

ORIGINAL PAPERS

*Genome analysis*  
**Method for identifying transcription factor binding sites in yeast** 1675  
 H.-K.Tsai, G.T.-W.Huang, M.-Y.Chou, H.H.-S.Lu and W.-H.Li

**Combining multiple microarrays in the presence of controlling variables** 1682  
 T.Park, S.-G.Yi, Y.K.Shin and S.Y.Lee

*Sequence analysis*  
**Ribosomal RNA as molecular barcodes: a simple correlation analysis without sequence alignment** 1690  
 K.H.Chu, C.P.Li and J.Qi

*Structural bioinformatics*  
**PROXIMO—a new docking algorithm to model protein complexes using data from radical probe mass spectrometry (RP-MS)** 1702  
 S.K.Gerega and K.M.Downard

**Molecular complexes at a glance: automated generation of two-dimensional complex diagrams** 1710  
 K.Stierand, P.C.Maaß and M.Rarey

**Ensemble classifier for protein fold pattern recognition** 1717  
 H.-B.Shen and K.-C.Chou

**SCARNA: fast and accurate structural alignment of RNA sequences by matching fixed-length stem fragments** 1723  
 Y.Tabei, K.Tsuda, T.Kin and K.Asai

*Gene expression*  
**How accurately can we control the FDR in analyzing microarray data?** 1730  
 S.-H.Jung and W.Jang

**Context-specific infinite mixtures for clustering gene expression profiles across diverse microarray dataset** 1737  
 X.Liu, S.Sivaganesan, K.Y.Yeung, J.Guo, R.E.Bumgarner and M.Medvedovic

**A Mixture model with random-effects components for clustering correlated gene-expression profiles** 1745  
 S.K.Ng, G.J.McLachlan, K.Wang, L.B.-T.Jones and S.-W.Ng

**A probabilistic dynamical model for quantitative inference of the regulatory mechanism of transcription** 1753  
 G.Sanguinetti, M.Rattray and N.D.Lawrence

**Identification of humoral immune responses in protein microarrays using DNA microarray data analysis techniques** 1760  
 S.Sundaresh, D.L.Doolan, S.Hirst, Y.Mu, B.Unal, D.H.Davies, P.L.Felgner and P.Baldi

*Systems biology*  
**Observing local and global properties of metabolic pathways: ‘load points’ and ‘choke points’ in the metabolic networks** 1767  
 S.A.Rahman and D.Schomburg

**Synchronizing a multicellular system by external input: an artificial control strategy** 1775  
 R.Wang, L.Chen and K.Aihara

APPLICATIONS NOTE

*Genome analysis*  
**Combo: a whole genome comparative browser** 1782  
 R.Engels, T.Yu, C.Burge, J.P.Mesirov, D.DeCaprio and J.E.Galagan

*Sequence analysis*  
**TSSub: eukaryotic protein subcellular localization by extracting features from profiles** 1784  
 J.Guo and Y.Lin

*Phylogenetics*  
**HoSeqI: automated homologous sequence identification in gene family databases** 1786  
 A.-M.Arigen, G.Perrière and M.Gouy

**TRFMA: a web-based tool for terminal restriction fragment length polymorphism analysis based on molecular weight** 1788  
 Y.Nakano, T.Takeshita and Y.Yamashita

*Structural bioinformatics*  
**TopoICE-R: 3D visualization modeling the topology of DNA recombination** 1790  
 I.K.Darcy and R.G.Scharein

**FCP: functional coverage of the proteome by structures** 1792  
 R.García-Serna, L.Opatowski and J.Mestres

**SHARP<sup>2</sup>: protein–protein interaction predictions using patch analysis** 1794  
 Y.Murakami and S.Jones

**Analysis assistant for single-molecule force spectroscopy data on membrane proteins—MPTV** 1796  
 F.Mueller, D.J.Muller and D.Labudde

**Prelude&Fugue, predicting local protein structure, early folding regions and structural weaknesses** 1800  
 J.M.Kwasigroch and M.Rooman

**BDT: an easy-to-use front-end application for automation of massive docking tasks and complex docking strategies with AutoDock** 1803  
 M.Vaqué, A.Arola, C.Aliagas and G.Pujadas

*Systems biology*  
**BIOCHAM: an environment for modeling biological systems and formalizing experimental knowledge** 1805  
 L.Calzone, F.Fages and S.Soliman

**ISMB 2006 CONFERENCE PROCEEDINGS, AUGUST 6–10, FORTALEZA, BRAZIL**  
 These are online only papers and are freely available as part of Issue 14 online

**EDITORIAL**

**ISMB 2006** e1  
 P.E.Bourne and S.Brunak

**ISMB 2006 Organization** e3

**ORIGINAL PAPERS**

<b>Automatic clustering of orthologs and inparalogs shared by multiple proteomes</b> A.Alexeyenko, I.Tamas, G.Liu and E.L.L.Sonnhammer	e9	<b>MotifCut: regulatory motifs finding with maximum density subgraphs</b> E.Fratkin, B.T.Naughton, D.L.Brutlag and S.Batzoglou	e150
<b>DynaPred: A structure and sequence based method for the prediction of MHC class I binding peptide sequences and conformations</b> I.Antes, S.W.I.Siu and T.Lengauer	e16	<b>Bistable switching and excitable behaviour in the activation of Src at mitosis</b> H.Fuß, W.Dubitzky, S.Downes and M.J.Kurth	e158
<b>ProfilePSTMM: capturing tree-structure motifs in carbohydrate sugar chains</b> K.F.Aoki-Kinoshita, N.Ueda, H.Mamitsuka and M.Kanehisa	e25	<b>Context-specific independence mixture modeling for positional weight matrices</b> B.Georgi and A.Schliep	e166
<b>The iRMSD: a local measure of sequence alignment accuracy using structural information</b> F.Armougom, S.Moretti, V.Keduas and C.Notredame	e35	<b>Improved Pruning algorithms and Divide-and-Conquer strategies for Dead-End Elimination, with application to protein design</b> I.Georgiev, R.H.Lilien and B.R.Donald	e174
<b>A model-based approach for mining membrane protein crystallization trials</b> S.Asur, P.Raman, M.E.Otey and S.Parthasarathy	e40	<b>Predicting the prognosis of breast cancer by integrating clinical and microarray data with Bayesian networks</b> O.Gevaert, F.De Smet, D.Timmerman, Y.Moreau and B.De Moor	e184
<b>Integrating structured biological data by Kernel Maximum Mean Discrepancy</b> K.M.Borgwardt, A.Gretton, M.J.Rasch, H.-P.Kriegel, B.Schölkopf and A.J.Smola	e49	<b>ZPRED: Predicting the distance to the membrane center for residues in <math>\alpha</math>-helical membrane proteins</b> E.Granseth, H.Viklund and A.Elofsson	e191
<b>ACIAP, Autonomous hierarchical agglomerative Cluster Analysis based protocol to partition conformational datasets</b> G.Bottegoni, W.Rocchia, M.Recanatini and A.Cavalli	e58	<b>Hairpins in a Haystack: recognizing microRNA precursors in comparative genomics data</b> J.Hertel and P.F.Stadler	e197
<b>A top-level ontology of functions and its application in the Open Biomedical Ontologies</b> P.Burek, R.Hoehndorf, F.Loebe, J.Visagie, H.Herre and J.Kelso	e66	<b>Modelling sequential protein folding under kinetic control</b> F.P.E.Huard, C.M.Deane and G.R.Wood	e203
<b>Comparative footprinting of DNA-binding proteins</b> B.Contreras-Moreira and J.Collado-Vides	e74	<b>BNTagger: improved tagging SNP selection using Bayesian networks</b> P.H.Lee and H.Shatkay	e211
<b>A probabilistic approach to protein backbone tracing in electron density maps</b> F.DiMaio, J.Shavlik and G.N.Phillips	e81	<b>Finding the evidence for protein-protein interactions from PubMed abstracts</b> H.Jang, J.Lim, J.-H.Lim, S.-J.Park, K.-C.Lee and S.-H.Park	e220
<b>CONTRAFold: RNA secondary structure prediction without physics-based models</b> C.B.Do, D.A.Woods and S.Batzoglou	e90	<b>Learning MHC I—peptide binding</b> N.Jojic, M.Reyes-Gomez, D.Heckerman, C.Kadie and O.Schueler-Furman	e227
<b>springScape: visualisation of microarray and contextual bioinformatic data using spring embedding and an ‘information landscape’</b> T.M.D.Ebbels, B.F.Buxton and D.T.Jones	e99	<b>Comparative genomics reveals unusually long motifs in mammalian genomes</b> N.C.Jones and P.A.Pevzner	e236
<b>Analysis of sample set enrichment scores: assaying the enrichment of sets of genes for individual samples in genome-wide expression profiles</b> E.Edelman, A.Porrello, J.Guinney, B.Balakumaran, A.Bild, P.G.Febbo and S.Mukherjee	e108	<b>Distance based algorithms for small biomolecule classification and structural similarity search</b> E.Karakoc, A.Cherkasov and S.C.Sahinalp	e243
<b>Dense subgraph computation via stochastic search: application to detect transcriptional modules</b> L.Everett, L.-S.Wang and S.Hannenhalli	e117	<b>Rapid knot detection and application to protein structure prediction</b> F.Khatib, M.T.Weirauch and C.A.Rohl	e252
<b>Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle</b> A.Fauré, A.Naldi, C.Chaouiya and D.Thieffry	e124	<b>Annotating proteins by mining protein interaction networks</b> M.Kirac, G.Ozsoyoglu and J.Yang	e260
<b>Semi-supervised LC/MS alignment for differential proteomics</b> B.Fischer, J.Grossmann, V.Roth, W.Gruissem, S.Baginsky and J.M.Buhmann	e132	<b>A decompositional approach to parameter estimation in pathway modeling: a case study of the Akt and MAPK pathways and their crosstalk</b> G.Koh, H.F.C.Teong, M.-V.Clément, D.Hsu and P.S.Thiagarajan	e271
<b>Statistical mechanical modeling of genome-wide transcription factor occupancy data by MatrixREDUCE</b> B.C.Foat, A.V.Morozov and H.J.Bussemaker	e141	<b>Finding novel genes in bacterial communities isolated from the environment</b> L.Krause, N.N.Diaz, D.Bartels, R.A.Edwards, A.Pühler, F.Rohwer, F.Meyer and J.Stoye	e281
		<b>A combinatorial pattern discovery approach for the prediction of membrane dipping (re-entrant) loops</b> G.Lasso, J.F.Antoniw and J.G.L.Mullins	e290

<b>Interpreting anonymous DNA samples from mass disasters—probabilistic forensic inference using genetic markers</b>	e298	<b>Decoding non-unique oligonucleotide hybridization experiments of targets related by a phylogenetic tree</b>	e424
T.Lin, E.W.Myers and E.P.Xing		A.Schliep and S.Rahmann	
<b>Peptide sequence tag-based blind identification of post-translational modifications with point process model</b>	e307	<b>Integrating copy number polymorphisms into array CGH analysis using a robust HMM</b>	e431
C.Liu, B.Yan, Y.Song, Y.Xu and L.Cai		S.P.Shah, X.Xuan, R.J.DeLeeuw, M.Khojasteh, W.L.Lam, R.Ng and K.P.Murphy	
<b>Identifying cycling genes by combining sequence homology and expression data</b>	e314	<b>Relative contributions of structural designability and functional diversity in molecular evolution of duplicates</b>	e440
Y.Lu, R.Rosenfeld and Z.Bar-Joseph		B.E.Shakhnovich	
<b>Quantification of transcription factor expression from Arabidopsis images</b>	e323	<b>Integrating image data into biomedical text categorization</b>	e446
D.L.Mace, J.-Y.Lee, R.W.Twigg, J.Colinas, P.N.Benfey and U.Ohler		H.Shatkay, N.Chen and D.Blostein	
<b>Mutation parameters from DNA sequence data using graph theoretic measures on lineage trees</b>	e332	<b>On counting position weight matrix matches in a sequence, with application to discriminative motif finding</b>	e454
R.Magori-Cohen, Y.Louzoun and S.H.Kleinstejn		S.Sinha	
<b>Indel seeds for homology search</b>	e341	<b>An ontology for a Robot Scientist</b>	e464
D.Mak, Y.Gelfand and G.Benson		L.N.Soldatova, A.Clare, A.Sparkes and R.D.King	
<b>Efficient identification of DNA hybridization partners in a sequence database</b>	e350	<b>ARTS: accurate recognition of transcription starts in human</b>	e472
T.P.Mann and W.S.Noble		S.Sonnenburg, A.Zien and G.Rätsch	
<b>An experimental metagenome data management and analysis system</b>	e359	<b>A computational approach toward label-free protein quantification using predicted peptide detectability</b>	e481
V.M.Markowitz, N.Ivanova, K.Palaniappan, E.Szeto, F.Korzeniewski, A.Lykidis, I.Anderson, K.Mavrommatis, V.Kunin, H.G.Martin, I.Dubchak, P.Hugenholtz and N.C.Kyrpides		H.Tang, R.J.Arnold, P.Alves, Z.Xun, D.E.Clemmer, M.V.Novotny, J.P.Reilly and P.Radivojac	
<b>An equilibrium partitioning model connecting gene expression and cis-motif content</b>	e368	<b>An integrative approach for causal gene identification and gene regulatory pathway inference</b>	e489
J.Mellor and C.DeLisi		Z.Tu, L.Wang, M.N.Arbeitman, T.Chen and F.Sun	
<b>Identification of metabolic units induced by environmental signals</b>	e375	<b>Computational inference of the molecular logic for synaptic connectivity in <i>C. elegans</i></b>	e497
J.C.Nacher, J.-M.Schwartz, M.Kanehisa and T.Akutsu		V.Varadan, D.M.Miller III and D.Anastassiou	
<b>Informative priors based on transcription factor structural class improve <i>de novo</i> motif discovery</b>	e384	<b>Novel Unsupervised Feature Filtering of Biological Data</b>	e507
L.Narlikar, R.Gordán, U.Ohler and A.J.Hartemink		R.Varshavsky, A.Gottlieb, M.Linial and D.Horn	
<b>Apples to apples: improving the performance of motif finders and their significance analysis in the Twilight Zone</b>	e393	<b>Constructing Near-Perfect Phylogenies with multiple homoplasy events</b>	e514
P.Ng, N.Nagarajan, N.Jones and U.Keich		R.V.Satya, A.Mukherjee, G.Alexe, L.Parida and G.Bhanot	
<b>Create and assess protein networks through molecular characteristics of individual proteins</b>	e402	<b>SNP Function Portal: a web database for exploring the function implication of SNP alleles</b>	e523
Y.Ofran, G.Yachdav, E.Mozes, T.Soong, R.Nair and B.Rost		P.Wang, M.Dai, W.Xuan, R.C.McEachin, A.U.Jackson, L.J.Scott, B.Athey, S.J.Watson and F.Meng	
<b>BaCellLo: a balanced subcellular localization predictor</b>	e408	<b>Protein classification using ontology classification</b>	e530
A.Pierleoni, P.L.Martelli, P.Fariselli and R.Casadio		K.Wolstencroft, P.Lord, L.Taberner, A.Brass and R.Stevens	
<b>Semi-supervised analysis of gene expression profiles for lineage-specific development in the <i>Caenorhabditis elegans</i> embryo</b>	e417	<b>Inferring Functional Pathways from Multi-Perturbation Data</b>	e539
Y.Qi, P.E.Missiuoro, A.Kapoor, C.P.Hunter, T.S.Jaakkola, D.K.Gifford and H.Ge		N.Yosef, A.Kaufman and E.Ruppin	
		<b>Accessing bioscience images from abstract sentences</b>	e547
		H.Yu and M.Lee	
		<b>A sequence-based filtering method for ncRNA identification and its application to searching for riboswitch elements</b>	e557
		S.Zhang, I.Borovok, Y.Aharonowitz, R.Sharan and V.Bafna	
		<b>Author Index</b>	e567