

УДК 575+573+519.8

DOI 10.18413/2687-0932-2020-47-2-372-379

**ОБЗОР МЕТОДОВ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ СВОЙСТВ
ГЕНЕТИЧЕСКОГО КОДА****A REVIEW OF MATHEMATIC MODELING METHODS OF GENETIC CODE
PROPERTIES****А.И. Гарянина, Н.И. Червяков**
A.I. Garianina, N.I. ChervyakovФГАОУ ВО «Северо-Кавказский федеральный университет»,
Россия, 355017, г. Ставрополь, ул. Пушкина, 1

North-Caucasus Federal University, 1 Pushkina St, Stavropol, 355017, Russia

E-mail: garyanina.anastasia@gmail.com

Аннотация

Рассматриваются подходы к математическому моделированию свойств генетического кода. В основе моделирования лежат методы картирования генов. Современный генетический код представляется как система для обработки, хранения и поиска генетической информации. Производится сравнение искусственно сгенерированных генетических кодов, называемых теоретическими, и стандартного генетического кода. Данная статья дает краткий обзор разработанных математических моделей для изучения свойств генетического кода. Изученные методы помогают найти ответ на вопрос, почему природа предпочитает стандартный код теоретическим генетическим кодам.

Abstract

The paper gives brief analysis of the developed statistical models studying the properties of the genetic code. Authors generate the codes, called theoretical genetic codes having the same properties as trivial genetic code. The main idea of the analyzed mathematical models is to find the optimal code from theoretical genetic codes. Based on the following criterias of amino acids: hydrophilic properties, isoelectric properties, affinity to water and molecular mass authors compare the theoretical genetic codes to the trivial genetic code. They also used the extra criterion – transition or transversion, meaning the nature of point mutation. The retrospect of the key elements of these approaches to modeling is given. The subject of modeling in the studied methods is gene mapping that play an important role in the study of differences in individual genomes. Prediction of the properties of theoretical codes and the creation of new approaches to their models will allow us to come closer to understanding the structure of the genetic code.

Ключевые слова: генетический код, картирование генов, точечные мутации, моделирование, перекрывающиеся гены.

Keywords: genetic code, gene mapping, point mutations, modeling, overlapping genes.

Введение

Открытие структуры ДНК и расшифровка генетического кода дали толчок быстрому и непрерывному развитию генетики и молекулярной биологии, а также возникновению новых отраслей науки, например, молекулярных биотехнологий. Стремительное развитие таких технологий совершенно изменило методологию, особенно в науках с биологическим предметом исследования [Trenchev I. et al., 2017].

По мнению С.В. Петухова [2008], генетический код занимает в науке совершенно особое место, как хорошо формализуемая с математической точки зрения биологическая составляющая, единая для всего живого. Сегодня последовательность азотистых оснований

ДНК признается главным источником биологической информации. Однако расшифрованные на сегодняшний день множества последовательностей нуклеотидов все еще не дают нам понимания того, как реализуются изученные свойства генетического кода.

Современные технологии позволяют генерировать код, обладающий свойствами и возможностями природного. Такие коды называются теоретическими, их число составляет 10^{84} [Trenchev I. et al., 2017]. Несмотря на глубокие исследования в данной области, природный генетический код все еще является единственным. Прогнозирование свойств теоретических кодов и создание новых подходов к их моделированию позволит приблизиться к пониманию особенностей структуры генетического кода.

Можно предположить, что наше время представляет молекулярную биологию и генетику, и начинается с открытия двойной спирали ДНК в 1953 Уотсоном и Криком [Ahmad M. et al., 2017]. Получили развитие следующие математические модели в биологии: прогноз вторичной структуры РНК, синтеза белка, в котором информация, инвестируемая в код, переводится, много проблем в биологической сущности человеческой патологии, исследование препаратов [Балабанова Т.Н., Трапезникова И.В., 2018] и так далее. Это помогло исследованию генетической природы некоторых наследственных болезней, врожденных уязвимых мест к болезням и прочее, связанное с мутациями.

Таким образом, в данной статье проведем анализ методов картирования генов, играющих важную роль в изучении различий в индивидуальных геномах. В основе исследования лежат разработанные математические модели для изучения свойств генетического кода. Прогнозирование, контроль и понимание возникают, главным образом, в процессе моделирования свойств генетического кода.

Подходы к моделированию структуры генетического кода

Многие ученые из различных областей занимались исследованиями структуры стандартного генетического кода (СГК). Например, впервые модель СГК была представлена физиком Джорджем Гамовым в 1954 году [Alff-Steinberger, 1969; Amirnovin, 1997]. В литературе эта модель известна как «алмазный код Гамова». Идея в том, чтобы представить кодон в форме ромба (алмаза) с нуклеотидами по четырем углам, расположенными в следующем порядке: трем вершинам присваивается случайный нуклеотид, а четвертой – нуклеотид, комплементарный к тому, что находится посередине (Рис. 1). Например, нуклеотиды 2, 3 и 4 определены случайным образом, а нуклеотид 1 комплементарен к 2. Несмотря на некоторые погрешности, алмазный код Гамова послужил отправной точкой для будущих исследований. Крик [Crick, 1968; Ardell, 1998] использовал эту модель в своих последующих работах. В последующих исследованиях Гамов [Crick, 1968; Crick, 1976; Cairns, 1998] сделал подробный обзор различных математических методов, которые он мог использовать для изучения СГК, а также определил статистические и экспериментальные методы, Методы Монте-Карло для исследований ГК [Trenchev I. et al., 2017].

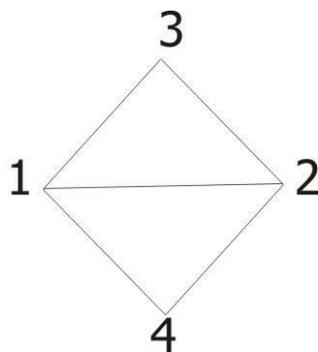


Рис. 1. Схематичное представление алмазного кода Гамова

Fig. 1. Gamow's diamond code

В целом свойства ГК изучаются в связи со средним количеством точечных мутаций и обычно делается вывод, что ГК близок к оптимальному. Оптимум некоторого критерия большинства исследований использует 2 подхода: статистический и альтернативный. Первый подход использует сравнение генетического кода со случайно сгенерированными кодами. Этот подход устанавливает, насколько оптимален ГК к некоторому предопределенному критерию. Альтернативный подход производит сравнение SGK с наиболее возможными теоретическими кодами. В большинстве случаев первый подход дает более реалистичные результаты, чем второй, потому что альтернативный подход часто использует линейные приближения и игнорирует возможные оптимальные коды. Принципиальным предметом в теории оптимизации генетических кодов является возникновение возможных изменений генетической последовательности (мутации).

В целом то, к чему приведет изменение, изучено, и возможная эволюция кода выведена. Неудобство этих двух методов в том, что, неважно, какие критерии исследования мы выбираем, коды случайным образом сгенерированы, и вероятность пропуска оптимального кода высока [Trenchev I. et al., 2017]. Например, Д. Хейг и Л. Херст [Haig, Hurst, 1991] сгенерировали 10 000 случайных ТКГ, в соответствии с некоторыми критериями, с 64 кодонами, выделенными в 21 синонимичном наборе, и тремя кодонами остановки, определенными в одном из наборов. Эти теоретические генетические коды были поставлены в сравнение к стандартному генетическому коду, и исследовалось среднее квадратичное отклонение свойств аминокислот, таких как гидрофильные свойства (R_f), изоэлектрические свойства [Gamow, 1954; Trenchev I. et al., 2017], растворимость в воде $\log\left(\frac{1-R_f}{R_f}\right)$ [Epstein, 1966; Trenchev I. et al., 2017] и молекулярная масса.

Авторы определили среднее квадратичное отклонение посредством следующей формулы:

$$\sum_{xyz \in \{A,C,G,T\}^3} \frac{(I) + (II) + (III)}{C},$$

где C – это число точечных мутаций в ТКГ без изучаемых кодонов остановки. (I), (II) и (III) определены следующим образом:

$$\sum_{x' \in \{A,C,G,T\} - \{x\}} D(c(xyz).c(x'yz)) \quad (1)$$

$$\sum_{y' \in \{A,C,G,T\} - \{y\}} D(c(xyz).c(xy'z)) \quad (2)$$

$$\sum_{z' \in \{A,C,G,T\} - \{z\}} D(c(xyz).c(xyz')) \quad (3)$$

где сумма вычислена только для кодонов, кодирующих аминокислоты, т. е. без кодонов остановки. Если X и Y – аминокислоты и $X = c(xyz)$, а $Y = c(x'yz)$, значит $D(X, Y) = [w(X) - w(Y)]^2$, $w(X)$ и $w(Y)$ – свойства аминокислот – гидрофильные свойства, молекулярная масса, изоэлектрические свойства и отношение растворимости в воде. Были вычислены 4 среднее квадратичных расстояния: полное $MS(c)$; для первого положения – $MS_1(c)$; для второго – $MS_2(c)$; и для третьего – $MS_3(c)$. Для вычисления полного расстояния использовались 4 формулы, для $MS_1(c)$ – расстояние первой позиции: первая формула для среднего расстояния и формула (1). В первой формуле C заменяется C_1 , числом точечных мутаций в первом положении. Оставшаяся пара расстояний $MS_2(c)$ и $MS_3(c)$ были

вычислены подобным образом, кроме использования формул (2) и (3), и количества точечных мутаций, приводящих ТК к лучшим индексам, чем СК относительно простых мутаций. Авторы подтвердили, что тип закодированной аминокислоты зависит от первого положения кодона. Другой положительный результат моделирования – значение третьего положения в кодоне. В данном исследовании не было подтверждено, что второе положение в кодоне описывает полярность аминокислоты. Гипотеза, что СК есть продукт выбора между очень похожими кодами в целях минимизации ошибок, не была подтверждена. Возможно, причина в том, что Д. Хейг и Л. Херст [Haig, Hurst, 1991] не использовали вероятности возникновения аминокислот в среднем наборе белка и не точно определили вероятности мутаций в различных положениях кодонов [Trenchev I. et al., 2017].

Аналогично, оптимальный код СК проанализирован в работе [Freeland, Hurst, 1998]. Был использован подход, описанный выше, но авторы ввели дополнительный критерий, определяющий природу точечных мутаций: транзигия или трансверсия, если быть точными. Авторы сгенерировали 10^6 СК и посчитали среднеквадратичное отклонение. Изучение кода проводилось в следующей манере: (1) обобщение специфической восприимчивости кода как значения ошибки кода; (2) определение возможных кодов, в которые СК может быть преобразован; (3) сравнение значений кодов и значений СК с их определенным критерием.

С помощью данных исследований они продемонстрировали, что существует только один код более высокого качества, чем СК, результаты представлены в таблице 1, данный код имеет очень хорошие характеристики в отношении точечных мутаций.

Неудача авторов заключалась в том, что они изучали коды отдельно, то есть не сравнивали ТК и СК по критерию. Они доказали в своих последующих исследованиях, в которых С. Фриланд активно принимал участие, что СК является продуктом выбора кода посредством минерализации ошибок в декодировании [Gilis, 2001]. Они изучили ГК со следующей формулой, используемой для определения индивидуальной ошибки кода для каждого кодона:

$$\Delta_j = \frac{\sum_{i=1}^{210} (w\alpha_i + \beta_i)\varepsilon_i}{\sum_{i=1}^{210} (w\alpha_i + \beta_i)},$$

$$j = 1..64$$

где ε_i – вес, связанный с соответствующей исследуемой аминокислотой, и может быть вычислен с помощью матрицы мутаций Pam_{74-100} . Другой путь вычисления ε_i – рассматривать его как произведение вычитания двух различных гидрофобий аминокислот в результате изменения соответствующего кодона. W – это вес, число транзигий и трансверсий, зависящих от положения кодона. α_i, β_i – число транзигий и трансверсий, в зависимости от замещения в конкретной аминокислоте «другим», образованным «мутировавшим» кодоном. Например, замена аминокислоты Ile на Met $AUG \rightarrow AUA$ $\alpha_{Met \rightarrow Ile} = 1$, $AUG \rightarrow AUY$ $\beta_{Met \rightarrow Ile} = 2$. Авторы подтвердили вывод из предыдущего исследования, что вероятность появления нового кода с лучшими свойствами, чем СК, составляет 10^{-6} . Еще одним успехом стало их исследование оптимального СК с точки зрения его «адаптивности» к минимизации ошибок декодирования. Авторы сделали подробный анализ своих выводов, но только через призму гипотез, практически не проверенных. Другим недостатком подхода С. Фриланда [Freeland, Hurst, 1998] является то, что он сделал вполне определенные выводы, например, что он «изучил все возможные ТК», эволюция кодона может быть предсказана и т. д. (Таблица 1).

Количественная мера была использована в исследованиях Фриланда и др. в 1999 году, введенных Голдманом [Goldman, 1993] для изучения SGK, но выводы не доказали гипотезы происхождения ГК его эволюции.

Таблица 1
Table 1

Улучшенные теоретические генетические коды Фриланда в сравнении со стандартным генетическим кодом

Freeland's improved theoretical genetic codes compared to the standard genetic code

Первая позиция	Вторая позиция				Третья позиция
	U	C	A	G	
U	Ile	Ala	Gln	His	U
	Ile	Ala	Gln	His	C
	Cys	Ala	Stop	Stop	A
	Cys	Ala	Stop	Gly	G
C	Cys	Leu	Thr	Ser	U
	Cys	Leu	Thr	Ser	C
	Cys	Leu	Phe	Ser	A
	Cys	Leu	Phe	Ser	G
A	Trp	Pro	Asp	Ala	U
	Trp	Pro	Asp	Ala	C
	Trp	Pro	Glu	Ser	A
	Val	Pro	Glu	Ser	G
G	Tur	Met	Asn	Arg	U
	Tur	Met	Asn	Arg	C
	Tur	Met	Lys	Arg	A
	Tur	Met	Lys	Arg	G

Гидрофильные/гидрофобные свойства SGK рассчитывали в исследовании [Giulio, 1997], и было доказано, что они на 68 % лучше, чем полученные им TГК, с использованием шкалы, определенной Гумбелом и др. [Gumbel, 2015]. Он использовал аналогичный подход к вышеупомянутому, однако среднеквадратичное отклонение было вычислено с помощью иной формулы. Наименьшее значение было рассчитано на основе минимизации G-функции методом Лагранжа.

В данном исследовании авторы пришли к выводу, что SGK на 68 % оптимизирован и гидрофильные свойства аминокислот имеют, возможно, значимую роль в эволюции SGK. Подход [Giulio, 1997] был лучше, т. к. по сравнению с Фриландом [Freeland, Hurst, 1998] большее количество TГК было сгенерировано. Здесь можно отметить тот же недостаток: автор не использовал частоту возникновения аминокислоты в усредненном белке.

Среднеквадратичное отклонение Хейга и Херста [Haig and Hurst, 1991; Knight et al., – 1999] было использовано в исследовании Н. Голдмана [Goldman, 1993] для изучения SGK. Он сохранил силу синонимических множеств для генерации TГК, как в SGK, то есть 3 аминокислоты кодируются 6 кодонами, 4 аминокислоты – 4 кодонами, 1 аминокислота – 3 кодонами, 9 аминокислот – 2 кодонами, 2 аминокислоты – 1 кодоном и 3 стоп-кодонами. Голдман [Goldman, 1993] пришел к выводу на основе своей модели, что основной тенденцией в развитии SGK была простая мутационная ошибка минимизации, но он не сделал никаких выводов о мощности синонимического набора.

Джулио [Giulio, 1997] всерьез раскритиковал работы С. Фриланда и его команды [Knight R. et al., 1999; Lin K. et al., 2001] и не одобрил подход к изучению ГК. Его тезис заключался в том, что количество TГК намного больше, чем генерируемое С. Фриландом, и

выводы, сделанные в отношении небольшого числа генерируемых ТГК, являются ненадежными [Kuruoglu E. E., Arndt P. F., 2017]. В своих исследованиях он доказал, что число ТГК составляет около 10^{84} , и гораздо больше, чем количество, изученное С. Фриландом в его исследованиях, имеется в виду 10^{18} [Forger M., Sachse S., 2000; Lenstra R., 2014; Nemzer L.R., 2017], и дополнительно доказал, что возможные ТГК составляют около 270 млн.

С другой стороны, Фриланд и Джулио сделали свои выводы по репрезентативной выборке набора ТГК, а не работая над всем набором. Они утверждали, что вероятность нахождения лучшей ТГК с лучшими свойствами, чем СГК, удовлетворяющей определенным критериям, составляла от 10^{-6} до 10^{-9} [Rodin A.S., Branciamore S., 2013]. То, что разделяют оба автора, это предложение о том, что вероятность не может быть установлена с уверенностью, потому что невозможно изучить весь набор ТГК. Таким образом, обсуждение четкого описания всего набора ТГК остается открытым.

Заключение

Данный обзор позволяет оценить стремительное развитие молекулярных биотехнологий и разнообразие методов, применяемых в исследованиях генетического кода. Многие ученые обратились к общим исследованиям о происхождении живых организмов и структуре современного генетического кода. Это привело к созданию разнообразных эволюционных моделей с применением математических подходов.

Стоит отметить, что в данном исследовании мы рассмотрели подходы к моделированию свойств генетического кода с использованием теоретических генетических кодов в сравнении со стандартным генетическим кодом. Другие авторы также рассматривают свойства вырожденности генетического кода, помехоустойчивости, опираясь на исследование некоторых необычных способов записи генов, кодирующих белки. В природе один и тот же участок ДНК может кодировать несколько белков, однако в природе действует запрет на альтернативные чтения, которые могут привести к мутациям. Подобные гены называются перекрывающимися, и суть исследований состоит в том, чтобы понять, как природе удалось создать столь совершенный код.

Например, Н.Н. Козлов показал, что перекрытия генов (замена кодона) хоть и не влияют на тип кодируемого белка, но могут изменить его свойства. Им были установлены интегральные характеристики генетического кода, которые использует природа для выполнения двух функций: перекрытий пар генов, а также блокировки генов.

На основе исследований похожих свойств генетических кодов С.В. Петухов показывает, что природа обладает особыми свойствами упорядоченности, использующими неизвестные виды алгебр, а создание нового понятийного аппарата и «правильных» методов позволит решить задачи понимания и исследования живого вещества.

Список литературы

1. Балабанова Т.Н., Трапезникова И.В. 2018. Синтез панорамных изображений и их использование в цитогенетических исследованиях. Научные ведомости БелГУ, 45: 760–768.
2. Петухов С.В. 2008. Матричная генетика, алгебры генетического кода, помехоустойчивость. М.; Ижевск: Регулярная и хаотическая динамика, 316 с.
3. Ahmad, M., Jung, L., and Bhuiyan A. 2017. From DNA to protein: Why genetic code context of nucleotides for DNA signal processing? Biomedical Signal Processing and Control, 34: 44–63.
4. Alff-Steinberger, C. 1969. The genetic code and error transmission. Natl. Acad. Sci. USA, 64: 584–591.
5. Amirnovin, R. 1997. An analysis of the metabolic theory of the origin of the genetic code. J Mol Evol, 44: 473–476.
6. Ardell D.H. 1998. On error minimization in a sequential origin of the standard genetic code. Journal of molecular evolution, 47: 1–13.
7. Cairns, J., Overbaugh, J. and Miller, S. 1998. The origin of mutants. Nature (London), 335: 142–145.

8. Crick, FHC. 1968. The origin of the genetic code. *J Mol Biol.*, 38: 367–379.
9. Crick, FHC., et al. 1976. A speculation on the origin of protein synthesis. *Orig. Life.*, 7: 389–397.
10. Epstein, CJ. 1966. Role of the amino-acid “code” and of selection for conformation in the evolution of proteins. *Nature*, 210: 25–28.
11. Freeland S.J., Hurst L.D. 1998. The genetic code is one in a million. *Journal of molecular evolution*, 47: 238–248.
12. Forger M., Sachse S. 2000. Lie superalgebras and the multiplet structure of the genetic code. I. Codon representations. *Journal of Mathematical Physics*, 41: 5407–5422.
13. Gamow G., Rich A., Yčas M. 1956. The problem of information transfer from the nucleic acids to proteins. *Advances in biological and medical physics.* – Elsevier, 4: 23–68.
14. Gamow, G. 1954. Possible mathematical relation between deoxybonucleic acid and proteins. *Det Kongelige Danske Videnskabernes. Biologiske Meddelelser*, 22:1–13.
15. Gilis D. et al. 2001. Optimality of the genetic code with respect to protein stability and amino-acid frequencies. *Genome biology*, 2: research0049. 1
16. Giulio, MD. 1997. On the origin of the genetic code. *J. Theor. Biol.*, 187: 573–581.
17. Giulio, D. 2013. The origin of the genetic code in the ocean abysses: New comparisons confirm old observations. *Journal of Theoretical Biology*, 333: 109–116.
18. Goldman N. 1993. Further results on error minimization in the genetic code. *Journal of molecular evolution*, 37: 662–664.
19. Gumbel M. et al. 2015. On models of the genetic code generated by binary dichotomic algorithms. *Biosystems*, 128: 9–18.
20. Haig, D., and Hurst, LD. 1991. A quantitative measure of error minimization in the genetic code. *J. Mol Evol.*, 33: 412–417.
21. Knight R.D., Freeland S.J., Landweber L.F. 1999. Selection, history and chemistry: the three faces of the genetic code. *Trends in biochemical sciences*, 24: 241–247.
22. Kuruoglu E.E., Arndt P.F. 2017. The information capacity of the genetic code: Is the natural code optimal? *Journal of theoretical biology*, 419: 227–237.
23. Lenstra R. 2014. Evolution of the genetic code through progressive symmetry breaking. *Journal of theoretical biology*, 347: 95–108.
24. Lin K., May A. C. W., Taylor W. R. 2001. Amino acid substitution matrices from an artificial neural network model. *Journal of Computational Biology*, 8: 471–481.
25. Nemzer L.R. 2017. A binary representation of the genetic code. *Biosystems*, 155: 10–19.
26. Pelc S.R. 1965. Correlation between coding-triplets and amino-acids. *Nature*, 207: 597–599.
27. Rodin A.S., Branciamore S. 2013. The Universal Genetic Code and Non-Canonical Variants. *Brenner's Encyclopedia of Genetics*, 263–264.
28. Trenchev I. et al. 2017. Mathematical Models for Studying the Properties of the Genetic Code. *Der Pharmacia Lettre*, 9 [8]: 40–54.

References

1. Balabanova T.N., Trapeznikova I.V. 2018. Synthesis of panoramic images and their use in cytogenetic studies. *Belgorod State University Scientific Bulletin*, 45: 760–768. (in Russian)
2. Petoukhov S.V. 2008. Matrix genetic, genetic code algebra, noise immunity. M.; Izhevsk: *Regulyarnaya I khaoticheskaya dinamika*, 316 c. (in Russian)
3. Ahmad, M., Jung, L., and Bhuiyan A. 2017. From DNA to protein: Why genetic code context of nucleotides for DNA signal processing? *Biomedical Signal Processing and Control*, 34: 44–63.
4. Alff-Steinberger, C. 1969. The genetic code and error transmission. *Natl. Acad. Sci. USA*, 64: 584–591.
5. Amirnovin, R. 1997. An analysis of the metabolic theory of the origin of the genetic code. *J Mol Evol*, 44: 473–476.
6. Ardell D.H. 1998. On error minimization in a sequential origin of the standard genetic code. *Journal of molecular evolution*, 47: 1–13.
7. Cairns, J., Overbaugh, J. and Miller, S. 1998. The origin of mutants. *Nature (London)*, 335: 142–145.
8. Crick, FHC. 1968. The origin of the genetic code. *J Mol Biol.*, 38: 367–379.
9. Crick, FHC., et al. 1976. A speculation on the origin of protein synthesis. *Orig. Life.*, 7: 389–397.
10. Epstein, CJ. 1966. Role of the amino-acid “code” and of selection for conformation in the evolution of proteins. *Nature*, 210: 25–28.

11. Freeland S.J., Hurst L.D. 1998. The genetic code is one in a million. *Journal of molecular evolution*, 47: 238–248.
12. Forger M., Sachse S. 2000. Lie superalgebras and the multiplet structure of the genetic code. I. Codon representations. *Journal of Mathematical Physics*, 41: 5407–5422.
13. Gamow G., Rich A., Yčas M. 1956. The problem of information transfer from the nucleic acids to proteins. *Advances in biological and medical physics*. – Elsevier, 4: 23–68.
14. Gamow, G. 1954. Possible mathematical relation between deoxybonucleic acid and proteins. *Det Kongelige Danske Videnskabernes. Biologiske Meddelelser*, 22:1–13.
15. Gilis D. et al. 2001. Optimality of the genetic code with respect to protein stability and amino-acid frequencies. *Genome biology*, 2: research0049. 1
16. Giulio, MD. 1997. On the origin of the genetic code. *J. Theor. Biol.*, 187: 573–581.
17. Giulio, D. 2013. The origin of the genetic code in the ocean abysses: New comparisons confirm old observations. *Journal of Theoretical Biology*, 333: 109–116.
18. Goldman N. 1993. Further results on error minimization in the genetic code. *Journal of molecular evolution*, 37: 662–664.
19. Gumbel M. et al. 2015. On models of the genetic code generated by binary dichotomic algorithms. *Biosystems*, 128: 9–18.
20. Haig, D., and Hurst, LD. 1991. A quantitative measure of error minimization in the genetic code. *J. Mol Evol.*, 33: 412–417.
21. Knight R.D., Freeland S.J., Landweber L.F. 1999. Selection, history and chemistry: the three faces of the genetic code. *Trends in biochemical sciences*, 24: 241–247.
22. Kuruoglu E.E., Arndt P.F. 2017. The information capacity of the genetic code: Is the natural code optimal? *Journal of theoretical biology*, 419: 227–237.
23. Lenstra R. 2014. Evolution of the genetic code through progressive symmetry breaking. *Journal of theoretical biology*, 347: 95–108.
24. Lin K., May A. C. W., Taylor W. R. 2001. Amino acid substitution matrices from an artificial neural network model. *Journal of Computational Biology*, 8: 471–481.
25. Nemzer L.R. 2017. A binary representation of the genetic code. *Biosystems*, 155: 10–19.
26. Pelc S.R. 1965. Correlation between coding-triplets and amino-acids. *Nature*, 207: 597–599.
27. Rodin A.S., Branciamore S. 2013. The Universal Genetic Code and Non-Canonical Variants. *Brenner's Encyclopedia of Genetics*, 263–264.
28. Trenchev I. et al. 2017. Mathematical Models for Studying the Properties of the Genetic Code. *Der Pharmacia Lettre*, 9 [8]: 40–54.

Ссылка для цитирования статьи For citation

Гарянина А.И., Червяков Н.И. 2020. Обзор методов математического моделирования свойств генетического кода. *Экономика. Информатика*. 47 (2): 372–379. DOI: 10.18413/2687-0932-2020-47-2-372-379.

Garianina A.I., Chervyakov N.I. 2020. A review of mathematic modeling methods of genetic code properties. *Economics. Information technologies*. 47 (2): 372–379 (in Russian). DOI: 10.18413/2687-0932-2020-47-2-372-379.