

Сетевая организация живых систем: от генных сетей к сетям социальным

Колчанов Н.А.*, Иванисенко В.А., Лашин С.А., Пономаренко М.П.

Институт цитологии и генетики СО РАН, Новосибирск, Россия

* kol@bionet.nsc.ru

В основе функционирования всех типов живых систем лежат генные сети (ГС) – группы координированно функционирующих генов, взаимодействующих друг с другом через продукты – РНК, белки, метаболиты и др., обеспечивающие на основе информации, закодированной в геномах, формирование всех фенотипических характеристик организмов (молекулярных, биохимических, клеточных, физиологических, морфологических, поведенческих и т. д.). Обязательные компоненты ГС – пути передачи сигналов, метаболические пути, сети белок-белковых взаимодействий, регуляторные контуры с отрицательными и положительными обратными связями. ГС могут включать от десятков до многих тысяч генов, РНК, белков, метаболитов и миллионы взаимодействий между ними. Функционирование одних ГС может быть ограничено объемом отдельных клеток, в то время как молекулярные компоненты других генных сетей могут быть распределены по множеству морфологических структур (клеток, тканей и органов). Реконструкция ГС, анализ структурно-функциональной организации и компьютерное моделирование их динамики необходимы для решения широкого круга задач в таких областях, как медицина, фармакология, сельское хозяйство, биотехнологии. К числу ключевых задач относятся: реконструкция молекулярно-генетических механизмов возникновения и развития патологий человека, поиск новых фармакологических мишеней для терапии заболеваний, поиск молекулярных маркёров для селекции сортов растений и пород животных с улучшенными характеристиками, конструирование новых штаммов микроорганизмов с целевыми свойствами – продуцентов биотехнологически ценных белков и метаболитов. В докладе рассмотрены: а) методы системной компьютерной биологии и глубокого машинного обучения для реконструкции ГС на основе анализа больших генетических данных, получаемых с помощью омиксных технологий (геномики, транскриптомики, протеомики, метаболомики); б) методы компьютерного анализа графов генных сетей для выявления ключевых элементов (генов, РНК, белков, метаболитов), наиболее значимых для функционирования ГС; в) методы моделирования динамики функционирования ГС для поиска молекулярных мишеней (генов, РНК, белков), воздействия на которые, включая фармакологические воздействия или влияние мутаций, приводит к выраженному изменению функционирования генных сетей; г) методы конструирования интерпретируемых нейронных сетей, описывающих динамику функционирования ГС, основанные на задании их топологии согласно онтологии структур, функций и процессов предметной области «Генные сети». В заключительной части доклада рассмотрены подходы к конструированию сетей социальных сообществ, основанные на онтологическом описании предметной области «Качество жизни».

Network organization of living systems: from gene networks to social networks

Kolchanov N.A.*, Ivanisenko V.A., Lashin S.A., Ponomarenko M.P.

Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia

* kol@bionet.nsc.ru

The functioning of all types of living systems is based on gene networks (GNs) – groups of coordinately functioning genes that interact with each other through products – RNA, proteins, metabolites, and others, providing the formation of all phenotypic characteristics of organisms (molecular, biochemical, cellular, physiological, morphological, behavioral, etc.) based on the information encoded in genomes. Essential components of GNs include signal transduction pathways, metabolic pathways, protein-protein interaction networks, and regulatory circuits with negative and positive feedback loops. GNs can include from tens to many thousands of genes,

RNAs, proteins, metabolites, and millions of interactions between them. The functioning of some GNs may be limited to individual cells, while the molecular components of other gene networks can be distributed across multiple morphological structures (cells, tissues, and organs). The reconstruction of GNs, analysis of their structural-functional organization, and computer modeling of their dynamics are necessary for solving a wide range of tasks in areas such as medicine, pharmacology, agriculture, and biotechnology. Key tasks include the reconstruction of molecular-genetic mechanisms of the emergence and development of human pathologies, the search for new pharmacological targets for disease therapy, the identification of molecular markers for selecting plant varieties and animal breeds with improved characteristics, and the construction of new strains of microorganisms with target properties – producers of biotechnologically valuable proteins and metabolites. The report considers: (a) methods of computational systems biology and deep machine learning for GN reconstruction based on the analysis of large genetic data obtained using omics technologies (genomics, transcriptomics, proteomics, metabolomics); (b) methods of computer analysis of gene network graphs to identify key elements (genes, RNAs, proteins, metabolites) most significant for GN functioning; (c) methods of modeling the dynamics of GN functioning to search for molecular targets (genes, RNAs, proteins) whose impact (including pharmacological effects or the influence of mutations) leads to significant changes in gene network functioning; (d) methods of constructing interpretable neural networks describing the dynamics of GN functioning based on the assignment of their topology according to the ontology of structures, functions, and processes of the “Gene networks” subject area. In the final part of the report, approaches to the construction of social community networks based on the ontological description of the subject area “Quality of life” are considered.

生命系統的網絡組織：從基因網絡到社會網絡
在。

Kolchanov N.A.*, Ivanisenko V.A., Lashin S.A., Ponomarenko M.P.

拉辛，國會議員 波諾馬連科

ICG SB RAS，俄羅斯新西伯利亞

* kol@bionet.nsc.ru

所有類型的生命系統的功能都基於基因網絡（GNs）－協調功能的基因組通過產物－RNA、蛋白質、代謝物等相互作用，在基因組中編碼的信息的基礎上提供生物體所有表型特徵的形成（分子、生化、細胞、生理、形態、行為等）。GS的強制性組成部分是信號轉導途徑、代謝途徑、蛋白質－蛋白質相互作用網絡、具有負反饋和正反饋的調節迴路。GS可以包括從數萬到數千個基因、RNA、蛋白質、代謝物以及它們之間的數百萬種相互作用。一些GS的功能可能受到單個細胞體積的限制，而其他基因網絡的分子成分可以分佈在各種形態結構（細胞、組織和器官）中。HSs的重建、結構和功能組織的分析及其動力學的計算機模擬對於解決醫學、藥理學、農業和生物技術等領域的廣泛問題是必要的。重點任務包括：重構人類病理髮生和發展的分子遺傳機制，尋找治療疾病的新藥理靶點，尋找分子標記用於培育具有改良特性的植物品種和動物品種，構建具有目標特性的新型微生物菌株——具有生物技術價值的蛋白質和代謝物的生產者。該報告考慮了：(a) 基於使用組學技術（基因組學、轉錄組學、蛋白質組學、代謝組學）獲得的大遺傳數據分析重建GS的系統計算機生物學和深度機器學習方法；(b) 基因網絡圖的計算機分析方法，以確定對GS功能最重要的關鍵要素（基因、RNA、蛋白質、代謝物）；(c) 模擬GS功能以搜索分子目標（基因、RNA、蛋白質）的動力學方法，其影響（包括藥理作用或突變

的影響) 導致基因網絡功能的顯著變化; (d) 構建描述 GS 功能動態的可解釋神經網絡的方法, 基於根據基因網絡主題領域的結構、功能和過程的本體設置其拓撲結構。在報告的最後部分, 考慮了基於主題領域“生活質量”的本體論描述來構建社會社區網絡的方法。

DOI 10.18699/sblai2023-22

Нейросетевое моделирование распространения COVID-19 с учетом социально-экономических процессов

Криворотко О.И.^{1, 2*}, Кабанихин С.И.¹, Зятков Н.Ю.¹

¹ Институт математики им. С.Л. Соболева СО РАН, Новосибирск, Россия

² Институт вычислительной математики и математической геофизики СО РАН, Новосибирск, Россия

* krivorotko.olya@mail.ru

Сложность описания эпидемических процессов возрастает с появлением новых вирусов (новые штаммы COVID-19, туберкулеза и др.) и влиянием социально-экономических процессов (изоляция, масочный режим, массовая вакцинация и др.). Процесс и скорость распространения инфекционных заболеваний характеризуются данными, представляющими распределения эпидемиологических характеристик (среднее время течения болезни, инкубационного периода и др.), временные ряды статистических наблюдений (например, количество ежедневно выявленных, госпитализированных, протестированных), законами распространения, а также связанными с ними данными о социально-экономических процессах. Развитие методов машинного обучения и анализа данных позволяет на новом уровне построить модели описания динамики распространения заболеваний с учетом социально-экономических характеристик [1]. В работе построена и оптимизирована полносвязная рекуррентная нейронная сеть, позволяющая осуществлять недельные прогнозы по количеству ежедневно выявленных и критических (подключенных к аппарату ИВЛ) случаев на основе эпидемиологических и социально-экономических показателей [2, 3]. Численные прогнозы сопоставлены с результатами применения классических компартментных и агентных моделей на примере распространения новой коронавирусной инфекции COVID-19 в Новосибирской области [2, 4]. Показано преимущество использования оптимизированной нейронной сети в случае учета влияния социально-экономических показателей на распространение эпидемии при построении краткосрочных прогнозов.

Работа выполнена при поддержке Математического Центра в Академгородке, соглашение с Министерством науки и высшего образования Российской Федерации № 075-15-2022-281.

Литература

1. Nikparvar B., Rahman Md.M., Hatami F., Thill J.-C. Spatio-temporal prediction of the COVID-19 pandemic in US counties: modeling with a deep LSTM neural network, Scientific Reports, 2021, vol. 11, P. 21715.
2. Krivorotko O.I., Zyatkov N.Y. Data-driven regularization of inverse problem for SEIR-HCD model of COVID-19 propagation in Novosibirsk region, Eurasian Journal of Mathematical and Computer Applications, 2022, vol. 10, iss. 1, P. 51-68.
3. Криворотко О.И., Кабанихин С.И. Математические модели распространения COVID-19. Препринт Института математики им. С.Л. Соболева СО РАН, № 300, Новосибирск: ИМ СО РАН, 2022.
4. Krivorotko O., Sosnovskaia M., Vashchenko I., Kerr C., Lesnic D. Agent-based modeling of COVID-19 outbreaks for New York state and UK: parameter identification algorithm, Infectious Disease Modelling, 2022, vol. 7, P. 30-44.