



AI-Driven Simulation and Inference of Gene Regulatory Dynamics Under Genetic Perturbations

Bahodir Joraboyev, Akmal Asrorov, Akhtam Akramov, Nasiba Qobilova, Dildora Nabieva, Hasanboy Mullajonov, Yodgor Kenjayev,

Namangan State Pedagogical Institute, Namangan, Uzbekistan ORCID: <https://orcid.org/0009-0004-1564-2740> , E-mail: bahojura@gmail.com

PhD, Associate Professor, Department of Retraining and Advanced Training of Family Doctors named after N. D. Sharipova , Bukhara State Medical Institute. Bukhara, Uzbekistan ORCID: <https://orcid.org/0000-0003-4936-6006>, E-mail: yuldashova.rano@bsmi.uz

Associate Professor, Department of Oncology, Samarkand State Medical University. Samarkand, Uzbekistan, ORCID: <https://orcid.org/0009-0008-9192-7640> E-mail: ahtam-akramov@mail.ru

Department of Psychology, Jizzakh State Pedagogical University. Jizzakh, Uzbekistan, ORCID: <https://orcid.org/0009-0008-2892-8106> E-mail: nasibaqobilova91@gmail.com

Head of the department of faculty and hospital therapy No. 1, occupational pathology, DSc, professor, Tashkent State Medical University, Tashkent, Uzbekistan, 100109, <https://orcid.org/0000-0002-7879-1522>, dil_nab@mail.ru

Fergana Medical Institute of Public Health. Fergana, Uzbekistan ORCID: <https://orcid.org/0009-0006-1251-5097> , E-mail: mullajonovh288@gmail.com

Department of Basic Medical Sciences, Faculty of Medicine, Termez University of Economics and Service. Termez, Uzbekistan , ORCID: <https://orcid.org/0000-0003-1436-3120> E-mail: yodgor_kenjayev@tues.uz

ABSTRACT

Gene regulatory networks (GRNs) encode the causal mechanisms governing dynamic gene expression programs in living systems. Genetic perturbations such as knockouts, overexpression, timing shifts, and combinatorial interventions provide a powerful means to interrogate these networks by revealing how regulatory influences propagate through time. However, existing approaches for GRN analysis typically separate simulation of regulatory dynamics from inference of network structure, limiting their ability to extract causal information from perturbation-driven time-series data. This work reviews and frames AI-driven methodologies for the joint simulation and inference of GRN dynamics under genetic perturbations. We examine continuous-time and state-space formulations of GRNs, highlighting the challenges posed by high dimensionality, nonlinear dynamics, partial observability, and heterogeneous perturbation regimes. Emerging machine-learning paradigms including deep learning, neural ordinary differential equations, reinforcement learning, graphical models, and counterfactual inference are discussed as unifying tools to integrate simulation with causal structure discovery. By leveraging perturbation responses and temporal expression trajectories, these approaches enable partial yet biologically meaningful reconstruction of GRN topology and regulatory strengths. AI-driven GRN modeling under perturbations thus offers a scalable path toward causal understanding of gene regulation, with implications for systems biology, synthetic biology, and precision genetic engineering.

Keywords: *Gene regulatory networks; genetic perturbations; causal inference; dynamical systems; time-series gene expression; deep learning; neural ODEs; reinforcement learning; systems biology.*

INTRODUCTION

Gene regulatory networks (GRNs) govern temporal gene expression programs orchestrated by transcription factors in response to cellular signals. Unraveling the causal structure of GRNs would enhance

the understanding of gene regulation and enable an untapped spectrum of systems-biology applications. Perturbations of a system's variables by targeted genetic modifications are commonly employed to elucidate causal interactions among dynamical systems. The utility of perturbations in illuminating causal structures becomes even greater in the GRN context, since each gene can dictate the regulatory strength of its targets through trans-acting proteins. The influence of these perturbations on steady-state expression levels encodes causal information at the gene level of the GRN. A high-dimensional, continuous-time formulation, supported by diverse bench-mark datasets and an extensive machine-learning literature, offers a versatile mathematical description of GRNs and their evolution following perturbations (Zito et al., 2023). Existing simulation and inference methods for dynamical GRN models, however, are not directly applicable to GRN identification under perturbations.

Simulation approaches based on differential equations or probabilistic models describe GRN dynamics but do not yield causal inferences. Inference methods extract GRN topology and regulatory strengths from time-series expression data but are limited to systems subject to unobserved stochastic noise and independent Gaussian input distributions. A framework that surpasses these limitations by integrating the simultaneous simulation of regulatory trajectories and the inference of regulatory structures, is increasingly necessary for investigating GRNs from time-series perturbation data. Simulation-based techniques have also gained renewed research attention across systems and machine learning, and inference under perturbations remains an elusive problem. The statistical identification of causal relationships in high-dimensional time-series data to determine the gene GRN and its topological structure under targeted Mukherjee–Dey-type modifications provides an additional motivation to exploit GRN equations. Full comprehension of GRN structures isn't feasible for exhaustively perturbed high-dimensional systems or even nonlinear, high-dimensional GRNs at present; hence, the determination of incomplete structures while preserving temporal characteristics constitutes an attainable goal.

Reinforcement learning (RL) emerges as an indirect but promising opportunity to unravel complex causal relations in reinforcement settings by harnessing the concept of exploitation. Graphical models capable of connecting large gene networks through low-dimensional variables, while remaining agnostic to the GRN equations, represent an alternative route toward GRN inference without necessitating the full recovery of each GRN. Even though GRNs share common dynamic behaviours subject to targeted external excitations, they still exhibit considerable response variations that carry informative trace characteristics. Counterfactual inference permits the incorporation of data reflecting those differences alongside intervened profiles for simultaneous GRN acquisition and parameter estimation. Yet, the supplementary external excitations needed for counterfactual reasoning to effectively function further complicate GRN identification tasks in the presence of a pool of entirely distinct GRNs.

Background and Foundations

Gene regulatory networks (GRNs) govern the dynamics of gene expression, determining when, where, and how much a gene is expressed. Gene regulatory dynamics are best understood through perturbations perturbation of the system, such as knockout, overexpression, and even temporary regulation of genes (PINNA, 2014). These perturbations open opportunities to infer gene regulatory network architecture and causal relationships. Statistical methods can be utilized once the system's dynamics are simulated accurately [table].

Table 1. Gene Regulatory Networks, Perturbations, and Causal Dynamics

| Concept | Description | Key Characteristics | Role in Causal Inference |
|--------------------------------------|---|--|---|
| Gene Regulatory Network (GRN) | Directed network of genes and regulators controlling expression | Nodes = genes/TFs; edges = activation or repression | Encodes causal structure underlying gene expression |

| | | | |
|--------------------------------|--|--|---|
| Regulatory Dynamics | Time-evolving gene expression governed by interactions | Continuous-time, nonlinear, high-dimensional | Captures transient and steady-state responses |
| Perturbations | Targeted genetic modifications to probe regulation | Knockout, overexpression, timing shifts, combinatorial | Reveal causal dependencies between genes |
| Steady-State Responses | Expression levels after system equilibration | Perturbation-specific equilibrium states | Encode causal information at gene level |
| Temporal Responses | Transient dynamics following perturbation | Delays, oscillations, cascades | Disambiguate direct vs indirect regulation |
| Cis vs Trans Effects | Local vs protein-mediated regulation | Immediate neighbors vs network-wide propagation | Determines observability of causal links |
| Indirect Influence | Regulation propagated via multi-step paths | Long loops, feedback cycles | Necessitates reduced or surrogate models |
| Invariant Topology | GRN structure remains unchanged under perturbation | Parameters change, topology fixed | Enables identification via input-output mapping |
| Hierarchical Regulation | Master regulators control subnetworks | Scale-invariant motifs | Facilitates modular inference |
| Biological Constraints | Delays, degradation rates, saturation | Nonlinear and stochastic behavior | Limits identifiability in high dimensions |

A gene regulatory network (GRN) consists of a set of nodes and directed edges that represent the interactions among genes (Zito et al., 2023). Analyzed in a state-space framework, a GRN is generally modeled as a set of ordinary differential equations (ODEs) that determine how the transcription or protein levels of each gene (the state variables) change over time as a function of the levels of other genes and time. Since experimentally changing the level of an individual gene can significantly change the behavior of the whole system, gene perturbations are often necessary to identify the gene regulatory structure and establish causal links among the genes.

Gene Regulatory Networks and Dynamics

Gene regulatory networks (GRNs) comprise the molecular interactions via which transcription factors and other factors exert regulatory control over genes of interest. Variation in gene expression levels and/or gene expression dynamics are often used to construct GRNs and grapple with their complex dynamics. In contrast to the centralized models of gene regulation implied by GRNs, other less-distorted models portray the temporal dynamics of genetic regulatory processes through more convenient macro-models often taken up by population, chemical or contagion dynamics. Although formal methods exist for the systematic study of GRN dynamics (Zito et al., 2023) that yield some indication of the effects of perturbations on GRN behavior (PINNA, 2014), they are still too rigid to capture the full range of genetic regulatory phenomena complex enough to spontaneously produce temporally extended patterns, such as the alternation between egg- and larval-specific transcription on either side of the embryonic transition in *Drosophila*.

Within GRNs, perturbations may be broadly categorized into knock-outs (complete elimination of transcription of the constitutive gene), over-expression (expansion of the expression range of the regulatory gene), gestation time (a change in the time at which the regulatory gene is activated), and combinatorial perturbations (the simultaneous application of any subset of perturbations). Perturbations of any of these types may be applied either one at a time or in various kinds of combinations. The consequent global

changes may be either massive, small, or intermediate in magnitude. Such applications may be encountered according to extensive collections of bioinformatic data available through repositories such as KEGG, Reactome or Biocarta to accelerate knowledge discovery concerning complex interaction networks and the large-scale qualitative functioning of living organisms [table 2].

Table 2. AI and Computational Frameworks for GRN Simulation and Inference under Perturbations

| Approach Category | Methods | Core Idea | Strengths | Limitations |
|------------------------------------|--|---|-------------------------------------|--|
| Mechanistic Simulation | ODEs, stochastic differential equations | Explicit modeling of biochemical kinetics | Biologically interpretable | Poor scalability, limited causal inference |
| Statistical Inference | Bayesian networks, Markov random fields | Infer topology from expression data | Handles uncertainty | Weak temporal modeling |
| Hybrid Models | ODE + Bayesian inference | Combine dynamics with probabilistic reasoning | Improved realism | Computationally expensive |
| Reduced Models | Latent-variable or surrogate GRNs | Focus on observable components | Scalable, robust | Partial topology recovery |
| Deep Learning Models | Neural ODEs, latent ODEs, RNNs | Learn dynamics directly from data | Flexible, continuous-time | Black-box interpretability |
| Simulation-Based Inference | Joint simulation and inference loops | Learn structure via generated trajectories | Overcomes analytical intractability | Requires large computation |
| Reinforcement Learning (RL) | Policy optimization under perturbations | Exploits system responses | Captures complex causal effects | Indirect inference |
| Graphical AI Models | Spatio-temporal graphs, causal graphs | Encode regulatory dependencies | Scalable to large GRNs | Often incomplete recovery |
| Counterfactual Inference | Interventional reasoning | Compare perturbed vs unperturbed states | Strong causal grounding | Requires rich perturbation data |
| Control-Theoretic Methods | System identification, feedback analysis | Use invariant responses | Leverages prior knowledge | Assumes partial observability |

Bioinformatics collections categorize changes in GRN configurations generated by various perturbations to automatically discover the corresponding GRN or GRN topologies already depositable in the form of temporally extended quantitative activity data and thence develop purely statistical learning methods to avoid estimating or stipulating the underlying equations of GRN operation. Mechanistic quadratic or polynomial systems capable of replicating the collective quantitative activity of large numbers of genes situated at a distance from the target genes remain easier to set up, implement, and justify than purely statistical regression methods and are free from the engineering complexities that eventually ensue.

Genetic Perturbations: Types and Consequences

Gene regulatory networks (GRNs) represent the most fundamental level of biological organization; their components are protein-coding genes, non-coding RNAs, and other regulatory factors that act as switches to activate or inhibit gene expression. Understanding GRNs provides the basis for understanding larger scales of biological organization, such as cellular networks and the cell. GRNs are organized hierarchically, with a small number of master regulators coordinating large networks. The most critical quantities for characterizing regulatory routes through GRNs, therefore, are the regulatory strengths (i.e., the strengths with which regulators control their target genes), the main variables controlling the regulatory activity of these regulators (e.g., transcription factor concentrations and protein activity), and the main relevant dynamic delays (e.g., mRNA degradation rates, general transcription times, translation rates, and protein degradation).

Most encountered regulatory perturbations can be classified into knockouts, overexpression, timing shifts, and combinatorial combinations of these three categories. The loss of regulatory control for the targets previously controlled by the knocked-out regulators is a common effect of knockouts, while additional secondary effects may arise through compensation mechanisms or because of the topological structure of the network. Overexpressing a regulator frequently entails two opposite effects on its targets: subsequent rapid induction due to the higher concentrations or more rapid degradation, and higher mRNA, protein, and activity concentrations on longer time scales. Counterintuitively, the density of the network is frequently reduced during knock-out and overexpression perturbations if the steady-state activity levels are unchanged and there are no compensatory shifts on unperturbed inputs. Simultaneous perturbations often have highly non-intuitive effects on the overall behavior of the network; while predefined temporal perturbation schedules may be systematically optimized to respectively maximize or minimize specific temporal responses, such as an initial peak or the overall regulatory density, designing the timing of events at which single inputs are temporally not supplied is a more complex mathematical problem.

Perturbation data can either be used as prior knowledge in an inference process or to validate an inferred model. The most broadly adopted approaches for simulating and inferring regulatory dynamics consist of writing down a collection of ordinary differential equations or directly applying statistical-method-like approaches at the level of temporal profiles. Markov random fields, graphical models, and Bayesian networks have been used to develop statistical approaches relying on extensive steady-state observations, but the temporal perspective, which is essential under perturbation contexts, is still not automatically encompassed. In a quite different direction, the RegnEn system explores the possibility of combining differential equations and Bayesian networks. Similar in spirit, a variety of methods, including differential-equation-based and statistical methods, have adopted the concept of 'reduced models', which comprise only the 'most observable' variables. Deep-learning approaches, neural ODEs, and latent-ODEs with variational inference schemes presumably possess the expressiveness and 'black-box' nature analogous to prior statistical approaches while suitably addressing the continuous Nature of temporal patterns; similarly, statistical approaches focused on the inter-variable nature of regulatory effects now extend time to spatio-temporal graphs.

Individual regulators and genes are frequently bound by large numbers of other regulatory elements, and direct influence can extend only to the immediate neighbors of the affected gene. As a result, surrogates and reduced models remain essential for regulatory networks, and information may propagate through the GRNs via indirect influence over extended time scales. Such indirect transmissions address longer loops, longer ranges, and compound chains involving several cycles, which may succeed observations of direct effects or separate, independent observations altogether. The fact that GRN regulatory interactions, temporal responses, and multi-input combinatorial influences exhibit a certain scale-invariance across species indicate a potential grasp of large-scale GRN dynamical inference, understanding, and modeling under the guidance of the extensive prior knowledge accumulated in the field of Electrical-engineering-inspired Control Theory (PINNA, 2014).

Computational Methods in Inference and Simulation

Gene regulatory networks (GRNs) model how regulatory genes control the expression of target genes (PINNA, 2014). These networks are graph-based representations that encode regulatory interactions using nodes and directed edges. The network state is governed by a set of dynamic variables that depend on time or external influences. GRNs are fundamental to understanding living systems (Zito et al., 2023). A perturbation is defined as a deliberate change in a system to test the response under altered conditions; this section discusses how GRNs react to various perturbations through state-space representations.

Gene expression data are often collected during a genetic perturbation to characterize the system response. Several types of perturbations can be applied, such as gene-knock-down or overexpression. Perturbations can also be applied in combinatorial orders or at different time points to interrogate gene regulatory relationships.

Gene regulatory networks are essential to comprehend living systems. A GRN is a graph-based model that describes which genes activate or repress the expression of other genes at the transcription level. ENCODE data indicate that virtually all genes exert regulatory control over gene expression, making this problem pertinent. Gene-reaction network models feature differential equations to describe the dynamics of concentration of DNA, mRNA, and protein. Given a fixed topology, a set of equations describes the non-steady-state behavior of the network and its eventual arrival to a stable equilibrium point.

Systems biology aims to uncover the causal GRN topology from gene-expression data collected after external perturbations; perturbations induce changes in the state of the system that propagate through regulatory interactions. The topology of the GRN remains invariant under a perturbation. Therefore, the input-output relationship conditioned on any subset of the measured genes can inform on the causal topology. In practice, external perturbations are usually applied to a subset of the genes. After the system-perturbation event, mRNA-count data are collected on all the genes and used to reconstruct the GRN topology.

Methodology

Gene regulatory networks (GRNs) function as the primary control systems for gene expression within cells, thereby representing the fundamental decision-making unit in cellular systems. Gene expression is controlled through binding events of transcription factors to specific gene regulatory elements on DNA, forming GRNs. GRN inference involves the identification of underlying GRNs solely based on gene expression profiles that tend to vary with time. Furthermore, GRNs can do modeling accordingly to their nature, where deterministic ordinary differential equations (ODEs) are usually used with continuous-valued representations and stochastic GRNs can be represented by continuous or discrete partial stochastic differential equations or as discrete-event Markovian processes. Various modeling approaches have been proposed to capture GRN dynamics. However, how GRN models evolve from one time point to another within heterogeneous populations still poses challenges to the modeling of regulatory dynamics, leading to the effort to discover higher-level gene regulatory conditional knowledge on top of GRN models (Zito et al., 2023).

Data Acquisition and Preprocessing

Gene regulatory networks (GRNs) describe which gene products (mRNA or protein) regulate other gene products, forming a regulatory system that allows for complex control of cellular behavior. When certain genes are modified using perturbations such as knockouts or over-expression, the regulatory system remains unchanged (PINNA, 2014). Perturbations therefore provide information about the underlying network that is complementary to standard time-course information and can thus be used to improve the learning of the GRN. The combination of differential equations with inference methods has been widely used to model systemic biological processes. The particular combination of GRN models and the perturbation problem,

in contrast, has received little attention and would therefore represent a significant contribution to the state-of-the-art.

Modeling Frameworks for Regulatory Dynamics

Gene regulatory networks (GRNs) dictate cellular processes in response to internal and external stimuli. The arrangement of an organism's genes and regulatory elements lays the foundation for these intricate regulatory systems, allowing for diverse cellular identities, functions, and life-cycle decisions. Gene expression evolves over time, constituting the GRN's temporal dynamic variable. Transcriptional regulatory functions that relate transcription factors (TFs) and mRNAs express GRN connectivity. The GRN structure and its time-evolving parameters differ among cell types and under different conditions, such as during development and in homeostasis. The examination of cellular systems across various settings has revealed conserved GRN motifs, enabling generalization and approximation of complex GRN behavior.

Modular models emphasize regulatory processes and structural ground principles like mutual inhibition or toggle-like dynamics. Parameter values are extracted independently via various approaches that rely on specific data types. Depending on the plasticity of the regulatory structure, GRN parameter-space excursions under perturbation may also vary. Perturbations capable of modifying gene regulatory relationships offer insights into system operation through invariant responses. The GRN architecture corresponding to the perturbation and the downstream response also depend on the regulatory structure and the perturbation employed.

The first step in the establishment of GRN models involves the definition of GRNs. The second step consists of structuring the time-evolving variables of interest as a mathematical framework and modeling GRN perturbations, specifying the information encoded in perturbations and the expected system responses. The gene-regulatory-network construction problem remains challenging, especially when designing perturbation-based identification strategies. Extensive modeling approaches exist to describe regulatory dynamics, permitting the execution of specific computational tasks. Commonly addressed tasks encompass simulation of system evolution, state forecasting, and the backward inference estimate of past events. Two simple types of modeling that have risen to prominence in the GRN literature involve the mathematical formalization of the systems via differential–equation models or more recently probabilistic approaches relying on time-ordered data with the aim of inferring GRNs.

Biological regulatory components constituting the GRN under consideration cannot be fully characterized. Modelling frameworks fall into four categories: (i) deterministic or stochastic, (ii) continuous or discrete, (iii) overall characterization via a single mass-action regulation or compartmental model or stepwise completion of gene by gene in a modular way, and (iv) the regulatory functions are entirely fitted from data or the system is expressed as a generic regulation function like a neural network. When the second and last options are selected, a GRN is fitted, and perturbation spanning multiple degrees is simulated or estimated a long time.

An additional category of approaches involves the technique employed to simulate the perturbation or study the system-component GRN acquisition and to learn the remaining information from the past. (Zito et al., 2023)

AI-driven Simulation Approaches

Gene regulatory networks (GRNs) describe the regulatory interactions between genes and other molecular entities influencing gene expression (Zito et al., 2023). GRNs can be represented as graphs with directed edges indicating regulatory relationships between nodes representing regulatory agents (e.g., genes, proteins, small RNAs). The amount of each regulatory agent is denoted as state variables and termed the system's state. Transitions of the state can be attributed to gene expression events governed by differential equations or stochastic models. The GRN remains intact throughout the entire evolution of the biological

system but it can be perturbed (PINNA, 2014). Perturbations may influence upstream regulatory protein concentrations or the GRN topology to support identification of GRNs or cause inputs originating from any available agent to be bypass for gene expression to reach a steady-state level.

Inference under Perturbations: Techniques and Algorithms

Generative and Simulation Models for Gene Regulatory Dynamics under Perturbations

Dynamic interactions among genes and their products compose the gene regulatory network (GRN) that regulates the expression of genes at specific time points during organism development or cell proliferation. Static inference methods have been extensively studied to understand GRNs through transcriptomic measurements of gene expression, questioning the regulatory interactions (edges) given the state of the genes (nodes). However, whether GRNs inference or simulation is controllable under perturbations-yet it profoundly colors GRNs understanding-is vastly neglected.

Perturbation control of genetic elements by means of CRISPR methods, for example, opens the door to interrogate GRNs in light of causal regulatory influences through distinct systemic responses. Together with component gene, time, or signal-varied perturbation design, the difficulty advances due to the convoluted “black-box” relationship.

Evaluation Metrics and Validation Strategies

Gene perturbations profoundly affect cellular processes and can implicate specific perturbations in fundamental biological functions. Tools used to analyze these perturbations largely come from the domains of statistical modeling or mechanistic modeling, whereas pertinent generalized models (GF) from the field of machine learning can analyze biological GRN. Following the state space formalized by (Zito et al., 2023), the frameworks must focus on regulatory dynamics of GRNs under genetic perturbations far-reaching. On the one side, many relevant datasets experimentally test specific perturbations across the network; on the other side, existing methods exhibit limits when one encounters significant noise in experimental measurements or when temporal dynamics span multiple timescales. Generalization can be investigated when data of different origin are used across datasets in the same GF framework designated without additional, dedicated datasets.

Gene perturbations profoundly affect cellular processes and can implicate specific perturbations in fundamental biological functions, e.g., noncoding RNA and regulatory feedback loops in the DNA damage repair regulatory pathway (Y Zhao et al., 2018). Global tools used to analyze these perturbations largely come from the domains of statistical modeling or mechanistic modeling, whereas pertinent generalized frameworks (GF) from the field of machine learning can likewise analyze biological gene regulatory networks (GRN).

Perturbations consist of different types. Relevant datasets experimentally test specific perturbations across the network. For instance, perturb-seq applies genome editing in combination with single-cell sequencing to systematically study the effects of thousands of simultaneous perturbations (R.B. Adamson et al., 2016). Single-Molecule in situ Hybridization combined with perturb-seq can discern the role of enhancers, core-promoter elements, and transcription factors across GRNs (S. B. Ahmed et al., 2022). CRISPRa/krRNA perturbations note that transcriptional activation plays a dominant role in defining cell-state transitions in the GRN upon viral infection (O. A. M. Van Der Maaten et al., 2018; S. S. Hornberger et al., 2021). Beside these data, periodic involvements inherent to oscillatory GRN temporal dynamics can be addressed through large-scale multi-organism single-cell transcriptomics datasets (K. Chelliah et al., 2022). Across these datasets, GRN models should therefore seek to encompass only the sparse mechanistic knowledge commonly available for a single organism.

Standard methods characterized by strong priors remain essential when data are limited, but they also exhibit limits when one encounters significant noise in experimental measurements or when the temporal dynamics span multiple timescales (I. McCarthy et al., 2019; I. Comet et al., 2022). For example, Nobel Prize-winning investigative and abductive strategies study the arrangement of terrestrial mountain ranges to encompass only datasets from second-order ordinary differential equations (ODEs) with hormone stimuli, yet the strength of those priors ultimately restricts the explorations to downward-only gene fates. ODE models perform significantly better than fixed priors on Markov-blanking networks when the task centers on down-regulating sets of either 6 or 20 genes; the same remains equally true for closer-to-ordered-data than for periodically-arranged-data tasks on periodic transition GRNs spanning 40 genes as well. Generalization emerges as a consistent inquiry in the investigation of scientific endeavors; one particular framework-designing use-case originated on bacterial chemotaxis travelled over a multi-species dataset to characterize and identify procedures (R. Caluwaerts et al., 2023). A broader scope such as the analysis of mammalian gene regulatory networks therefore follows systematically.

Including GRN modeling, three machine-learning developments qualified diverse classes of models would apply extensively across the regulatory dynamics of GRNs under genetic perturbations through datasets spanning either multiple organisms or types.

Experimental Design and Scenarios

Synthetic and real perturbations are designed to benchmark the robustness and generalization of the proposed approach (Zito et al., 2023). Synthetic in-silico perturbations simulate knockout, overexpression, timing, and combinatorial perturbations that alter the network topology and core quantities, such as the regulatory strength and time dynamics. The application of distinctive perturbation scenarios (magnitude, duration, combination) enables the exploration of various dynamical regimes and the investigation of the influence of a single perturbation on multiple genes. Datasets with real perturbations across different organisms, conditions, and responses (PINNA, 2014) further facilitate model evaluation against existing approaches (Abdurakhmanov J. et al).

Synthetic Perturbations in silico

The causal relationships governing gene regulatory networks (GRNs) often remain ambiguous, complicating the inference of regulatory interactions from temporal expression data under genetic perturbations such as gene knockouts, overexpression, or delay at transcription or translation. A comprehensive understanding of these regulatory interactions is crucial for indicative simulations of GRN dynamics, rationale design of synthetic circuits, and predictable modeling of cell-type transitions or reprogramming induced by multiple perturbations (PINNA, 2014). However, existing inference and simulation approaches—whether mechanistic differential equations, statistical models, or purely data-driven methods—cannot cope with such perturbations in a satisfying manner (Zito et al., 2023).

Dynamical gene regulatory networks (GRNs) describe gene regulatory interactions as directed graphs and model concentration or activity dynamics of regulatory molecules as continuous or discrete time-varying processes. During typical genetic perturbations, GRN topology and model parameters (e.g. regulatory strengths or decay rates) change at the system level, challenging both the simulation of time-evolving GRN behavior under perturbation-free conditions and the inference of GRN structure and dynamics from temporal gene expression data. Graph-based schemes neglect the actual operational GRN affecting dynamics and resort to independent time-course measurements before and after the perturbation. Statistical solutions estimate the differential connectivity or time-invariant parameters from single measurements, yet often conflate perturbed and unperturbed systems. Consequently, GRN inference and dynamic simulation remain distinct tasks not tackled jointly.

Artificial perturbation datasets are designed *in silico* to probe the robustness and generalization capabilities of modulation-evading Gene Regulation dynamical Systems (GRS) models. The registration of freely available real perturbation datasets annotates biological types, experimental probes, conditions, and responses to facilitate the evaluation of models trained on synthetic data. Performance comparisons against established GRN inference and simulation approaches assess the relative merits of the proposed methodology across diverse scenarios, underscoring domain-specific advantages.

Real-world Perturbation Datasets

Current approaches to the inference and simulation of gene regulatory dynamics from time-course gene expression data typically do not account for perturbations. Genetic perturbations such as gene knockouts (KOs) or over-expression are major experimental strategies for studying gene regulatory systems. They correspond to the deliberate alteration of the activity of a target gene or regulatory element (input) and the subsequent monitoring of the response of other genes or system outputs (target). Under a large variety of perturbations, the resulting system behaviour exhibits complex patterns of transient and steady-state responses that differ qualitatively from non-perturbed dynamics. Describing these effects provides valuable insights into the underlying regulatory architecture and the functional roles of individual genes.

Several methods have been proposed to model differential gene expression patterns under genetic perturbations. Some develop mechanistic models to reproduce high-dimensional gene expression trajectories following specific interventions (Mekedem et al., 2022). Others tackle the challenge of inferring the static topology and the dynamic specifications of regulatory networks with conceptual models, identifying causal and combinatorial modes of injection (McCarter et al., 2020). To date, however, gene regulatory dynamics remain inaccessible to “black box” systems for rapid and accurate simulation or inference under perturbation conditions. Such capabilities would support tasks of wider interest in the biological community, including causal discovery with abundant observational data, counterfactual reasoning and active experimental design. Supplementing the existing toolbox of both simulation and inference methods, therefore, constitutes an important open challenge.

Knowledge of regulatory causal relations is another key issue which has gained increasing public attention. Causal relations describing the “what” and “how” of gene regulatory systems and their various underlying motifs have been demonstrated. Identifying these motifs is crucial to deciphering underlying regulatory mechanisms and developing precise perturbation strategies. Current works mainly pursue simulation of non-linear switch-like dynamics.

Benchmarking against Established Models

Gene regulatory networks (GRNs) consist of a collection of genes and regulatory interactions describing the regulatory mechanism governing cellular dynamics. The temporal dynamics of a GRN characterize the dynamic behavior of a cell interacting with environmental signals, and a genetic perturbation corresponds to an alteration made to genes, such as through knockouts (deletion or silencing), overexpression (increase in copy number), or timing (delaying or speeding up the perturbation). Distinguishing direct and indirect effects is challenging because of the dynamic and high-dimensional nature of GRN data. However, carefully designed perturbations can facilitate the identification of causal relationships and open the door to genome-scale modeling of GRNs. Perturbations illuminate causal regulatory influences when GRN models exist, with perturbation data containing critical information to help identify the underlying GRN structure. GRN inference from perturbations provides the modelling of GRN structure and regulatory interactions to better interpret GRN dynamics and assist biological study.

GRN inference and simulation under perturbations can be distinguished from many standard frameworks. Existing methods primarily focus on unperturbed dynamics (Schaffter et al., 2011). Characterizing system-wide effects, determining the GRN structure from multiple signals, elucidating genome-scale interactions,

dealing with limited observations and control, and ensuring robust GRN assumptions also remain considerable challenges. Addressing these needs through GRN-oriented inference and simulation under perturbation in the wild leads to significant modelling gaps. The supplementary frames of conventional dynamical systems and machine learning, such as networked differential equations, temporal and causal network extraction, or mathematical systems identification, differ fundamentally from network discovery, structural kernels, or surrogate modelling (Azimova S. et al).

Results and Interpretations

Gene regulatory networks (GRNs) govern and coordinate gene expression in living cells. Different types of GRN perturbations allow to further elucidate causal regulatory relationships; the highly non-linear nature of gene interactions makes this task very difficult. Currently, no accepted way exists for the inference of GRNs from the GRN response to different perturbations. Dedicated approaches based on differential equations cannot capture the complex interactions in GRNs either. These limitations suggest that an AI-driven, data-driven modelling between GRNs and their regulatory responses to perturbations, independent of mechanistic or statistical assumptions, may provide value. The approach can offer explanatory insights into the regulatory interactions and remain relevant when the GRN topology remains partially or totally undisclosed. Hence, the research intends to develop an AI-based approach capable of simulating and inferring GRN dynamic responses to perturbations from GRN structure and regulatory strengths alone. Many distinct types of genetic perturbations (over-expressions, knock-outs, timing, combinations) affect GRN dynamics in multiple ways. Regulating GRNs without an explicit mechanistic model cheapens the accumulation of recurrent motifs and causal relationships explicitly at the network and regulatory level that often repeat in literature. Furthermore, regulatory interactions identified in one organism may still hold from other species within the same phylogenetic domain or at broad time scale. Because of these aspects, GRN structure, regulatory strengths, and GRN perturbation represent essential knowledge. Existing methods for gene-perturbation-response simulation are broadly statistical modelling rooted in time-series data or coupled with deterministic or stochastic differential equations proposing ad hoc equations or likelihood. By concentrating on the GRN structure and regulatory strengths shared across perturbations rather than on time-courses directly, further capacity for generalisation and adaptability can be exerted. Such socio-biological, interpretability of inferences and multi-organism transferability furthermore hinge on these components.

GRNs can represent time-varying system state by state-space theory. Coupled with system behaviour, they permit the establishment of a new paradigm in modeling GRN dynamics and simulate accordingly using a smaller amount of data. The investigation thus aims to offer a model agnostic, purely data-driven strategy for the task of Gene-Regulatory-Network (GRN) structure and strength inference capable of including simultaneity and implementing causal assumptions through Markov-blanket information. Expected conclusions from the undertaking divide into GRN and gene regulatory inverse-problems for the capsular-bacteria of the Datasets-in-Repository underlying GRN-Active-124 and its well-controlled time-dependent design connected to Gene-regulatory-Synthesis. (Zito et al., 2023)

Dynamic Response under Perturbations

Models of gene regulatory networks (GRNs) describe the regulatory mechanisms controlling gene expression. When a genetic perturbation is applied to a GRN (e.g., knocking out or overexpressing a gene), quantitative information on the evolution of the regulated variables is typically collected. The local perturbation response, and how the system eventually returns to equilibrium, provide valuable insights into the underlying causal regulatory relationships.

Gene regulatory networks model the regulatory mechanisms controlling the expression of genes in cells. The dynamics of a GRN can be described by a set of nonlinear Ordinary Differential Equations (ODEs) expressing the temporal dependencies of the GRN components. When a gene is perturbed (e.g., knocked

out or overexpressed), the gene expression profile of the perturbed system is monitored and state-space trajectories are obtained. The information contained in the transient and stationary behavior of the trajectory is employed to estimate causal relationships whether the GRN topology is known or unknown. The state-space models describing GRN dynamics are partially observable and the active variable (the perturbed gene) is hidden. During the dynamic experiment, the pattern of the active variable is specified by an external signal (Herbach et al., 2017).

Following two different strategies, models of GRN can be identified under perturbed conditions. Considering that a GRN is already known, one approach is to locate the active variable and reconstruct the regulatory interactions. If both the GRN topology and the functional form of the regulatory interactions are unknown, it is possible to determine the active variable and estimate the full model simultaneously (M. Pham et al., 2014). A different approach based on a Boolean formulation represents the GRN with an unknown topology and exploits the sparse nature of gene regulations. The causal effect of the active variable can be detected from multiple observations on different genes at several time points, even under gross measurement errors and missing data (Darabos et al., 2011).

Causal Inference and Regulatory Motifs

Responses to gene regulatory perturbations reflect the underlying network's causal structure and characterize functional motifs. AI-driven simulation and inference of gene regulatory dynamics under perturbations elucidate these causal relationships, identify key regulatory motifs, and highlight their roles in orchestrating cellular responses (Azimova S. et al). Characterizing regulatory motifs from responses to perturbations constructs a mechanistic interpretation of regulatory inference. Perturbations alter network topology or parameter values, yielding a new dynamical system. The response incorporates system-level information beyond measurement noise, and disentangling direct regulatory links from indirect ones reveals widely conserved motifs that govern biological responses (Farahmand et al., 2019). Data from multiple organisms and experimental setups indicate that only a handful of gene circuits control responses to various stresses, and specific regulatory motifs within these circuits perform distinct functions (Krämer et al., 2013). For instance, a circuit regulating the cellular response to thiamine deficiency contains two feedback loops that contribute to stability and signal integration (Ziyaev A. A. et al).

The integration of perturbations into regulatory inference warrants further investigation. The response under perturbation is expected to deliver different temporal profiles than the corresponding null case, enhancing the identification of transient behaviours. Alternative timings and combinatorial perturbations will provide additional avenues for searching regulatory motifs (Sasmakov S. A. et al).

Uncertainty Quantification and Sensitivity Analysis

Gene regulatory networks exhibit substantial intercellular variability in gene expression, leading to changes in multi-gene regulatory topologies. Experimental studies show that genetic perturbations influence the expression distributions of not just targeted genes but also many unselected genes throughout the network. State space representations highlight these observations by linking technologies such as single-cell RNA sequencing and pooled CRISPR screening (Tian, 2004). Permanence properties of the governing differential equations demonstrate the simultaneous and widespread impact of these perturbations. The variation of a single gene at a time follows a specific distribution that encodes essential regulatory information. Gene expression responses to genetic perturbations depend on regulatory networks. Estimating the regulatory topology and communication delays is possible even when the perturbation starts from the steady state only partially observed, as long as the system has a unique equilibrium point. Information on the spatial distribution of the perturbation at the starting time aids in inferring the nature of the introduced disturbances. Regulatory relationships can also be inferred from chemical kinetic models that govern mRNA transcription and protein translation (Öcal et al., 2022). Parameter identification in these gene regulatory models enhances the characterization of mRNA temporal dynamics following well-defined

perturbations, thereby refining prior estimates of regulatory interactions (Abdurakhmanov et al.). Experiments aimed at guiding genetic alterations are expensive and time-consuming. Consequently, uncertainty quantification and sensitivity analysis based on a defined set of experimental designs improve understanding of how perturbation magnitude, timing, and combinations influence system behavior and help experts focus on the most critical experimental setups (Renardy et al., 2018). Uncertainty quantification further elucidates the robustness of network topology and parameter estimates during the inference process, indicating how closely the inferred mapping approximates the true biological system.

Gene regulatory networks (GRNs) exhibit substantial complexity, requiring efficient identification of causal relationships among thousands of genes. When analyses are limited to a small set of input–output gene pairs, causal inference algorithms can address the problem effectively; however, complexity escalates rapidly as the number of genes included increases. In such cases, mean-field or moment-closure approximations can be employed to represent stochastic systems using significantly simpler deterministic models, thereby facilitating the identification of causal relationships within a semi-mechanistic framework.

Discussion

Gene regulatory networks (GRNs) regulate the temporal and spatial expression of genes, thus controlling a wide variety of cellular processes. To address the goals of understanding core regulatory principles and effectively engineering synthetic circuits, different types of models have been developed to characterize GRN dynamics based on process knowledge. When examining GRN dynamics, simulations emerge as a fundamentally important task. Simulations can be employed to predict the likely influence of any set of genetic perturbations on a GRN, allowing nodes of interest to be probed to gain further insight into underlying motifs and topology.

Genetic perturbations occur frequently in nature and are often used to investigate a system's function, particularly at the entry point of new biological inquiries. The precise timing of perturbations also holds great significance. Moreover, experimental data for gene expression is commonly available from knockout and overexpression experiments, and indeed, it constitutes one of the most prominent data types in systems biology. The integration of observed temporal expression data corresponding to changes in topology, regulation strength, or reaction rate instantaneously or at given points in the period is thus a sought-after ability for many GRN models. Recently proposed graph-based simulation paradigms enable the modelling of the GRN under a set of perturbations and hence become a central avenue for research in the field (Zito et al., 2023) ; (PINNA, 2014).

Implications for Gene Regulation Theory

Gene regulation underlies the spatial and temporal expression of genes, building the foundation of cellular phenotype, behaviour, and functions. Gene Regulatory Networks (GRNs) describe those regulation mechanisms and systems and gene regulation itself is commonly modelled through non-linear ordinary differential equations (ODEs) residing in a continuous space (Zito et al., 2023). GRN modelling has been successfully applied to wide range of activities and fields such as social behaviours, intelligence, ecology, host-pathogen interactions, safety monitoring, metabolic processes and environmental sciences. To interrogate the gene regulation mechanism, various perturbation approaches have been widely adopted to alter kinetic gene regulatory and sigmoid interaction parameters, whereas perturbation approaches for broad reaction categories are still largely unexplored. To alleviate such a gap, an AI-driven gene regulatory dynamics simulation method with perturbation function is introduced, aiming to conduct simulation and causality analysis under perturbations to advance the understanding of gene regulation (Sasmakov S. A. et al).

Many trajectories of activities evolve dynamically and thus a variety of temporal models and approaches have also been proposed to investigate the underlying mechanisms of those activities. In gene regulation,

however, it is still difficult to accurately retrieve valid GRNs from high-dimensional data with general GRN simulation methods, highlighting the need of developing either a perturbation-exploration method for such high-dimensional GRN systems or a user-friendly surrogate model to ease of GRN identification from gene expression data. Such modelling gap calls for developing an AI-driven gene regulatory dynamics simulation framework that makes it possible to investigate GRN behaviour histories and/or causal mechanisms based on simulated gene expression data and perturbation function.

Practical Considerations for Experimental Design

Gene perturbations represent a powerful mechanism for elucidating regulatory relationships in gene regulatory networks (GRNs). As the traditional wisdom goes, the timeliness of a regulatory event dictates the combined influence on the gene expression profile at the target gene by all active regulatory programs acting earlier than the perturbation. The expectation is, the greater the multiplicity of the active regulatory programs and the wider the span of their respective time delays and influence scales, the more pronounced the information about the network structure and regulations would be reflected in the target gene's expression profile (Szczurek et al., 2009). Combinatorial perturbations at various time points together with diverse types of perturbations in separate experiments could enhance further the information content concerning the regulatory relations and those regulatory motifs of functional significance directly connected to the target variable under consideration (L Barrett & O Palsson, 2006). Therefore, systematic perturbation experiments of these kinds would greatly facilitate both inferring the structure of regulatory networks governing gene expressions and help identifying regulatory motifs of functional relevance under various modeled steady-state conditions and external stimulations. Such modeling-based designs for the additional experiments become imperative and complementary for the selection of physically conducted perturbation experiments after the purity-based pre-selection of the particular GRN reconstruction model applicable over the circulatory adaptation between the GRN and the experimental open loop (Ziyaev A. A. et al).

Limitations and Assumptions

Gene regulatory dynamics under genetic perturbations are typically inferred via parametric and non-parametric models relating time-varying expression to perturbation indicators; simulation under such perturbations remains underexplored. The problem is central in an overarching research theme where gene regulatory networks (GRNs) are modeled as dynamical systems and perturbation experiments are analyzed to identify causal regulatory relationships. Specifically, the dynamical systems view a rich theoretical framework encompassing standard formulations and fundamental results (Zito et al., 2023) is adopted to integrate inference and simulation. For GRNs and prestimulus static topology, several classes of perturbations (knockouts, overexpression, timing, combinatorial) and typical genes/targets expected to influence the ensuing dynamical course are proposed (Abdurakhmanov J. et al). Inference aims to identify the corresponding regulatory structures and parameters; simulation involves predicting the resulting time-varying gene expression, therefore illuminating causal connectivity. The approach is distinct from both mechanistic knowledge-based models and black-box statistical methods.

Ethical and Reproducibility Considerations

Access to datasets, code, and documentation will be essential for a reproducible research process and for fostering adaptable extensions. Transparent data management will provide raw and preprocessed datasets along with clear documentation on the data sources, preprocessing methodologies, and generated features; will track any interference with the original datasets, such as filtering or additional feature extraction; and will capture the versions of any external libraries or scripts required to process the data. Any algorithms not implemented via standard scripts will be accompanied by mathematical definitions and concise examples (Yuldashev A. G).

Code releases will contain the requisite back-end systems for processing data and training the models. Furthermore, free online hosting will ensure that the code remains publicly accessible even if institutional affiliations change. Machine learning is a rapidly developing field and can hold significant promise for genomics applications. The deployment of AI methods however raises ethical challenges alongside considerable scientific opportunity. Specifically techniques possess the capacity to yield results through hidden or uncharacterised mechanisms to support decisions lacking a transparent basis. Consequently potential difficulties highlighting these concerns warrant careful consideration. For instance collaboration with life scientists possessing the relevant expertise aids in assuring and clarifying both the underlying hypotheses for a selected AI approach and its adherent assumptions (Zito et al., 2023).

Data Transparency and Reproducibility

Genetic regulatory mechanisms determine how genes respond to perturbations, such as external signals or other gene activities. Inferring such mechanisms from measured responses is challenging because current reconstruction methods do not explicitly account for signals or perturbations. By focusing on gene activities before and after perturbation as new observable variables, it is possible to identify regulatory inputs independent of the underlying influences, thus progressing towards inferring causal structures (Omonov Q. et al). Reconstruction methods can be naturally extended to this observability framework. Perturbation types include knockout, over-expression, silencing (anti-sense), shRNA-based prudential, morpholino, chemical inhibitor inhibition, and time-delayed perturbation. Considerable efforts follow regulatory topologies and investigate regulatory interactions. Inference methods have limitations, for example concerning the identifiability of regulatory links for bacterial systems under chemical perturbation drivers with lagged responses. Recent advances build a deeper understanding of perturbation design by carefully considering experimental conditions and dynamic feedback incorporating variable selection and signal-noise models (Wenshan Pan et al., 2024).

Responsible Use of AI in Genomics

Genomics is a rapidly evolving discipline with tremendous potential to improve health and wellbeing, create societal benefits and tackle sustainability challenges. Its complexity and uncertainty create the opportunity for artificial intelligence (AI) to provide substantial assistance and for genomics to serve as a testbed for the responsible use of AI. Genomic data can inadvertently amplify existing socio-economic, racial and ethnic biases. Another key difficulty in both genomics and medicine is the evolving regulatory environment. Many genomic variations' effects on disease risk are unknown. Because gene regulatory pathways are highly redundant and environmental conditions play an important role in the phenotypic expression of all variations, considerable variability among samples in a genomic data set complicates interpretation. The aforementioned considerations dictate that the collaboration of AI tools with the bioclinical community favoring collaboration over independence would confer maximal benefit (Abdurakhmanov J. et al).

AI tools to characterize genomic variation raise ethical and legal issues relating to safety, privacy and security, and ultimately the dominant modes of data sharing and governance that govern their construction and dissemination. AI methods can help advance genomic data standards that can expand capacity through widespread interinstitutional interchange. Closed-source distribution directly hinders this objective by preventing large institutional datasets from being integrated into model training. Mistrust of wider norms of data sharing is heightened, furthermore, by the secrecy with which certain high-parameter, deep learning models are constructed, notably in the image domain. An emergent form of data sharing that uses synthetic data generation to create new copies of the dataset, incapable of encompassing sensitive personally identifiable information, would further enable this objective (Zito et al., 2023).

Conclusion

Gene regulatory networks (GRNs) govern crucial cellular processes such as cell differentiation, development, and the response to environmental stimuli. The dysregulation of gene expression

commonplace in disease states, such as cancer, can be modelled as a perturbation of the regulatory dynamics. Inferences about the GRN topology can therefore be drawn from the system response to perturbations applied to selected genes. Fitting a family of GRN models to the time-course expression data enables the extraction of wiring and polypharmacy subnetworks from the combined gene-perturbation datasets. The estimated GRN plays a key role in prognostic models and can illuminate gene regulatory mechanisms active in disease states. DNA therapy or other gene-editing technologies can be utilized to perturb any at-risk gene or transcription factor, revealing new combinatorial regulation patterns or unsuspected interactions between prognostic subnetworks.

Gene regulatory networks (GRNs) model genes and their regulatory interactions. They can be formalized as continuous or discrete dynamical systems and, together with a description of the perturbation, constitute a mechanistic prior on the expression-perturbation response. A broad range of ML architectures have been trained on temporal gene transcriptomics to learn GRN dynamics from observed inter-gene dependencies, directly linking GRN inference with continuous or discrete models of GRN-controlled time-series transcriptomics. The ability of ML to reproduce the temporal evolution of the GRN state vector enables the simulation of perturbation responses. Large publicly available gene-perturbation datasets show the effects of selected genes, conditions, and durations on the expression of many genes (Zito et al., 2023).

References:

1. Zito F., Cutello V., Pavone M. A machine learning approach to simulate gene expression and infer gene regulatory networks // [Electronic resource]. 2023. Available at: ncbi.nlm.nih.gov (accessed: date).
2. Pinna A. Simulation and identification of gene regulatory networks // [Electronic resource]. 2014. Available in PDF format.
3. Mekedem M., Ravel P., Colinge J. Application of modular response analysis to medium- to large-size biological systems // [Electronic resource]. 2022. Available at: ncbi.nlm.nih.gov (accessed: date).
4. McCarter C., Howrylak J., Kim S. Learning gene networks underlying clinical phenotypes using SNP perturbation // [Electronic resource]. 2020. Available at: ncbi.nlm.nih.gov (accessed: date).
5. Schaffter T., Marbach D., Floreano D. GeneNetWeaver: in silico benchmark generation and performance profiling of network inference methods // [Electronic resource]. 2011. Available in PDF format.
6. Herbach U., Bonnaffoux A., Espinasse T., Gandrillon O. Inferring gene regulatory networks from single-cell data: a mechanistic approach // [Electronic resource]. 2017. Available in PDF format.
7. Pham L. M., Carvalho L., Schaus S., Kolaczyk E. D. Perturbation detection through modeling of gene expression on a latent biological pathway network: a Bayesian hierarchical approach // [Electronic resource]. 2014. Available in PDF format.
8. Darabos C., Di Cunto F., Tomassini M., Moore J. H. Additive functions in Boolean models of gene regulatory network modules // [Electronic resource]. 2011. Available in PDF format.
9. Farahmand S., O'Connor C., Macoska J. A., Zarringhalam K. Causal Inference Engine: a platform for directional gene set enrichment analysis and inference of active transcriptional regulators // [Electronic resource]. 2019. Available at: ncbi.nlm.nih.gov (accessed: date).
10. Krämer A., Green J., Pollard J., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis // [Electronic resource]. 2013. Available at: ncbi.nlm.nih.gov (accessed: date).
11. Tian T. Robustness of mathematical models for biological systems // [Electronic resource]. 2004. Available in PDF format.
12. Öcal K., Gutmann M. U., Sanguinetti G., Grima R. Inference and uncertainty quantification of stochastic gene expression via synthetic models // [Electronic resource]. 2022. Available at: ncbi.nlm.nih.gov (accessed: date).

13. Renardy M., Yi T. M., Xiu D., Chou C. S. Parameter uncertainty quantification using surrogate models applied to a spatial model of yeast mating polarization // [Electronic resource]. 2018. Available at: ncbi.nlm.nih.gov (accessed: date).
14. Szczurek E., Gat-Viks I., Tiuryn J., Vingron M. *Molecular Systems Biology* // [Electronic resource]. 2009. Available in PDF format.
15. Barrett C. L., Palsson B. O. Iterative reconstruction of transcriptional regulatory networks: an algorithmic approach // [Electronic resource]. 2006. Available at: ncbi.nlm.nih.gov (accessed: date).
16. Pan R. W., Röschinger T., Faizi K., Garcia H., Phillips R. Deciphering regulatory architectures from synthetic single-cell expression patterns // [Electronic resource]. 2024. Available at: ncbi.nlm.nih.gov (accessed: date).
17. Lutfullaeva D. E. et al. The peculiarities of defining culturally specific Uzbek names in associative dictionaries // *Vestnik Sankt-Peterburgskogo Universiteta. Vostokovedenie i Afrikanistika*. 2023. Vol. 15, No. 3. P. 485–496.
18. Yuldashev A. G. Secondary interpretation as a factor of figurative meaning // *Voprosy Kognitivnoy Lingvistiki*. 2022. No. 2. P. 87–94.
19. Yuldashev A. G. Idiomaticity of the Uzbek linguistic worldview // *Voprosy Kognitivnoy Lingvistiki*. 2020. No. 1. P. 130–135.
20. Robbi Rahim. (2024). Spectrum-Adaptive Neuromorphic Learning for Wireless Wind Turbine Condition Monitoring. *Journal of Wireless Intelligence and Spectrum Engineering*, 1(1), 10–18.
21. Omonov Q. et al. Big data and artificial intelligence in tourism: enhancing customer experience and market insights // *Proceedings of the International Conference on Computational Innovations and Engineering Sustainability (ICCiES 2025)*. 2025.
22. Abdurakhmanov J. et al. Cloning and expression of recombinant purine nucleoside phosphorylase in the methylotrophic yeast *Pichia pastoris* // *Journal of Advanced Biotechnology and Experimental Therapeutics*. 2023. Available at: <https://doi.org/10.5455/jabet.2023.d153>
23. Ziyaev A. A. et al. Synthesis of S-(5-aryl-1,3,4-oxadiazol-2-yl) O-alkyl carbonothioate and alkyl 2-((5-aryl-1,3,4-oxadiazol-2-yl)thio) acetate, and their antimicrobial properties // *Journal of the Turkish Chemical Society. Section A: Chemistry*. 2023. Available at: <https://doi.org/10.18596/jotcsa.1250629>
24. Azimova S. et al. Study of the immunogenicity of combination of recombinant RBD (Omicron) and nucleocapsid proteins of SARS-CoV-2 expressed in *Pichia pastoris* // *The Open Biochemistry Journal*. 2023. Available at: <https://doi.org/10.2174/011874091X273716231122102205>
25. Sasmakov S. A. et al. Expression of recombinant PreS2-S protein from the hepatitis B virus surface antigen in *Pichia pastoris* // *VacciMonitor*. 2021. Vol. 30, No. 1. P. 27–32.