Analysis of sources that support the EGA-Cellular-Reprogramming project.

Abstract
Buriticá et al. (2018) compare deterministic ordinary differential equation (ODE) models and stochastic Gillespie simulations to capture kinetic rates—activation, inactivation, and synthesis—in gene expression. Rodríguez Penas (2017) uses Differential Evolution and Scatter Search to estimate high-dimensional ODE parameters, explicitly enforcing biologically plausible bounds. John and Meza-Chaves (2019) apply genetic algorithms—with elitist selection, diversity control via Hamming distance, and a fitness based on Bayesian posterior probability—to infer gene regulatory interactions in an Arabidopsis system. These studies support an approach in which evolutionary methods optimize parameters in models of gene regulatory networks. They show that fitness functions based on estimation error or posterior probability, combined with constraints such as explicit parameter bounds, can yield simulations that approximate desired gene expression profiles. Together, they provide a cross-disciplinary basis for using computational optimization to calibrate differential equation models in systems biology.

Results
Characteristics of Included Studies

Study	Research Focus	Methodology	Key Parameters	Application Domain
Buriticá et al., 2018	Comparison of deterministic and stochastic simulation methods for gene expression	Petri net and Ordinary Differential Equation (ODE) modeling; Gillespie algorithm (first reaction and direct methods); no optimization method applied	Activation inactivation/synthe is rates (λ, μ, v); state variables: inactive gene, active gene, protein	Simple gene expression system; protein synthesis
Rodríguez Penas, 2017	Development of parallel metaheuristics for parameter estimation in systems biology	Differential Evolution (DE), Scatter Search (SS); asynchronous, cooperative, self-adaptive parallel algorithms; nonlinear ODE models	Model parameters for ODEs (number and type not specified; high-dimensional); parameter bounds for biological plausibility	Computational systems biology; parameter estimation in kinetic models; Mixed-Integer Dynamic Optimization (MIDO) problems
John & Meza-Chaves, 2019	Inference of gene interaction networks from time-series data using Genetic Algorithms (GAs)	Genetic Algorithms (traditional and Cross-generational elitist selection, Heterogeneous recombination, and Cataclysmic mutation (CHC)); Bayesian network (Directed Acyclic Graph (DAG)) modeling; fitness based on Bayesian posterior	Network structure (edges), gene interaction probabilities; Hamming distance for diversity; parameter bounds	Gene regulatory network (GRN) inference; Arabidopsis thaliana auxin response

Methodology:
• Deterministic simulation approaches: 2 studies used Ordinary Differential Equation (ODE) modeling.
• Stochastic simulation: 1 study used the Gillespie algorithm.
• Petri net modeling: 1 study used Petri nets.
• Metaheuristics: 1 study used Differential Evolution and Scatter Search.
• Parallel algorithms for parameter estimation: 1 study.
• Genetic Algorithms for network inference: 1 study.
• Bayesian network modeling: 1 study, with fitness based on Bayesian posterior.
• No optimization methods: 1 study.
• Other machine learning/statistical inference: No mention found in these studies.

Key Parameters:
• Kinetic rates and state variables: 1 study estimated these in a gene expression model.
• High-dimensional ODE parameters: 1 study estimated these with biologically plausible bounds.
• Parameter bounds for biological plausibility: 2 studies included these.
• Network structure and gene interaction probabilities: 1 study inferred these.
• Diversity metric (Hamming distance): 1 study used this in parameter search.

Application Domain:
• Simple gene expression system and protein synthesis: 1 study.
• Parameter estimation in kinetic models and MIDO problems: 1 study.
• Gene regulatory network inference in Arabidopsis thaliana: 1 study.
• Other biological systems or diseases: No mention found in these studies.

Thematic Analysis
Genetic Algorithm Optimization Approaches

Study	Parameter Optimization Strategy	Fitness Function Implementation	Biological Constraints Handling
Buriticá et al., 2018	Not applicable (no optimization method applied)	Not applicable	Not applicable
Rodríguez Penas, 2017	Differential Evolution and Scatter Search; asynchronous, cooperative, self-adaptive parallelization	Fitness based on parameter estimation error (details not specified); focus on solution quality and convergence	Parameter bounds for biological plausibility; adaptive strategies for robustness
John & Meza-Chaves, 2019	Traditional and CHC Genetic Algorithms; population of Directed Acyclic Graphs (DAGs); Hamming distance for diversity	Fitness: Bayesian posterior probability of model fit to data; multi-run consistency	Parameter bounds; mutation only when diversity is lost; avoids crossover between similar individuals

• Parameter optimization strategies:Reported in 2 of 3 studies; both used evolutionary algorithms (Differential
Evolution, Scatter Search, or Genetic Algorithms). One study used asynchronous, cooperative, self-adaptive
parallelization, and another used diversity strategies such as Hamming distance.
• Fitness function implementation:Described in 2 studies: one used parameter estimation error, and one used
Bayesian posterior probability of model fit to data. No mention found of the fitness function in 1 study.
• Biological constraints handling:Addressed in 2 studies, both using parameter bounds for biological plausibil-
ity. One study also used adaptive strategies for robustness, and another used diversity strategies (mutation
only when diversity is lost, avoiding crossover between similar individuals). No mention found of biological
constraints handling in 1 study.

Mathematical Modeling of Gene Regulatory Networks

Study	ODE System Representation	Dynamic Behavior Analysis	Parameter Sensitivity
Buriticá et al., 2018	Petri net and Ordinary Differential Equations (ODEs) for gene expression; stochastic simulation via Gillespie algorithm	Comparison of deterministic (ODE) and stochastic (Gillespie) dynamics; analysis of state distributions	Sensitivity analysis via parameter variation; Analysis of Variance (ANOVA) for significance
Rodríguez Penas, 2017	Nonlinear ODEs for dynamic biological systems; high dimensional models	Focus on parameter estimation and model calibration; scalability to large systems	Adaptive strategies to enhance robustness; statistical analysis of parameter effects
John & Meza-Chaves, 2019	Bayesian networks (Directed Acyclic Graphs) for gene interactions; not ODE-based	Dynamic inference from time-series data; model fit assessed via Bayesian posterior	Diversity maintenance in Genetic Algorithm; validation on engineered and real data

• ODE System Representation:
– ODE-based models: 2 of 3 studies.
∗ Of these, 1 also used Petri nets and stochastic simulation (Gillespie algorithm).
∗ 1 used nonlinear, high-dimensional ODEs.
– Non-ODE approach (Bayesian networks/DAGs): 1 of 3 studies.
• Dynamic Behavior Analysis:
– Deterministic vs stochastic comparisons: 1 study.
– State distribution analysis: 1 study.
– Parameter estimation and model calibration: 1 study.
– Scalability analysis: 1 study.
– Dynamic inference from time-series data: 1 study.
– Bayesian model fit assessment: 1 study.
• Parameter Sensitivity:
– Sensitivity analysis via parameter variation: 1 study.
– Analysis of Variance (ANOVA) or other statistical tests for parameter significance: 1 study.
– Adaptive strategies to enhance robustness: 1 study.
– Statistical analysis of parameter effects: 1 study.
– Genetic Algorithm diversity maintenance: 1 study.
– Validation of parameter sensitivity on engineered and real data: 1 study.
No mention found of studies combining all three modeling approaches (ODEs, Petri nets, and Bayesian networks) in a single framework. Each study reported some form of parameter sensitivity analysis, but the specific methods varied.

Biological Applications and Validation

Study	Cell Reprogramming Application	Model Validation Approach	Biological Relevance of Results
Buriticá et al., 2018	Not addressed	Comparison with literature (Goss & Peccoud, 1998); sensitivity analysis; Analysis of Variance (ANOVA)	Highlights importance of stochasticity in gene expression; deterministic models insufficient for population variability
Rodríguez Penas, 2017	Not addressed	Benchmarking on parameter estimation and Mixed-Integer Dynamic Optimization (MIDO) problems; statistical analysis	Demonstrates computational feasibility for large-scale biological models; no mention found of direct biological insights
John & Meza-Chaves, 2019	Not directly addressed, but relevant to gene regulatory network (GRN) inference	Validation on engineered signals and experimental Arabidopsis data; comparison with known relationships	Identifies known and novel gene interactions in auxin response; potential for broader gene regulatory network discovery

• Cell reprogramming:Not directly addressed in any of the three studies; two studies did not address it, and one study was relevant to gene regulatory network inference but did not directly address reprogramming.

• Model validation approaches:

– Comparison with literature, sensitivity analysis, and Analysis of Variance (ANOVA): 1 study.

– Benchmarking on parameter estimation and Mixed-Integer Dynamic Optimization (MIDO) problems, with statistical analysis: 1 study.

– Validation on engineered signals and experimental Arabidopsis data, and comparison with known relationships: 1 study.

• Biological relevance:

– Importance of stochasticity in gene expression and the insufficiency of deterministic models for population variability: 1 study.

– Computational feasibility for large-scale biological models, but no mention found of direct biological insights: 1 study.

– Identification of known and novel gene interactions in auxin response, with potential for broader gene regulatory network discovery: 1 study.

Implementation Considerations

Study	Parameter Types	Value Ranges	Biological Significance	Optimization Impact
Buriticá et al., 2018	Activation, inactivation, synthesis rates	Not specified; varied in sensitivity analysis	Reflects gene activation/inactivation and protein synthesis	Affects simulation outcomes; not used for optimization
Rodríguez Penas, 2017	ODE model parameters (varied, high-dimensional)	Bounded for biological plausibility	Represents kinetic rates, regulatory strengths	Parameter bounds critical for feasible solutions; adaptive strategies improve robustness
John & Meza-Chaves, 2019	Network structure (edges), interaction probabilities	Parameter bounds applied; Hamming distance for diversity	Encodes gene-gene regulatory relationships	Bounds and diversity maintenance prevent overfitting and ensure biological plausibility

• Parameter Types:
– Kinetic or ODE model parameters (activation/inactivation/synthesis rates, or general ODE parameters): 2 studies.
– Network structure (edges) and interaction probabilities as parameters: 1 study.
• Value Ranges:
– Explicit parameter bounds for biological plausibility: 2 studies (one also used diversity maintenance).
– Parameters varied in sensitivity analysis but no explicit bounds specified: 1 study.
• Optimization Impact:
– Parameter bounds (and, in one case, diversity maintenance) were critical for feasible or robust optimization: 2 studies.
– Parameters not used for optimization, only for simulation and sensitivity analysis: 1 study.
No mention found of studies using unbounded parameters for optimization. All studies that performed optimization emphasized the importance of parameter bounds or diversity constraints.

References

D. Buriticá, Bárbara Valeria Mejía Bohórquez, Lina Rojas, and Hector Leandro Sáenz Castro. "Análisis Estocástico de Un Sistema Génico Simple Para La Síntesis de Una Proteína Implementando Los Métodos de Gillespie." Revista Cuarzo, 2018.

David J. John, and Kenneth David Meza-Chaves. "Algoritmos Genéticos Como Mecanismo Para Modelar Interacciones Genéticas Utilizando Información Observada En Distintos Tiempos." Revista Tecnología En Marcha, 2019.

David Rodríguez Penas. "Optimization in Computational Systems Biology via High Performance Computing Techniques," 2017.