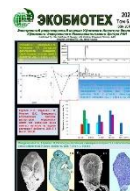




# ЭКОБИОТЕХ

ISSN 2618-964X

<http://ecobiotech-journal.ru>


## МОДЕЛИРОВАНИЕ СИСТЕМЫ ЭПИГЕНЕТИЧЕСКОЙ ПАМЯТИ С ТРАНСКРИПЦИОННОЙ РЕПРЕССИЕЙ И МЕТИЛИРОВАНИЕМ ДНК

Галимзянов А.В.

Уфимский Институт биологии Уфимского федерального  
исследовательского центра РАН, Уфа, Россия  
E-mail: [ganim@anrb.ru](mailto:ganim@anrb.ru)

Разработана *in silico* модель синтезированной в *Escherichia coli* системы эпигенетической памяти [Maier et al., 2017, Nat. Comm., 8:15336]. Система включает в себя искусственный оперон памяти с ДНК-метилтрансферазой CcrM из *Caulobacter crescentus*, метилирующей промотор этого оперона, и ген инженерного транскрипционного репрессора ZnF, блокирующего тот же самый промотор в неметилированном состоянии. Управление опероном памяти, который располагается на основной плазмиде, осуществляется с триггерной плазмиды, несущей специфичный оперон, индуцируемый арабинозой. Искусственный триггерный оперон функционирует в составе расширенного за счет него арабинозного регулона. Для описания работы последнего используется дополнительная модель. Объединенная модель системы управления памятью учитывает позитивный и негативный контроль регуляторным белком araC, отрицательную авторегуляторную и отрицательную прямую, чувствительную к метилированию регуляцию белком ZnF и метилирование ДНК. Модели построены на языке формализма динамической теории молекулярно-генетических систем управления. Расчет моделей проведен в программном комплексе SETIES для анализа динамики управляющих генных сетей, динамических эпигенов и клеточных ансамблей. В компьютерных экспериментах показано, что переключатель имеет два альтернативных режима функционирования, каждый из которых стабилен в ряду клеточных делений. Оба режима охарактеризованы с точки зрения динамики молекулярных компонент искусственного оперона памяти, триггерного оперона и природного оперона катаболизма L-арабинозы.

**Ключевые слова:** генная сеть ♦ транскрипционный репрессор ♦ ДНК-метилтрансфераза ♦ арабинозный регулон ♦ динамика

Поступила в редакцию: 18.12.2023

[Цитировать | Cite as](#)

DOI: [10.31163/2618-964X-2023-6-4-262-272](https://doi.org/10.31163/2618-964X-2023-6-4-262-272)

EDN: PULEBH

## MODELING OF THE EPIGENETIC MEMORY SYSTEM WITH TRANSCRIPTIONAL REPRESSION AND DNA METHYLATION

Galimzyanov A.V.

Ufa Institute of Biology of the Ufa Federal Research Centre of  
the Russian Academy of Sciences, Ufa, Russia  
E-mail: [ganim@anrb.ru](mailto:ganim@anrb.ru)

The *in silico* model of the epigenetic memory system synthesized in *Escherichia coli* [Maier et al., 2017, Nat. Comm., 8:15336] was developed. The system includes an artificial memory operon with DNA methyltransferase CcrM from *Caulobacter crescentus* methylating the promoter of this operon, and an engineered transcriptional repressor ZnF that blocks the same promoter in an unmethylated state. The memory operon, which is located on the maintenance plasmid, is controlled by a trigger plasmid carrying a specific arabinose-inducible operon. The artificial trigger operon functions as part of an expanded arabinose regulon. An additional model is used to describe the operation of the latter. The unified model of the memory control system takes into account positive and negative control by the regulatory protein araC, negative feedback self-regulation and negative direct, methylation-sensitive regulation by the ZnF protein, and DNA methylation. The models are built in the language of the formalism of the dynamic theory of molecular genetic control systems. The models were calculated in the SETIES software package to analyze the dynamics of control gene networks, dynamic epigenes and cell ensembles. Computer experiments have shown that the switch has two alternative functional modes, each of which is stable in a series of successive cell divisions. Both modes are characterized in terms of the molecular component dynamics of the artificial memory operon, the trigger operon, and the natural L-arabinose catabolism operon.

**Keywords:** gene circuit ♦ transcriptional repressor ♦ DNA methyltransferase ♦ ara operon ♦ dynamics

Принято в печать: 28.12.2023



## ВВЕДЕНИЕ

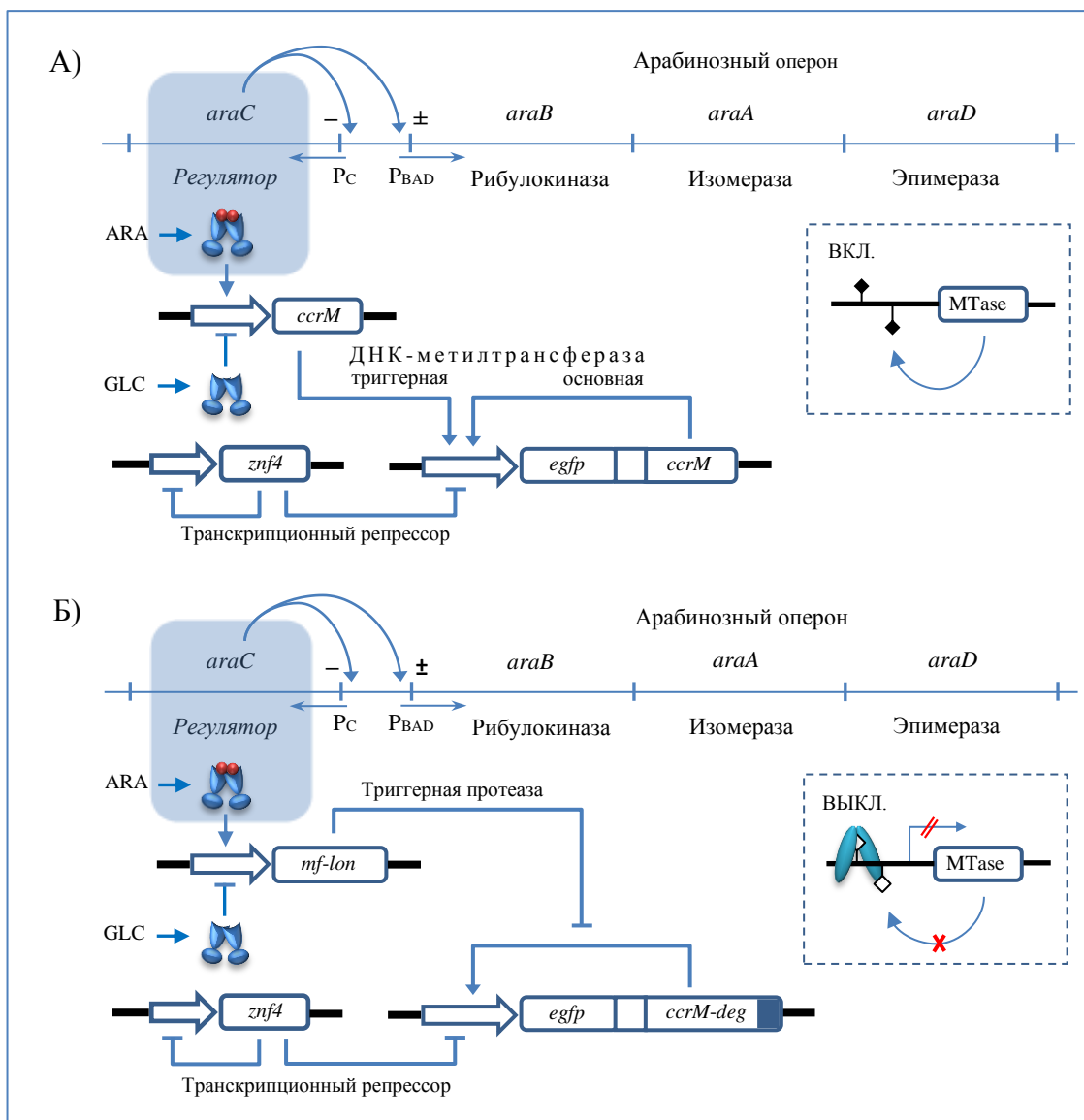
Дизайн искусственных генетических регуляторных систем является ведущим направлением современной синтетической биологии [Gao et al., 2023; Vazquez-Vilar et al., 2023]. Среди них особый интерес представляют генетические переключатели как системы, имеющие известную ценность и прогнозируемый потенциал в плане онтогенетических, экологических и биотехнологических функций. В живой природе переключатели могут быть носителями динамической наследственной памяти [Галимзянов и др., 2019] или выполнять роль стартеров молекулярно-генетических программ [Lee et al., 2022]. В практическом применении они могут служить основой для эпигенетических биосенсоров, чувствительных к присутствию антибиотиков [Ullrich et al., 2020], или выступать в качестве агентов контролируемого синтеза биофармацевтических препаратов в биореакторах [Strittmatter et al., 2021]. При этом решение задач анализа и синтеза таких систем предполагает использование портретных математических и компьютерных моделей [Şimşek et al., 2023].

В настоящей работе для компьютерного моделирования выбрана малая искусственная генная сеть, регуляторный контур которой образован чувствительным к метилированию ДНК транскрипционным репрессором и опероном-мишенью с ДНК-метилтрансферазой и геном-репортером. Сеть принадлежит семейству систем памяти, синтезированных в [Maier, 2017]. Отличительной чертой этих систем является способность хранить эпигенетическую информацию в форме обратимых паттернов метилирования ДНК. Другой особенностью является то, что сигналы на активацию или сброс памяти формируются по схеме индукции арабинозного оперона. Все это дает материал для развития портретных моделей в контексте сопряжения динамики искусственных конструкций с конститутивными молекулярно-генетическими системами клетки, что актуально как в проблематике минимизации нагрузки на физиологию бактериальной клетки, так и в аспекте эколого-генетической безопасности.

## МАТЕРИАЛЫ И МЕТОДЫ

Система эпигенетической памяти, синтезированная и функционирующая в клетках *E. coli* [Maier et al., 2017], представляет собой плазмидный генно-сетевой переключатель, состоящий из двух генов (рис. 1). Синтетический ген *znf4* кодирует транскрипционный репрессор ZnF4 (Zn – zinc finger, 4 – гомотетрамер), чувствительный к метилированию ДНК целевого промотора. Природный (у *Caulobacter crescentus*) ген *ccrM* кодирует ДНК-метилтрансферазу CcrM (cell cycle-regulated methyltransferase), метилирующую остатки аденина в последовательностях GANTC.

Ген *ccrM* вместе с маркерным геном *egfp* образуют искусственный оперон памяти. Промотор оперона, сформированный на основе sigma70-промотора VBa\_J23107, метилируется белком CcrM, являющимся продуктом самого этого оперона, что создает положительную обратную связь в регуляции *ccrM*. Репрессор ZnF4 также способен связываться с промотором оперона памяти и, тем самым, подавлять его транскрипцию. Однако сайты связывания ZnF4 в области промотора частично перекрываются с сайтами GANTC, поэтому процессы репрессии и метилирования здесь альтернативны: когда промотор заблокирован белком ZnF4 и CcrM отсутствует, *ccrM* устойчиво репрессируется (OFF-состояние); когда промотор метилирован CcrM и связывание с ним репрессора невозможно, *ccrM* устойчиво экспрессируется (ON-состояние). Кроме того, для предотвращения избыточных концентраций ZnF4 кодирующий его ген поставлен под контроль промотора LacIQ1, активность которого также блокируется ZnF4 и не зависит от метилирования (негативная авторегуляция).



**Рис. 1. Эпигенетический переключатель с репрессором ZnF4 и ДНК-метилтрансферазой CcrM под управлением арабинозного регулона [Maier et al., 2017; Schleif, 2010].**

А) Инициализация памяти посредством экспрессии гена метилтрансферазы на «пусковой» плазмиде в результате индукции *ara*-регулона (установка в ON-состояние). Б) Сброс памяти посредством синтеза *mf-Lon*, деградации *CcrM* и возобновления репрессии *ccrM* со стороны *ZnF4* (возврат в OFF-состояние).

Для управления функциональными состояниями этой бистабильной системы используется температурный сигнал (переход 30 → 37°C), ослабляющий связывание *ZnF4* с промотором гена *ccrM*, дублирующая экспрессия копии гена метилтрансферазы под контролем независимого от *ZnF4* и *CcrM* промотора, а также модификация белка *CcrM* со вставкой, делающей его чувствительным к протеазе *Lon* из *Mesoplasma florum* (*mf-Lon*). В двух последних случаях опероны, состоящие из кодирующих *CcrM* или *mf-Lon* генов, а также маркерных генов, локализованы на пусковой (trigger) плазмиде, используемой дополнительно к основной (maintenance), и находятся под контролем промотора P<sub>BAD</sub>, с которым связывается регуляторный белок *AraC* (ген *araC* также клонирован в пусковую плазмиду). При этом плазмидная цепь P<sub>BAD</sub>-*araC*-*AraC* является искусственным расширением природного арабинозного регулона [Schleif, 2010], утилизирующего L-арабинозу в клетках *E. coli*, и индуцируется L-арабинозой вместе с индукцией хромосомного *ara*-оперона.

Арабинозный регулон у *E. coli* состоит из оперонов *araBAD*, *araC*, *araFG* и *araE*, локализованных в трех различных участках (рис. 1) [Schleif, 2010]. Операторные зоны оперонов *araBAD* и *araC* перекрываются, образуя общий управляющий участок *araCBAD*, причем опероны транскрибируются в противоположных направлениях. Гены *araB*, *araA* и *araD* кодируют, соответственно, рибулокиназу, арабинозоизомеразу и L-рибулозофосфат 4-эпимеразу, расщепляющие L-арабинозу до D-ксилоулозо-5-фосфата. Ген *araC* кодирует регуляторный белок, в нативном состоянии находящийся в форме димеров. В транспортную систему регулона входят гены *araE*, *araF* и *araG*, кодирующие L-арабинозопермеазу, связывающий белок и белок, ассоциированный с мембраной. Опероны *araE* и *araFG*, функционирующие сходным образом, можно рассматривать совместно, как единый транспортный *araT*-оперон, кодирующий “эффективный” транспортный Т-белок [Prokudina et al., 1991].

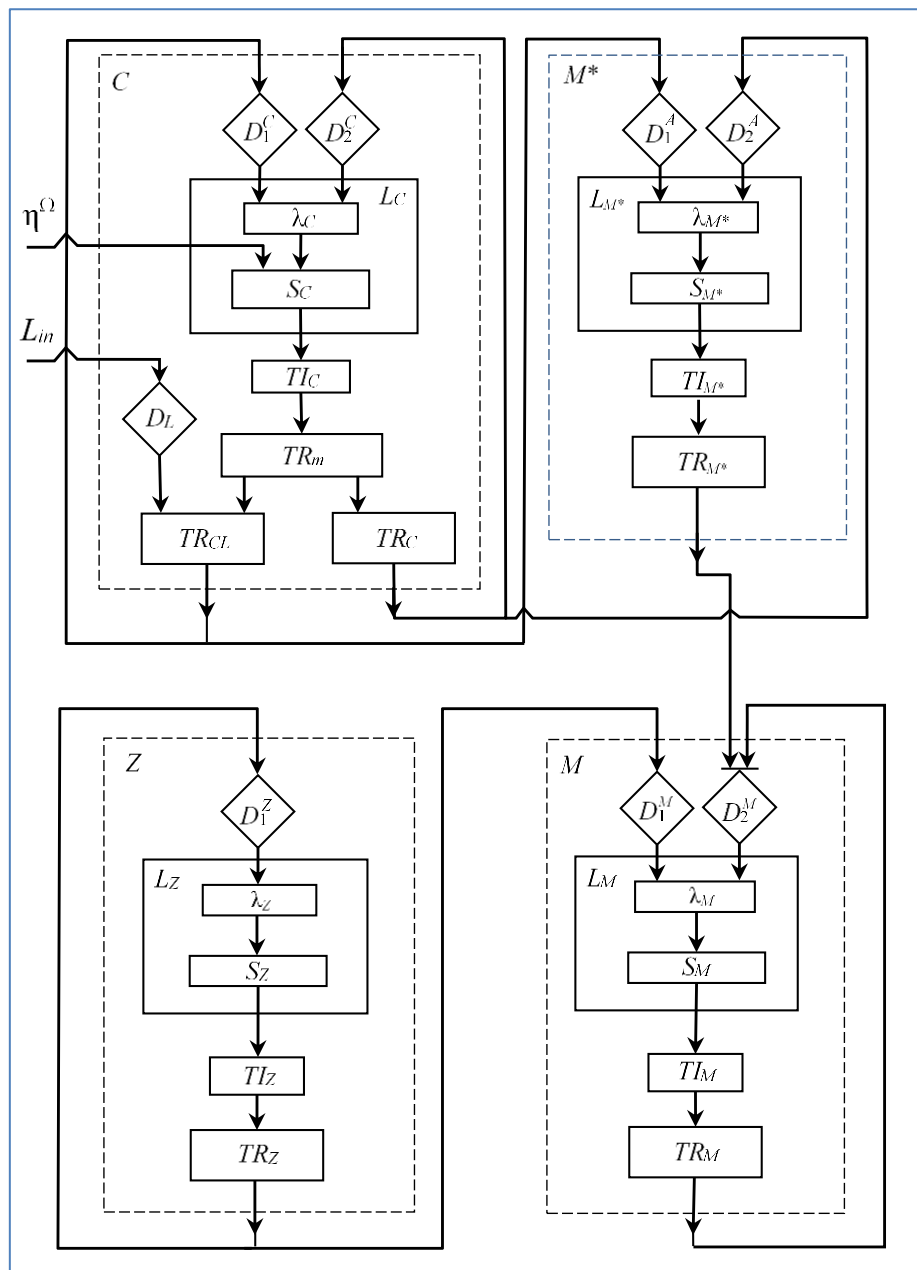
Построение портретных моделей генных сетей проводилось на языке формализма обобщенных пороговых моделей (ОПМ) [Tchuraev, Galimzyanov, 2009]. Расчет динамики в моделях осуществлялся на платформе «SETIES» для исследования молекулярно-генетических систем управления [Галимзянов, 2015а]. Для создания программных компонент использовалась интегрированная среда разработки Borland Delphi 7 (Embarcadero Technologies). В качестве языка описания информационных моделей генных сетей и их комплексной репрезентации применялся SETIES Standard Gene Network Language [Галимзянов, 2015б].

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Обобщенная пороговая модель эпигенетического переключателя состоит из генетического блока *Z* синтеза репрессора ZnF4, блока *M* синтеза ДНК-метилтрансферазы с основной плазмиды, блока *M\** синтеза ДНК-метилтрансферазы (или блока *P* синтеза протеазы для основной ДНК-метилтрансферазы) с «пусковой» плазмиды, а также блока *C* синтеза репрессора / активатора *ara*-оперона (рис. 2). Скорости синтеза и деградации генных продуктов представлены в табл. 1.

В клеточной управляющей подсистеме в качестве элементов рассматриваются *генетические блоки* (G-блоки). Генетический блок – это ген, взятый в совокупности с механизмами транскрипции, процессинга, транспорта и депо конечного продукта. В *j*-м генетическом блоке учитываются следующие величины:  $m_j(t)$  и  $r_j(t)$  – концентрации, соответственно, мРНК и белка, выраженные числом молекул на клетку;  $a_{1j}$  – единичная интенсивность транскрипции (сила промотора) (молек./мин);  $a_{2j}$  – единичная интенсивность трансляции (молек./мин);  $b_{1j}$  и  $b_{2j}$  – скорости деградации, соответственно, мРНК и белка (молек./мин), вычисляемые как  $\ln 2 / t_{1/2}$ , где  $t_{1/2}$  – время полужизни вещества;  $P_{ij}$  – константы, описывающие пороговую концентрацию белкового продукта *i*-го гена, необходимую для подавления синтеза транскриптов с *j*-го гена, которые определяются сродством регуляторных веществ к сайтам связывания белков и степенью мультимерии;  $T$  – длительность клеточного цикла (мин).

Компьютерная модель арабинозного регулона разработана на основе его математической модели, предложенной в работе [Prokudina et al., 1991]. Для описания динамики системы управления катаболизмом L-арабинозы методом ОПМ выделены генетические блоки *C*, *A* и *T*, соответствующие оперонам *araC*, *araBAD* и *araT* (рис. 3). Скорости синтеза и деградации генных продуктов представлены в табл. 2.

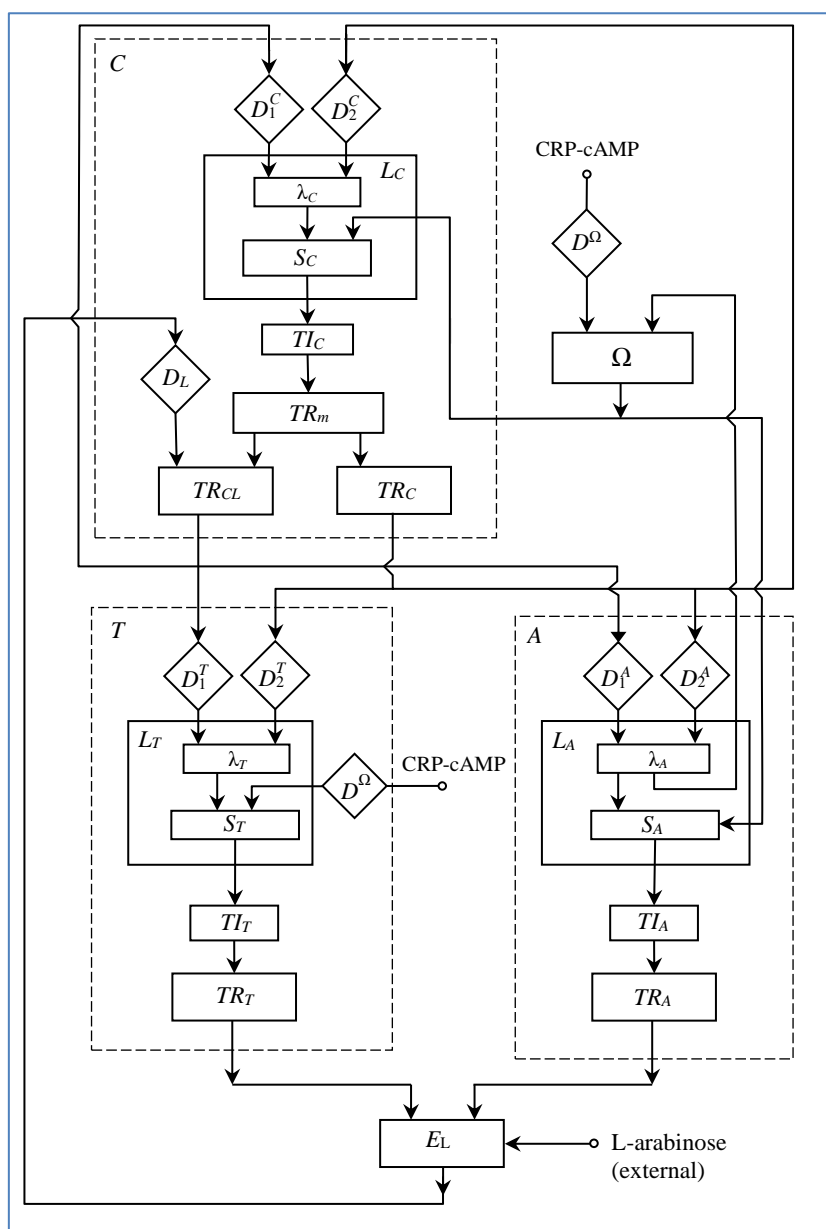


**Рис. 2. Структура модели эпигенетического переключателя.** Генетические блоки *Z*, *M*, *C* и *M\** содержат следующие элементы: *D* – дискриминатор пороговых значений концентраций регуляторов;  $\lambda$  – конечный автомат; *S* – комбинатор бинарных сигналов; *TI* – элемент задержки; *TR* – блок синтеза белка.

**Таблица 1. Константы синтеза и деградации для мРНК и белков переключателя**

Параметр	Генетические блоки			
	Z	M	M*	C
$a_{11}$ , молек. мин <sup>-1</sup>	1,5	1	1	0,04
$a_{12}$ , молек. мин <sup>-1</sup>	-	-	-	0,2
$a_2$ , молек. мин <sup>-1</sup>	2,5	1,7	2,1	1,5
$b_1$ , мин <sup>-1</sup>	0,2	0,231	0,231	0,3
$b_2$ , мин <sup>-1</sup>	0,1	0,087	0,087	0,01

Примечание. Скорости  $a_{11}$  и  $a_{12}$  – синтез мРНК,  $a_2$  – синтез белка,  $b_1$  – деградация мРНК,  $b_2$  – деградация белка.



**Рис. 3. Структура системы управления катаболизмом L-арабинозы [Prokudina et al., 1991].** Генетические блоки A, C и T содержат следующие элементы: D – дискриминатор пороговых значений концентраций регуляторов;  $\lambda$  – конечный автомат; S – комбинатор бинарных сигналов; TI – элемент задержки; TR – блок синтеза белка. Блок  $\Omega$  описывает процесс образования репрессорной петли в зоне *araCBAAD*. В блоке  $E_L$  вычисляется внутриклеточная концентрация арабинозы  $L_{in}$ .

**Таблица 2. Константы синтеза и деградации для мРНК и белков ара-регулона**

Параметр	Генетические блоки		
	C	A	T
$a_{11}$ , молек. мин <sup>-1</sup>	0,04	0,5	0,3
$a_{12}$ , молек. мин <sup>-1</sup>	0,2	3	3
$a_2$ , молек. мин <sup>-1</sup>	1,5	5	5
$b_1$ , мин <sup>-1</sup>	0,3	0,5	0,5
$b_2$ , мин <sup>-1</sup>	0,01	0,03	0,04

Примечание. Скорости  $a_{11}$  и  $a_{12}$  – синтез мРНК,  $a_2$  – синтез белка,  $b_1$  – деградация мРНК,  $b_2$  – деградация белка.

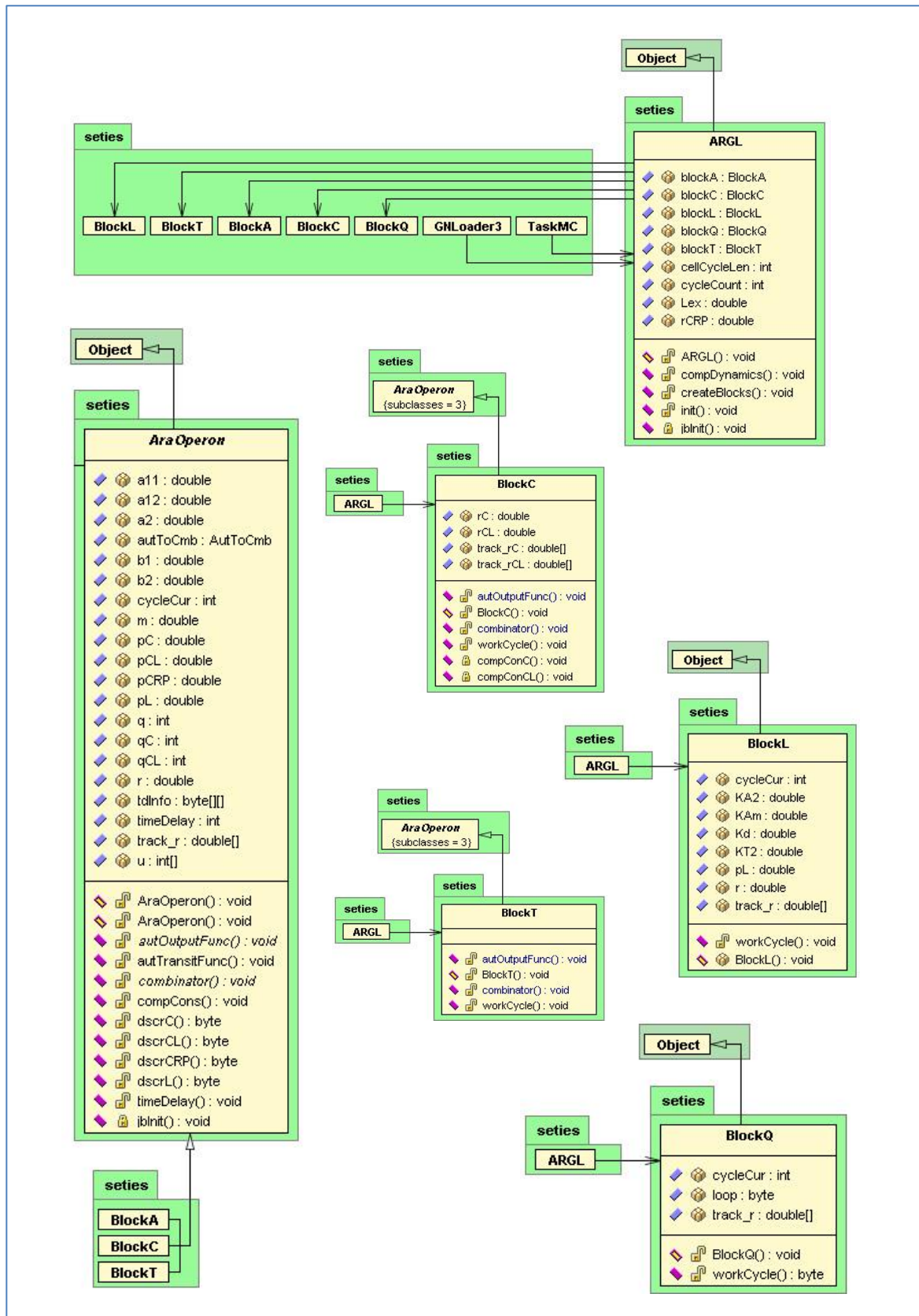
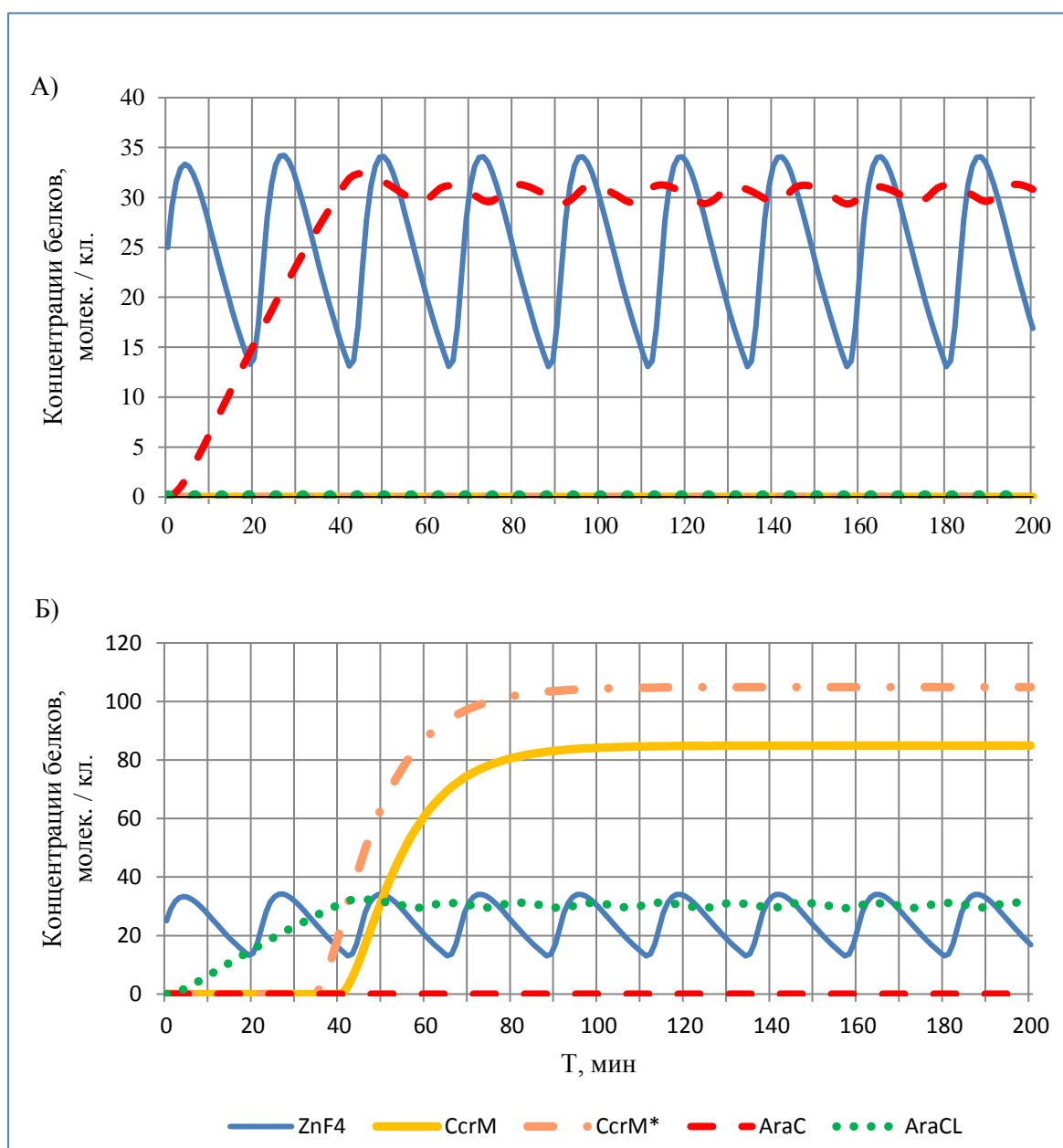


Рис. 4. UML-диаграммы классов объектно-ориентированной модели *ara*-регулона. Абстрактные классы программной реализации – виртуальный класс «AraOperon» и производные от него классы «BlockC», «BlockA», «BlockT», конкретизирующие характеристики моделируемых элементов.

Обе модели объединены на платформе SETIES для моделирования генных сетей, динамических эпигенов и клеточных ансамблей (рис. 4).



Модель функционирует следующим образом. В отсутствие L-арабинозы в клетке активны только гены *araC* и *znf4* (рис. 5А, 6А). Белок AraC (репрессорная форма) блокирует активность целевых оперонов. Белок ZnF4 репрессирует *ccrM1* и *znf4* (последний функционирует в режиме автоколебаний). Индукция арабинозного оперона и оперона на триггерной плазмиде происходит при достижении внутриклеточной концентрации L-арабинозы порога своего действия. Димеры белка AraC связываются с L-арабинозой и в форме активаторных комплексов способствуют экспрессии генов целевых оперонов – как самого арабинозного оперона, так и гена *ccrM* на триггерной плазмиде под контролем промотора P<sub>BAD</sub> (рис. 5Б, 6Б). Максимальная активность промоторов достигается при наличии надпороговой дозы комплекса CRP-сAMP.

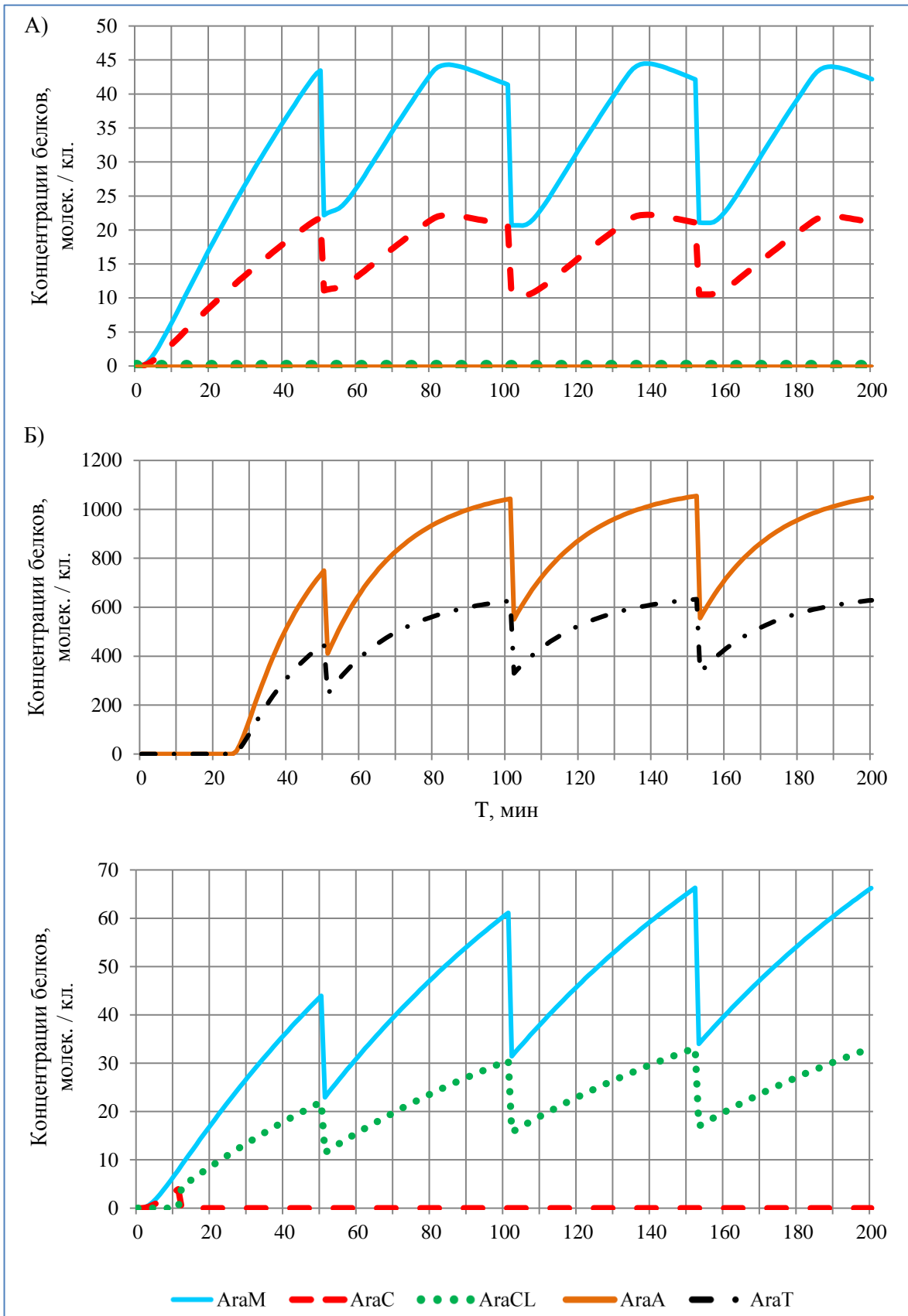


**Рис. 5. Динамика в модели эпигенетической памяти.**

А) *Ara*-регулон не индуцирован, система памяти неактивна.

Б) *Ara*-регулон индуцирован, система памяти устойчиво переключается в активное состояние.





**Рис. 6. Динамика в модели арабинозного регулона с учетом клеточных делений.**

А) *Ara*-регулон не индуцирован (концентрация арабинозы L и регуляторного комплекса CRP-cAMP ниже своих порогов), синтезируются мономеры AraM с образованием димеров AraC (репрессорная форма). Б) *Ara*-регулон индуцирован, ген *araC* активен на более высоком уровне, синтезируются мономеры AraM с образованием димеров AraC, связывающихся с арабинозой (AraCL, активаторная форма), а также белки AraA и AraT.

Генно-сетевая регуляция генома и метилирование ДНК, наряду с РНК-интерференцией, являются ведущими механизмами, обеспечивающими эпигенетические свойства бактериальных клеток и клеточных популяций. Совместное их моделирование требует, с одной стороны, интеграции фактического материала, а с другой – развития инструментария в области компьютерной системной биологии. В представленной модели бистабильной эпигенной сети фактор метилирования ДНК комбинируется с базовым свойством генных регуляторных сетей, а именно с циркулированием молекулярных сигналов в контурах прямых и обратных, положительных и отрицательных регуляторных связей. Изучение синтетических конструкций и управляющих ими схем с учетом метилирования регуляторных районов генов открывает возможности для получения и использования новых знаний о механизмах дифференцировки и передачи сигналов в клетке и, следовательно, об онтогенезе и адаптации живых систем к изменяющимся условиям среды.

### ФИНАНСОВАЯ ПОДДЕРЖКА

Работа выполнена в рамках государственного задания Минобрнауки России по теме № 220131100163-4 «Межвидовые взаимодействия в микробных сообществах и растительно-микробных ассоциациях естественных и техногенных экосистем (генетические, биохимические и биотехнологические аспекты)».

### СПИСОК ЛИТЕРАТУРЫ

1. Галимзянов А.В., Ступак Е.Э., Чураев Р.Н. Эпигенные сети: теория, модели, эксперимент // Успехи современной биологии. 2019. Т. 139 (2). С. 107–113. DOI: [10.1134/S0042132419020030](https://doi.org/10.1134/S0042132419020030)
2. Галимзянов А.В. SETIES 5.0 Java-платформа (программное ядро) «SETIES 5.0 J-Core» для исследования молекулярно-генетических систем // Офиц. бюл. РОСПАТЕНТА «Компьютерные программы. Базы данных. Топологии интегральных микросхем». 2015а. №10 (108). ID 2015660238. <http://fips.ru/Archive/EVM/2015/2015.10.20/index.htm>, EDN: KOYRYL
3. Галимзянов А.В. Командная оболочка «SETIES Gene Networking Shell» // Офиц. бюл. РОСПАТЕНТА «Компьютерные программы. Базы данных. Топологии интегральных микросхем». 2015б. № 10 (108). ID 2015660326. <http://fips.ru/Archive/EVM/2015/2015.10.20/index.htm>, EDN: JYGLYF
4. Gao Y., Wang L., Wang B. (2023). Customizing cellular signal processing by synthetic multi-level regulatory circuits // Nat. Commun. 2023. V. 14 (1). Art. 8415. DOI: [10.1038/s41467-023-44256-1](https://doi.org/10.1038/s41467-023-44256-1)
5. Lee H.J., Kim H.J., Lee S.J. Control of  $\lambda$  Lysogenic *Escherichia coli* Cells by Synthetic  $\lambda$  Phage Carrying *cI<sup>antisense</sup>* // ACS Synth. Biol. 2022. V. 11 (11). P. 3829–3835. DOI: [10.1021/acssynbio.2c00409](https://doi.org/10.1021/acssynbio.2c00409)
6. Maier J.A.H., Möhrle R., Jeltsch A. Design of synthetic epigenetic circuits featuring memory effects and reversible switching based on DNA methylation // Nat. Commun. 2017. V. 8. Art. 15336. DOI: [10.1038/ncomms15336](https://doi.org/10.1038/ncomms15336)
7. Prokudina E.I., Valeev R.Y., Tchuraev R.N. A new method for the analysis of the dynamics of the molecular genetic control systems. II. Application of the method of generalized threshold models in the investigation of concrete genetic systems // J. Theor. Biol. 1991. V. 151 (1). P. 89–110. DOI: [10.1016/s0022-5193\(05\)80145-6](https://doi.org/10.1016/s0022-5193(05)80145-6)

8. Schleif R. AraC protein, regulation of the L-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action // FEMS Microbiol. Rev. 2010. V. 34. P. 779–796. DOI: [10.1111/j.1574-6976.2010.00226.x](https://doi.org/10.1111/j.1574-6976.2010.00226.x)
9. Şimşek E., Yao Y., Lee D., You L. Toward predictive engineering of gene circuits // Trends Biotechnol. 2023. V. 41 (6). P. 760–768. DOI: [10.1016/j.tibtech.2022.11.001](https://doi.org/10.1016/j.tibtech.2022.11.001)
10. Strittmatter T., Egli S., Bertschi A., et al. Gene switch for l-glucose-induced biopharmaceutical production in mammalian cells // Biotechnol. Bioeng. 2021. V. 118 (6). P. 2220–2233. DOI: [10.1002/bit.27730](https://doi.org/10.1002/bit.27730)
11. Ullrich T, Weirich S, Jeltsch A. Development of an epigenetic tetracycline sensor system based on DNA methylation // PLoS One. 2020. V. 15 (5). Art. e0232701. DOI: [10.1371/journal.pone.0232701](https://doi.org/10.1371/journal.pone.0232701)
12. Vazquez-Vilar M., Selma S., Orzaez D. The design of synthetic gene circuits in plants: new components, old challenges // J. Exp. Bot. 2023. V. 74 (13). P. 3791–3805. DOI: [10.1093/jxb/erad167](https://doi.org/10.1093/jxb/erad167)

Цитировать как

Галимзянов А.В. Моделирование системы эпигенетической памяти с транскрипционной репрессией и метилированием ДНК // Экобиотех, 2023, Т. 6 № 4. С. 262-272. DOI: [10.31163/2618-964X-2023-6-4-262-272](https://doi.org/10.31163/2618-964X-2023-6-4-262-272)  
EDN: PULEBH

Cited as

Galimzyanov A.V. Modeling of the epigenetic memory system with transcriptional repression and DNA methylation. *Ekobiotech.* 2023, V. 6 (4). P. 262-272. DOI: [10.31163/2618-964X-2023-6-4-262-272](https://doi.org/10.31163/2618-964X-2023-6-4-262-272) EDN: PULEBH (In Rus.)