

1. Record Nr.	UNISA996601563203316
Autore	Ma Jian
Titolo	Research in Computational Molecular Biology : 28th Annual International Conference, RECOMB 2024, Cambridge, MA, USA, April 29-May 2, 2024, Proceedings
Pubbl/distr/stampa	Cham : , : Springer, , 2024 ©2024
ISBN	1-0716-3989-7
Edizione	[1st ed.]
Descrizione fisica	1 online resource (508 pages)
Collana	Lecture Notes in Computer Science Series ; ; v.14758
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Preface -- Organization -- Contents -- Fast Approximate IsoRank for Scalable Global Alignment of Biological Networks -- 1 Introduction -- 2 Algorithm -- 2.1 IsoRank -- 2.2 Approximate IsoRank -- 2.3 Computational Complexity -- 3 Experiments -- 3.1 Synthetic Networks -- 3.2 Biological Networks -- 3.3 Competing Methods -- 4 Discussion -- 4.1 Synthetic Networks -- 4.2 Biological Networks -- References -- Sequential Optimal Experimental Design of Perturbation Screens Guided by Multi-modal Priors -- 1 Introduction -- 2 Background -- 3 Method -- 3.1 Sequential Design of Perturb-Seq Experiment -- 3.2 Data-Driven Motivation for Incorporating Prior Knowledge -- 3.3 IterPert: A Multi-modal Prior-Guided Active Learning Strategy -- 4 Experiment -- 5 Related Works -- 6 Discussion -- A Data Processing on Multi-modal Priors -- B Fusion Operator -- C Error Bars for Baselines -- D Baseline Performance for Genome-Scale Perturbation Screen -- References -- Efficient Analysis of Annotation Colocalization Accounting for Genomic Contexts -- 1 Introduction -- 2 Methods -- 2.1 A Generative Model -- 2.2 The Context-Aware Markov Chain Null Model -- 2.3 Computing the Mean and Variance of the Overlap and Shared Bases Test Statistics -- 2.4 Mean and Variance of Any Separable Statistic -- 2.5 Multiple Chromosomes -- 3 Experiments -- 4 Discussion -- References -- Secure Federated Boolean Count Queries Using Fully-Homomorphic Cryptography -- 1 Introduction -- 2

Methods -- 2.1 Background -- 2.2 Union Cardinality -- 2.3
 Intersections Through Sampling -- 3 Results -- 3.1 Runtime
 Benchmarks -- 3.2 Accuracy Analysis -- 3.3 Security Analysis -- 4
 Discussion -- References -- FragXsiteDTI: Revealing Responsible
 Segments in Drug-Target Interaction with Transformer-Driven
 Interpretation -- 1 Introduction -- 2 Related Works -- 3 Methodology
 -- 3.1 Data Preparation.
 3.2 Feature Extraction -- 3.3 Classifier Module -- 4 Experiments -- 4.1
 Datasets -- 4.2 Implementation and Evaluation -- 4.3 Comparison on
 Target Datasets -- 4.4 DrugBank -- 4.5 Interpretation -- 5 Conclusion
 -- References -- An Integer Programming Framework for Identifying
 Stable Components in Asynchronous Boolean Networks -- 1
 Introduction -- 2 Preliminaries -- 2.1 Boolean Networks -- 2.2 Linear
 Threshold Functions -- 2.3 Network Dynamics -- 3 ILP Framework for
 Finding Quasi-Attractors -- 3.1 Lg Representation -- 3.2 Quasi-
 attractor Detection -- 3.3 External Nodes -- 3.4 Implementation
 Details -- 4 Results -- 4.1 Synthetic Networks -- 4.2 Real Biological
 Networks -- 5 Conclusions -- References -- ImputeCC Enhances
 Integrative Hi-C-Based Metagenomic Binning Through Constrained
 Random-Walk-Based Imputation -- 1 Introduction -- 2 Results -- 2.1
 Overview of ImputeCC -- 2.2 ImputeCC Achieved Accurate
 Preclustering for Contigs Containing Single-Copy Marker Genes -- 2.3
 ImputeCC Retrieved the Most High-Quality Genomes from the Mock
 metaHi-C Datasets -- 2.4 ImputeCC Markedly Outperformed Existing
 Bidders on Real metaHi-C Datasets -- 2.5 Running Time Analysis of the
 ImputeCC -- 3 Materials and Methods -- 3.1 Datasets -- 3.2 Data
 Preprocessing -- 3.3 The Framework of ImputeCC Binning -- 3.4
 Evaluating the Quality of Recovered MAGs from the Mock and Real
 metaHi-C Datasets -- 3.5 MAG Analyses on Real metaHi-C Datasets --
 3.6 Other Bidders Used in Benchmarking -- 4 Discussions --
 References -- Graph-Based Genome Inference from Hi-C Data -- 1
 Introduction -- 2 Inferring the Sample Genome from Hi-C Data with
 Genome Graphs -- 2.1 Problem Definition of Genome Inference -- 2.2
 The Hardness of the Problem -- 2.3 Computation of the Φ Function --
 2.4 Graph-Based Dynamic Programming Algorithm -- 2.5 Heuristics for
 Computing q -- 2.6 Accuracy of the Heuristic Algorithm.
 2.7 Practical Improvements to Efficiency and Accuracy -- 3
 Experimental Results -- 3.1 Construction of a Genome Graph with Hi-C
 Reads Mapped -- 3.2 Graph Hi-C Workflow Improves TAD Identification
 -- 4 Discussion -- References -- Meta-colored Compacted de Bruijn
 Graphs -- 1 Introduction -- 2 Preliminaries: Modular Indexing of
 Colored Compacted de Bruijn Graphs -- 3 Meta-colored Compacted de
 Bruijn Graphs -- 3.1 Definition -- 3.2 Data Structures Used and Two-
 Level Intersection Algorithm -- 3.3 The Optimization Problem -- 3.4
 The SCPO framework -- 4 Experiments -- 5 Conclusions -- References
 -- Color Coding for the Fragment-Based Docking, Design and
 Equilibrium Statistics of Protein-Binding ssRNAs -- 1 Introduction -- 2
 Method and Algorithms -- 2.1 Ensuring Self-avoidance Through Color
 Coding -- 2.2 Reducing Clashes Through Monochromatic Clique Covers
 -- 2.3 Rational ssRNA Design as a Relaxation of Docking -- 2.4
 Equilibrium Statistics -- 3 Results -- 3.1 Stability Analysis -- 3.2
 Impact of Monochromatic Clique Covers -- 3.3 Docking Through
 Energy-Minimization Under Different Fragment Definitions -- 3.4
 Design -- 4 Conclusions and Perspectives -- References -- Automated
 Design of Efficient Search Schemes for Lossless Approximate Pattern
 Matching -- 1 Introduction -- 2 Preliminaries -- 3 A Greedy Heuristic
 for Improving Search Schemes -- 4 Integer Linear Program Formulation
 -- 5 Dynamic Selection -- 6 Experiments and Results -- 6.1 Dataset

and Computational Environment -- 6.2 Better Search Schemes -- 6.3 Application to Lossless Read Mapping -- 7 Conclusion -- References -- CELL-E: A Text-to-Image Transformer for Protein Image Prediction -- 1 Introduction -- 2 Results -- 2.1 The CELL-E Model -- 2.2 Performance Evaluation -- 2.3 Analysis of NLS Using CELL-E -- 3 Discussion -- 4 Methods -- 4.1 Model Specifics -- 4.2 Nucleus Image Encoder. 4.3 Protein Threshold Image Encoder -- 4.4 Amino Acid Embedding -- 4.5 CELL-E Transformer -- 4.6 Probability Density Maps -- References -- A Scalable Optimization Algorithm for Solving the Beltway and Turnpike Problems with Uncertain Measurements -- 1 Introduction -- 2 Method -- 2.1 Problem Setting -- 2.2 Minorization-Maximization Scheme for Solving Turnpike -- 2.3 Extension to Other Variants -- 2.4 Initializer Sampling -- 3 Empirical Results -- 4 Conclusion -- References -- Overcoming Observation Bias for Cancer Progression Modeling -- 1 Introduction -- 2 Methods -- 2.1 Classical MHNs with Unaffected Observation -- 2.2 MHNs with Effects on Observation -- 2.3 Non-identifiability and Regularization -- 3 Results -- 3.1 Colon Adenocarcinoma -- 3.2 Lung Adenocarcinoma -- 4 Discussion -- References -- Inferring Metabolic States from Single Cell Transcriptomic Data via Geometric Deep Learning -- 1 Introduction -- 2 Problem Setup -- 3 Related Work -- 4 GEFMAP: Gene Expression-Based Flux Mapping and Metabolic Pathway Prediction -- 4.1 Metabolic Network Graph Generation -- 4.2 Inferring the Objective Function -- 4.3 Null Space Network for Solving the Objective Function -- 5 Experiments -- 5.1 Experimental Setup -- 5.2 Validating Our Objective Function on Core E. Coli Network -- 5.3 Learning FBA Solution Flux Estimations -- 5.4 Human Embryoid Networks -- 6 Conclusions and Future Work -- A Supplemental Information -- A.1 Supplemental Methods -- References -- Computing Robust Optimal Factories in Metabolic Reaction Networks -- 1 Introduction -- 2 Computing Optimal Parameter-Free Factories -- 2.1 Factories and Hypergraphs -- 2.2 Parameter-Free Shortest Factories -- 3 Characterizing Optimal Factories -- 3.1 The Structure of Shortest Factories -- 3.2 Hyperpaths Are Factories -- 3.3 Guaranteeing Nondegeneracy -- 4 Experimental Results -- 4.1 Experimental Setup. 4.2 Freeia Finds Factories Missed by the Prior State-of-the-Art -- 4.3 Speed of Computing Parameter-Free Factories -- 5 Conclusion -- References -- Undesignable RNA Structure Identification via Rival Structure Generation and Structure Decomposition -- 1 Introduction -- 2 RNA Design -- 2.1 Secondary Structure, Loop and Free Energy -- 2.2 MFE and Structure Distance -- 3 Undesignability -- 4 Theorems and Algorithms for Undesignability -- 4.1 Algorithm 0: Exhaustive Search -- 4.2 Theorem 1 and Algorithm 1: Identify One Rival Structure -- 4.3 Theorem 2 and Algorithm 2: Identify Multiple Rival Structures -- 4.4 Theorem 3 and Algorithm 3: Structure Decomposition -- 5 Experiments on Eterna100 Dataset -- 5.1 Setting -- 5.2 Results -- 5.3 Insights -- 6 Conclusions and Future Work -- References -- Structure- and Function-Aware Substitution Matrices via Learnable Graph Matching -- 1 Introduction -- 1.1 Contributions -- 2 Preliminaries -- 2.1 Graph Neural Networks -- 2.2 Graph Edit Distance -- 3 Methods -- 3.1 GSM Architecture -- 3.2 Training GSM -- 4 Experimental Evaluation -- 4.1 Datasets -- 4.2 Experimental Setup -- 4.3 Similarity-Based Classification -- 4.4 Retrieval -- 5 Analysis of the Edit Cost Matrices -- 6 Discussion and Conclusions -- References -- Secure Discovery of Genetic Relatives Across Large-Scale and Distributed Genomic Datasets -- 1 Methods -- 2 Results -- References -- GFETM: Genome Foundation-Based Embedded Topic Model for scATAC-seq Modeling --

1 Introduction -- 2 Methods -- 2.1 The ETM Component -- 2.2 The GFM Component -- 2.3 Leveraging the Peak Embedding from GFM in ETM -- 2.4 Transfer Learning of GFETM -- 3 Results -- References --

SEM: Size-Based Expectation Maximization for Characterizing Nucleosome Positions and Subtypes -- 1 Introduction -- 2 Methods -- 3 Results -- 4 Conclusion -- References.

Centrifuger: Lossless Compression of Microbial Genomes for Efficient and Accurate Metagenomic Sequence Classification.
