Mesh Ewald model with a cut-off of 14 A. Data analysis suggests this protection is due to electrostatic interactions between basic residues of the proinsulin B chain with acidic regions of Z domain which shifts R groups of Tyr16-Leu17, Phe25-Tyr26 (PI Protease cleavage sites) inside the protein and induces the formation of protected soluble conformations. This investigation can help design new fusion forms for pharmaceutical applications as well as expression system studies.
A Matrix Based Pattern Template for Protein Patterns Induced on Genome

Foroughmand Araab M. H. 1, Goliaei B. 1, Sadeghi M. 2, Alishah K. 3

1Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran

2National Institute for Genetic Engineering and Biotechnology, Tehran, Iran

3Sharif University of Technology, Tehran, Iran

Except in some special cases (e.g. RNA editing) protein sequence is exactly determined from its corresponding genomic sequence. Protein sequence motifs which are widely used in protein sequence analysis are also a mapping from genomic sequences, thus, it is expected to be found in genomic sequences too, of course with different manners. We are presenting a very simple representation of genomic patterns which is a matrix like the BLOSUM and PAM matrices, except that, its genomic representation side also stores intra-codon dependency of nucleotides.
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdollahi H.</td>
<td>Babaabasi B.</td>
</tr>
<tr>
<td>Abdollapour N.</td>
<td>171</td>
</tr>
<tr>
<td>Abedini R.</td>
<td>Babaei F.</td>
</tr>
<tr>
<td>Afsharifar A. R.</td>
<td>141</td>
</tr>
<tr>
<td>Afzalifar R.</td>
<td>Babaeian Jelodar N.</td>
</tr>
<tr>
<td>Aghaei Naser Gh.</td>
<td>149</td>
</tr>
<tr>
<td>Aghasadeghi M. R.</td>
<td>Bambaei B.</td>
</tr>
<tr>
<td>Ahadi A. M.</td>
<td>178</td>
</tr>
<tr>
<td>Ahangari Cohan R.</td>
<td>Baradaran M.</td>
</tr>
<tr>
<td>Ahmadi J.</td>
<td>209</td>
</tr>
<tr>
<td>Ahmadi M.</td>
<td>Bardania, H.</td>
</tr>
<tr>
<td>Ahmadpanah J.</td>
<td>213</td>
</tr>
<tr>
<td>Ahravian H.</td>
<td>Batebi S.</td>
</tr>
<tr>
<td>Akbari Tusi Sh H.</td>
<td>33, 48, 54, 56, 129</td>
</tr>
<tr>
<td>Akbarzadeh Sharbaf S.</td>
<td>Behjatnia S. A. A.</td>
</tr>
<tr>
<td>Alemzadeh A.</td>
<td>181</td>
</tr>
<tr>
<td>Aliakbari F.</td>
<td>Behzadi M.</td>
</tr>
<tr>
<td>Aliakbari M.</td>
<td>189</td>
</tr>
<tr>
<td>Aliasgharzad N.</td>
<td>Beigi M. M.</td>
</tr>
<tr>
<td>Alishah K.</td>
<td>200</td>
</tr>
<tr>
<td>Alisoltani A.</td>
<td>Bigdeli S.</td>
</tr>
<tr>
<td>Alizadeh H.</td>
<td>217</td>
</tr>
<tr>
<td>Amani J.</td>
<td>Bihamta M. R.</td>
</tr>
<tr>
<td>Amani Malkeshi N.</td>
<td>192</td>
</tr>
<tr>
<td>Amid Motlagh M. H.</td>
<td>Bonifer C.</td>
</tr>
<tr>
<td>Amin Afsar M.</td>
<td>199</td>
</tr>
<tr>
<td>Amiri R.</td>
<td>Borna H.</td>
</tr>
<tr>
<td>Arjomandnezhad M.</td>
<td>98</td>
</tr>
<tr>
<td>Asgharzadeh M.</td>
<td>Bozorgmehr M. R.</td>
</tr>
<tr>
<td>Ashrafi E.</td>
<td>182, 183</td>
</tr>
<tr>
<td>Aslanzadeh V.</td>
<td>C</td>
</tr>
<tr>
<td>Assoode A.</td>
<td>Chamani J.</td>
</tr>
<tr>
<td>Assad M. T.</td>
<td>135, 136, 141, 156, 157, 182, 183</td>
</tr>
<tr>
<td>Assareh M. H.</td>
<td>192</td>
</tr>
<tr>
<td>Atarodi P.</td>
<td>199, 208</td>
</tr>
<tr>
<td>Ayat H.</td>
<td>Chitsazian F.</td>
</tr>
<tr>
<td>Azadmanesh K.</td>
<td>189</td>
</tr>
<tr>
<td>Azimzadeh Jamalkandi S.</td>
<td>Christophe Nebel J.</td>
</tr>
<tr>
<td>Azimzadeh M.</td>
<td>35</td>
</tr>
<tr>
<td>Azmi P.</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>Dadras A. R.</td>
</tr>
<tr>
<td></td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>Amin Afsar M.</td>
</tr>
<tr>
<td></td>
<td>205</td>
</tr>
<tr>
<td></td>
<td>Amid Motlagh M. H.</td>
</tr>
<tr>
<td></td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>Amin Afsar M.</td>
</tr>
<tr>
<td></td>
<td>171</td>
</tr>
<tr>
<td></td>
<td>Amiri R.</td>
</tr>
<tr>
<td></td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>Arjomandnezhad M.</td>
</tr>
<tr>
<td></td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Asgharzadeh M.</td>
</tr>
<tr>
<td></td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>Ashrafi E.</td>
</tr>
<tr>
<td></td>
<td>65, 140</td>
</tr>
<tr>
<td></td>
<td>Aslanzadeh V.</td>
</tr>
<tr>
<td></td>
<td>43, 44, 102</td>
</tr>
<tr>
<td></td>
<td>Asoode A.</td>
</tr>
<tr>
<td></td>
<td>141</td>
</tr>
<tr>
<td></td>
<td>Assareh M. H.</td>
</tr>
<tr>
<td></td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>Atarodi P.</td>
</tr>
<tr>
<td></td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>Ayat H.</td>
</tr>
<tr>
<td></td>
<td>192</td>
</tr>
<tr>
<td></td>
<td>Azadmanesh K.</td>
</tr>
<tr>
<td></td>
<td>215</td>
</tr>
<tr>
<td></td>
<td>Azimzadeh Jamalkandi S.</td>
</tr>
<tr>
<td></td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>Azimzadeh M.</td>
</tr>
<tr>
<td></td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>Azmi P.</td>
</tr>
<tr>
<td></td>
<td>190</td>
</tr>
<tr>
<td></td>
<td>Babaei F.</td>
</tr>
<tr>
<td></td>
<td>171</td>
</tr>
<tr>
<td></td>
<td>Babaeian Jelodar N.</td>
</tr>
<tr>
<td></td>
<td>45, 142</td>
</tr>
<tr>
<td></td>
<td>Bambaei B.</td>
</tr>
<tr>
<td></td>
<td>176</td>
</tr>
<tr>
<td></td>
<td>Baradaran M.</td>
</tr>
<tr>
<td></td>
<td>139</td>
</tr>
<tr>
<td></td>
<td>Bardania, H.</td>
</tr>
<tr>
<td></td>
<td>212</td>
</tr>
<tr>
<td></td>
<td>Batebi S.</td>
</tr>
<tr>
<td></td>
<td>69, 170</td>
</tr>
<tr>
<td></td>
<td>Behjatnia S. A. A.</td>
</tr>
<tr>
<td></td>
<td>178</td>
</tr>
<tr>
<td></td>
<td>Behzadi M.</td>
</tr>
<tr>
<td></td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Beigi M. M.</td>
</tr>
<tr>
<td></td>
<td>145</td>
</tr>
<tr>
<td></td>
<td>Bigdeli S.</td>
</tr>
<tr>
<td></td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Bihamta M. R.</td>
</tr>
<tr>
<td></td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>Bonifer C.</td>
</tr>
<tr>
<td></td>
<td>131</td>
</tr>
<tr>
<td></td>
<td>Borna H.</td>
</tr>
<tr>
<td></td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Bozorgmehr M. R.</td>
</tr>
<tr>
<td></td>
<td>182, 183</td>
</tr>
<tr>
<td></td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>Chamani J.</td>
</tr>
<tr>
<td></td>
<td>135, 136, 141, 156, 157, 182, 183</td>
</tr>
<tr>
<td></td>
<td>192</td>
</tr>
<tr>
<td></td>
<td>199</td>
</tr>
<tr>
<td></td>
<td>Chitsazian F.</td>
</tr>
<tr>
<td></td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>Christophe Nebel J.</td>
</tr>
<tr>
<td></td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>Dadras A. R.</td>
</tr>
<tr>
<td></td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>Dalir Fardoori R.</td>
</tr>
<tr>
<td></td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>Daneshpour M.</td>
</tr>
<tr>
<td></td>
<td>164, 167, 168</td>
</tr>
<tr>
<td></td>
<td>Darvishalipoor S.</td>
</tr>
<tr>
<td></td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Dehghan M.</td>
</tr>
<tr>
<td></td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Dehghani A.</td>
</tr>
<tr>
<td></td>
<td>167</td>
</tr>
<tr>
<td></td>
<td>Dehghani Amirabad A.</td>
</tr>
<tr>
<td></td>
<td>53, 216</td>
</tr>
<tr>
<td></td>
<td>Dorri F.</td>
</tr>
<tr>
<td></td>
<td>190, 191</td>
</tr>
<tr>
<td></td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>Ebrahim-Habibi A.</td>
</tr>
<tr>
<td></td>
<td>206, 210</td>
</tr>
<tr>
<td></td>
<td>Ebrahimi A.</td>
</tr>
<tr>
<td></td>
<td>92, 96, 146, 148</td>
</tr>
<tr>
<td></td>
<td>Ebrahimie E...55, 65, 67, 68, 69, 70, 71, 132, 134, 138, 140, 159, 162, 163, 166, 169, 170</td>
</tr>
<tr>
<td></td>
<td>Eftekhar M.</td>
</tr>
<tr>
<td></td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Elahi E.</td>
</tr>
<tr>
<td></td>
<td>161, 191</td>
</tr>
<tr>
<td></td>
<td>Emadi M.</td>
</tr>
<tr>
<td></td>
<td>109</td>
</tr>
<tr>
<td>Authors</td>
<td>Pages</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Habibi M.</td>
<td>50, 193</td>
</tr>
<tr>
<td>Erfani Moghaddam M.</td>
<td>128</td>
</tr>
<tr>
<td>Erfanitabar F.</td>
<td>151</td>
</tr>
<tr>
<td>Eslahchi C.</td>
<td>37, 51, 52</td>
</tr>
<tr>
<td>Esfandiari H.</td>
<td>197</td>
</tr>
<tr>
<td>Esmaeilizadeh A.</td>
<td>180</td>
</tr>
<tr>
<td>Esmaeilizadeh M.</td>
<td>72</td>
</tr>
<tr>
<td>Esmaeilbeiki R.</td>
<td>35</td>
</tr>
<tr>
<td>Esmaeilizadeh A.</td>
<td>49, 64</td>
</tr>
<tr>
<td>F</td>
<td></td>
</tr>
<tr>
<td>Fallah Mehrabadi J.</td>
<td>111, 113, 114, 185</td>
</tr>
<tr>
<td>Farajzadeh D.</td>
<td>200</td>
</tr>
<tr>
<td>Farazmand A.</td>
<td>120</td>
</tr>
<tr>
<td>Farnia P.</td>
<td>195</td>
</tr>
<tr>
<td>Farsi M.</td>
<td>83, 126, 187</td>
</tr>
<tr>
<td>Fazeli Z.</td>
<td>103</td>
</tr>
<tr>
<td>Feizi A.</td>
<td>38</td>
</tr>
<tr>
<td>Ferdosi M. H.</td>
<td>64</td>
</tr>
<tr>
<td>Ferdosiyani M.</td>
<td>81</td>
</tr>
<tr>
<td>Forootan M.</td>
<td>201</td>
</tr>
<tr>
<td>Froughmand Araab M. H.</td>
<td>217</td>
</tr>
<tr>
<td>G</td>
<td></td>
</tr>
<tr>
<td>Gaderian M.</td>
<td>44</td>
</tr>
<tr>
<td>Ganjtabesh M.</td>
<td>31, 32</td>
</tr>
<tr>
<td>Gardaneh M.</td>
<td>105</td>
</tr>
<tr>
<td>Ghaderi M.</td>
<td>179</td>
</tr>
<tr>
<td>Ghaderi-Zeferehi M.</td>
<td>177</td>
</tr>
<tr>
<td>Ghaderi-Zeferehi M.</td>
<td>47</td>
</tr>
<tr>
<td>Ghaedi K.</td>
<td>97, 100, 107, 109</td>
</tr>
<tr>
<td>Ghaeminejad S.</td>
<td>50, 193</td>
</tr>
<tr>
<td>Ghaffarizadeh A.</td>
<td>49</td>
</tr>
<tr>
<td>Ghahremani M. H.</td>
<td>214</td>
</tr>
<tr>
<td>Ghahremanpour M. M.</td>
<td>128</td>
</tr>
<tr>
<td>Gharaei A.</td>
<td>94</td>
</tr>
<tr>
<td>Ghasem Aghaei N.</td>
<td>46, 61</td>
</tr>
<tr>
<td>Ghavami M.</td>
<td>150</td>
</tr>
<tr>
<td>Ghiasvand F. A.</td>
<td>56</td>
</tr>
<tr>
<td>Gifani P.</td>
<td>59, 63, 76</td>
</tr>
<tr>
<td>Goliæi B.</td>
<td>57, 204, 209, 217</td>
</tr>
<tr>
<td>Gomari H.</td>
<td>215</td>
</tr>
<tr>
<td>H</td>
<td></td>
</tr>
<tr>
<td>Habibi E.</td>
<td>133</td>
</tr>
<tr>
<td>Habibi M.</td>
<td>37, 52</td>
</tr>
<tr>
<td>Haddad R.</td>
<td>186</td>
</tr>
<tr>
<td>Hafezparast M.</td>
<td>43</td>
</tr>
<tr>
<td>Haghnavazi A.</td>
<td>43, 44, 102</td>
</tr>
<tr>
<td>Haghpahan V.</td>
<td>210</td>
</tr>
<tr>
<td>Hahsemi M.</td>
<td>109</td>
</tr>
<tr>
<td>Hashemifar S.</td>
<td>48</td>
</tr>
<tr>
<td>Hassani M.</td>
<td>123</td>
</tr>
<tr>
<td>Hassanzadeh R.</td>
<td>37</td>
</tr>
<tr>
<td>Heidari M.</td>
<td>101, 110</td>
</tr>
<tr>
<td>Heidari P.</td>
<td>132</td>
</tr>
<tr>
<td>Heidari S.</td>
<td>90</td>
</tr>
<tr>
<td>Heidari Sh.</td>
<td>132</td>
</tr>
<tr>
<td>Højati Z.</td>
<td>90</td>
</tr>
<tr>
<td>Højati Z.</td>
<td>107</td>
</tr>
<tr>
<td>Hoseinpourfard M.</td>
<td>84</td>
</tr>
<tr>
<td>Hosseini E. S.</td>
<td>151</td>
</tr>
<tr>
<td>Housaindokht M. R.</td>
<td>199</td>
</tr>
<tr>
<td>I</td>
<td></td>
</tr>
<tr>
<td>Ilkhanizadeh Sh.</td>
<td>93</td>
</tr>
<tr>
<td>Izadpanah K.</td>
<td>178</td>
</tr>
<tr>
<td>J</td>
<td></td>
</tr>
<tr>
<td>Jafarian M.</td>
<td>195</td>
</tr>
<tr>
<td>Jalal R.</td>
<td>203</td>
</tr>
<tr>
<td>Jamshidiha M.</td>
<td>214</td>
</tr>
<tr>
<td>Jedi-Tehrani M.</td>
<td>211</td>
</tr>
<tr>
<td>K</td>
<td></td>
</tr>
<tr>
<td>Kabiri M.</td>
<td>182</td>
</tr>
<tr>
<td>Kadkhodai S.</td>
<td>179</td>
</tr>
<tr>
<td>Kakhkeshani N.</td>
<td>42</td>
</tr>
<tr>
<td>Kahrizi D.</td>
<td>153, 154, 172, 194</td>
</tr>
<tr>
<td>Kalbassi M. R.</td>
<td>95, 188</td>
</tr>
<tr>
<td>Karami A.</td>
<td>59, 63, 75, 76</td>
</tr>
<tr>
<td>Karbalaei Kh.</td>
<td>97</td>
</tr>
<tr>
<td>Karegar M.</td>
<td>195</td>
</tr>
<tr>
<td>Karimi A.</td>
<td>65</td>
</tr>
<tr>
<td>Karimi J.</td>
<td>123</td>
</tr>
<tr>
<td>Kashani M.</td>
<td>66</td>
</tr>
<tr>
<td>Kashf B.</td>
<td>198</td>
</tr>
<tr>
<td>Kashfi R.</td>
<td>134, 163</td>
</tr>
<tr>
<td>Kashiri H.</td>
<td>125</td>
</tr>
<tr>
<td>Katanforoush A.</td>
<td>64, 161</td>
</tr>
<tr>
<td>Kavousi H. R.</td>
<td>83, 187</td>
</tr>
<tr>
<td>Kazemitabar S. K.</td>
<td>158</td>
</tr>
</tbody>
</table>
Sardari S.
Samadi Kafil H.
Saeedi Jam M.
Saeedi A
Sadeghi M.
Sabouri H.
Sabbaghi A.
Sabaghian M.
Sabaghian E.
S
Roodbar Shojaei T.
Rohani Nejad H.
Rezaei M.
Reissmann M.
Razzazi M. R.
Razi H.
Razavi H.
Ravash R.
Razavi H.
Rahpayma Sarvestani N.
Rajabalipour Cheshne Gaz A. A.
Rahpayma Sarvestani N.
Rahmani H.
Rahmani H.
Rahimi R.
Rahimi R.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rajabalipour Cheshne Gaz H.
Rajabalipour Cheshne Gaz H.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Razavi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rajabalipour Cheshne Gaz H.
Rajabalipour Cheshne Gaz H.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Rahmani H.
Rajabalipour Cheshne Gaz A. A.
Rahmani H.
Rahmani H.
Rahimi R.
Rahimi R.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rajabalipour Cheshne Gaz H.
Rajabalipour Cheshne Gaz H.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rajabalipour Cheshne Gaz H.
Rajabalipour Cheshne Gaz H.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rajabalipour Cheshne Gaz H.
Rajabalipour Cheshne Gaz H.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rahmani H.
Rajabalipour Cheshne Gaz A. A.
Rahmani H.
Rahmani H.
Rahimi R.
Rahimi R.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rahmani H.
Rajabalipour Cheshne Gaz A. A.
Rahmani H.
Rahmani H.
Rahimi R.
Rahimi R.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rahmani H.
Rajabalipour Cheshne Gaz A. A.
Rahmani H.
Rahmani H.
Rahimi R.
Rahimi R.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rahmani H.
Rajabalipour Cheshne Gaz A. A.
Rahmani H.
Rahmani H.
Rahimi R.
Rahimi R.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rahmani H.
Rajabalipour Cheshne Gaz A. A.
Rahmani H.
Rahmani H.
Rahimi R.
Rahimi R.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rahmani H.
Rajabalipour Cheshne Gaz A. A.
Rahmani H.
Rahmani H.
Rahimi R.
Rahimi R.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rahmani H.
Rajabalipour Cheshne Gaz A. A.
Rahmani H.
Rahmani H.
Rahimi R.
Rahimi R.
Rashidi S.
Ravaei H.
Ravash R.
Razi H.
Ravash R.
Razi H.
Razzazi M. R.
Reissmann M.
Rezaei M.
Rahmani H.
Rajabalipour Cheshne Gaz A. A.
Rahmani H.
V

Vaez Torshizi R. ............................................. 173
Vaezpour S. Y. .................................................. 82
Vahedian Movahed H. ...................................... 156
Vahedian-Movahed H. ...................................... 136
Valiellahi E. .................................................... 83, 126, 127, 187
Valiellahi M. ..................................................... 187
Vallian S. ......................................................... 97, 103, 117, 118, 174
Vatandoost J. .................................................... 104
Velayati A. A. ................................................... 195
Vera M. .......................................................... 95

W

Westhead D. ...................................................... 131
Wong L. .......................................................... 52

Y

Yaghmaei P. ...................................................... 100
Yakhchali B. ...................................................... 120, 200
Yakhchali B. ...................................................... 181

Z

Zakizadeh S. ...................................................... 91
Zamansani F. .................................................... 159
Zare-Mirakabad F. ................................. 33, 43, 44, 48, 54, 129
Zarrabi M. ...................................................... 151, 152, 186
Zebarjadi A. ...................................................... 172
Zeinali S. ........................................................ 181
Zolfaghari M. R. .............................................. 114
Zomorodipour A. R. ........................................ 104