

## Геном как (гипер)текст: от метафоры к теории

С. Т. Золян<sup>1</sup>, Р. И. Жданов<sup>2</sup>

<sup>1</sup> ИНСТИТУТ ФИЛОСОФИИ, СОЦИОЛОГИИ И ПРАВА НАН РЕСПУБЛИКИ АРМЕНИЯ

<sup>2</sup> МОСКОВСКИЙ ПЕДАГОГИЧЕСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИНСТИТУТ ПЕРСПЕКТИВНЫХ ИССЛЕДОВАНИЙ  
ИНСТИТУТ ФУНДАМЕНТАЛЬНОЙ МЕДИЦИНЫ И БИОЛОГИИ  
КАЗАНСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ

*Аннотация.* С самого начала зарождения генетики было осознано глубинное сходство между языком и процессами обработки и передачи генетической информации. Были предприняты попытки использовать выработанные в лингвистике методы для дешифровки кодирующих белки нуклеотидных последовательностей. Малая информативность такого рода расшифровок была вполне предсказуемой – методики, ограничивающиеся процессами кодирования, в лучшем случае позволяют составить словарь единиц, но не в состоянии описать язык и выявить определяющие для коммуникации структуры смысла и текста. Так, до сих пор остаются невыясненными функции порядка 95 % не кодирующих последовательностей ДНК. Только текст (а не знак) может рассматриваться как объект процесса создания – передачи – сохранения – преобразования информации. Общая теория текста должна быть в состоянии описать языковые тексты и процессы их структурирования, функционирования и трансформации, и в то же время представить первичную структуру генома как систему текстов – гипертекст, состоящий из упорядоченного подмножества других текстов. Текст может быть рассмотрен как квазиорганизм, обладающий памятью, креативно-когнитивными характеристиками и коммуникативным потен-

*Золян С. Т., Жданов Р. И.* Геном как (гипер)текст: от метафоры к теории // Критика и семиотика. 2016. № 1. С. 60–84.

ISSN 2307-1737. Критика и семиотика. 2016. № 1  
© С. Т. Золян, Р. И. Жданов, 2016

циалом, а клетка (рибосома) – как обладающая квазиинтеллектом и способностью оперировать абстрактными семиотическими последовательностями. Такой подход требует конструирования новой версии семиотики – мультимодальной и текстоцентричной, где исходными базовыми понятиями являются текст и контекст, а знаки и смыслы рассматриваются одновременно и как контекстно-зависимые переменные, и как изменяющие контекст операторы.

*Ключевые слова:* геном, генетический код, общая теория текста, геном как (гипер-)текст, биосемиотика.

УДК 81'22 / 81'42 / 575.1 / 631.523.13

*Контактная информация:* Золян Сурен Тигранович, доктор филологических наук, профессор, ведущий научный сотрудник ИФСП НАН РА, Ереван (ул. Арами, 44, Ереван, 0010, Армения, [surenzolyan@gmail.com](mailto:surenzolyan@gmail.com))

Жданов Ренад Ибрагимович, доктор химических наук, профессор, профессор КФУ, Казань; главный научный сотрудник Института перспективных исследований МПГУ (пр. Вернадского, 88, Москва, 119571, Россия, [zrenad@gmail.com](mailto:zrenad@gmail.com))

### **Геном человека как (гипер)текст К постановке проблемы**

Цель данной статьи – рассмотреть возможности, форму и принципы описания генома как текста. Молекулярную генетику принято определять как раздел биологии, изучающий на молекулярном уровне закономерности и механизмы хранения, передачи, реализации и изменения генетической информации. Очевидно, что уже с самого начала зарождения генетики интуитивно ощущалось глубинное сходство между языком и процессами передачи генетической информации. О возможности рассматривать ген и геном как текст или язык говорилось неоднократно, но скорее это были метафоры или даже омонимы: используя такие термины, как код, информация, язык, словарь, алфавит, семантика, транскрипция, чтение, полисемия, транскрипт, редактирование и т. п., обычно игнорировалось их собственно лингвистическое содержание (ср.: [Raible, 2001], где приведена статистика подобных метафорических употреблений в научных статьях по генетике, а также: [Emmeche, Hoffmeyer, 1991; Седов, 2000; 2001]).

Между тем, аналогии между структурами языка и генома гораздо глубже, чем возможность метафорического употребления лингвистической терминологии. Как отмечал еще Роман Якобсон, «вопрос об этих изоморфных характеристиках (языка и генетического кода. – С. З., Р. Ж.) становится особенно показательным, когда мы осознаем, что они не имеют аналогов ни в одной из систем коммуникации в животном мире. Генетиче-

ский код, первичное проявление жизни, и язык как всеобщее достояние человечества, обеспечивший его исключительный скачок от генетики к цивилизации – это два фундаментальных хранилища передаваемой от предков к потомкам информации. Молекулярная наследственность (heredity) и вербальное наследие (legacy) – необходимое предусловие культурной традиции» [Jakobson, 1970, p. 440].

Если исходить из того, что биомолекулярные процессы есть особый тип информационных процессов, то из этого следует, что они должны разделять наиболее существенные характеристики, общие для всех видов коммуникации. Одним из таких необходимых для осуществления коммуникации свойств является организация сигналов в тексты. При коммуникации имеет место структурирование информации на основе определенных функционально-смысловых характеристик; исходные единицы образуют значимые конфигурации, композиция которых формирует текст. Информативной единицей коммуникации, будь то социальная или биологическая, будет не отдельный сигнал, а некоторая целостная структура, текст. Именно текст есть форма организации информации, если исходить из ее качественных смысловых характеристик (в отличие от основанной на чисто количественных параметрах теории информации К. Шеннона). Речь может идти о единой основе механизмов передачи смысла, а формой и результатом этого процесса является текст, который может быть выражен как в знаках языка, так и в иных носителях информации, в том числе и биологических (генетических). Как мы попытаемся показать, в качестве такой общей основы может быть рассмотрена организация информации в форме текстов. Сама идея о текстуальной структурированности генетической информации не нова, однако лингвистические механизмы подобного структурирования и их экспликация ранее не рассматривались.

В свое время (1968) диалог двух выдающихся ученых – лингвиста Романа Якобсона и первооткрывателя информационной РНК, Нобелевского лауреата Франсуа Жакоба породил много идей, которые до сих пор не реализованы в полной мере. Одна из них – то, что вследствие глубинного сходства «...понимание одной из систем может помочь в анализе другой. Однако этот процесс носит односторонний характер, поскольку до сих пор генетика, кажется, не внесла ощутимого вклада в прогресс лингвистики. Лингвистика же, напротив, предложила генетике совершенную модель. Если лингвистическая модель доказала свою исключительную ценность при молекулярном анализе наследственности, то это, возможно, потому, что она в равной степени применима как к структуре, так и к функциям генетической материи. Нечасто модель, навязанная концепциями своего времени, находила столь точное применение» [Жакоб, 1992, с. 139–140]. Разумеется, речь идет не о решении генетических проблем методами лингвистики, а именно о моделировании (о том, что «понимание одной из систем может помочь в анализе другой»). Лингвистическая модель не может

иметь доказательной силы для генетики (в конечном счете здесь последнее слово остается за экспериментом, как и наоборот: в лингвистике объяснение экспериментальным данным не может быть дано вне теоретической модели), но может иметь объяснительную. При таком моделировании остаются в стороне вопросы, связанные с химическими и физическими механизмами обеспечения информационных процессов, но зато эксплицируются структурно-семиотические (подобно тому, как в лингвистических исследованиях, как правило, не рассматриваются физиологические или нейролингвистические механизмы языковой деятельности).

### **Об изоморфизме информационных процессов**

Идея о том, что различные процессы, протекающие, с одной стороны, в живой природе, а с другой – в обществе и человеческом сознании, могут быть описаны как процессы создания, передачи и хранения информации, имеет весьма давнюю историю. Еще с античных времен используются устойчивые метафоры, отождествляющие мир и природу с книгой и словом (языком). В новое время эта идея получает научное оформление. Появляются теории, предлагающие единообразное научное описание форм организации как природных явлений, так и социальных и интеллектуальных. Одно из первых – это учение о формах великого поэта и естествоиспытателя И.-В. Гете («Опыт объяснения метаморфоза растений», 1790 г.), благодаря которому появляется и сам термин – морфология. Хотя это учение было предназначено для описания природных объектов, прежде всего биологических, очень скоро оно было перенесено в сферу гуманитарных наук. В частности, в XIX в. оно дало имя (если не жизнь) новому разделу языкознания – морфологии языка, а в XX в. вдохновило на пионерское описание текста-нарратива как трансформации единой глубинной структуры [Пропп, 1928]. В XX в. системно-структурные методы становятся доминирующими при изучении как организмов, так и социальных систем. Начало этому положили теория языка Ф. де Соссюра, которая в дальнейшем становится методологической основой для всех гуманитарных наук, а также разработанная на основе биологии общая теория систем [Bertalanffy, 1932; 1968]. В 40–50-х гг. XX в. общая теория систем получает мощный дополнительный импульс в связи с развитием кибернетики, будучи распространена и на информационные системы и процессы [Shannon, 1948; Wiener, 1948]. В конце XX в., с одной стороны, в теории Н. Лумана [Luhmann, 1990; 1995], с другой – Х. Матурана и Ф. Варела [Maturana, Varela, 1980] происходит своеобразный синтез всех трех направлений – лингвистического, биологического и кибернетического. Так, было продемонстрировано, что организация сложных систем (автопоэтические или саморефлективные системы) не исчерпывается саморегуляцией и целепо-

лаганием, но предполагает также рефлексию и осмысление, т. е. текстурально-семантические процессы.

### Генетический код как знаковая система

Таким образом, как то и предполагалось Романом Якобсоном, есть основания говорить о единой информационной основе биологических (прежде всего генетических), социальных и когнитивных процессов. При этом речь может идти не только об общих закономерностях, присущих всем типам систем, но и о более специфических аналогиях, характеризующих знаковые процессы. Для их описания естественным было обращение к дисциплине, изучающей общие закономерности передачи информации посредством знаковых структур – семиотике. Так, говоря о сущности выдающегося открытия генетики XX в., Ф. Крик описывает его как семиотический процесс соотнесения двух языков: «Обнаружение генетического кода действительно является великим достижением. Это, в определенном смысле, ключ к молекулярной биологии, потому что показывает, как связаны друг с другом два великих полимерных языка: язык нуклеотидов и язык белков» [Crick, 1966, p. 9].

Впрочем, В. В. Иванов справедливо считает, что использование принципов семиотического анализа даже предшествовало фундаментальному открытию кода ДНК: «В середине XX века дешифровка генетического кода оказалась возможной, помимо всех других подготовивших ее открытий, еще и потому, что еще до окончательного успеха в описании двойной спирали Криком и Уотсоном была предложена в отчетливом виде Гамовым, модель, построенная по аналогии с дешифровкой языкового текста» [Иванов, 2010, с. 44]. Идея о знаковой природе гена была очевидна еще для первооткрывателя ДНК Фридриха Мишнера (1869), который позднее (1892) высказал идею о том, что «генетическая информация может существовать в виде молекулярного текста, линейной последовательности химических символов», а несколько малых химических единиц в составе крупных молекул могут играть наследственную роль, подобно тому, «как все слова и понятия всех языков могут быть выражены посредством всего лишь от 24 до 30 букв алфавита» [Trifonov, 2000, p. 5].

Вполне естественно, что еще полвека назад, синтез семиотики и биологии привел к рождению новой дисциплины – биосемиотики (сам термин впервые появляется в: [Rotshild, 1962]). За это время сформировались различные направления и ведущие школы, был достигнут ряд важных результатов<sup>1</sup>. Параллельно с биосемиотикой и частично совпадая с ней, развива-

<sup>1</sup> Подробный обзор истории и основных направлений биосемиотики см.: [Stjernfelt, 2002; Barbieri, 2007; 2008a; 2008b; 2009; Pattee, 2005; Favareau, 2007; Witzany, 2007; Kull et al., 2009; Galik, 2013], а также: [Степанов, 1971; Чебанов, 2004].

ется биоинформатика [Elumalai, Eswaraiyah, 2013], а в последнее время также и биолингвистика [Augustyn, 2013, 2013; Bickerton, 2014], биосемантика, биогерменевтика и даже «протеиновая лингвистика» [Gimona, 2008], где также были получены интересные результаты<sup>2</sup>. Однако, на наш взгляд, до сих пор лишь в малой степени прояснены именно те аспекты биосемиотики, где, казалось бы, аналогии с языком и его структурами наиболее очевидны, а именно генетики (ср.: [Якобсон, 1985, Гамкрелидзе, 1988; Gamkrelidze, 2009; Гельфанд, 1990; Жакоб, 1992]).

Возможно, это было связано с тем, что в биосемиотику был перенесен и характерный для семиотики первой половины XX в. «знакоцентризм», причем в его наиболее последовательной версии, идущей от Чарльза Пирса, согласно которой знак определяется механизмами его интерпретации. Семиотическая теория Фердинанда де Соссюра, в которой знак определяется в первую очередь посредством включающей и конституирующей его системы, насколько мы можем судить, оказала на биосемиотику куда меньшее воздействие, хотя, казалось бы, она адекватнее отображает характер генетического семиозиса<sup>3</sup>.

Парадоксально, но не была в должной мере использована и ключевая для языкознания XX в. дихотомия языка и речи, т. е. разграничение между некоторой абстрактной системой и реализацией этой системы в конкретном контексте (в генеративной лингвистике эта идея Соссюра получает свое развитие как дихотомия глубинных и поверхностных уровней, или структур). Между тем дихотомия языка и речи находит прямое соответствие в организации генетических процессов. Здесь уместно вспомнить, что еще в 1944 г., задолго до открытия генетического кода, Э. Шредингер предлагал рассматривать его одновременно и как систему, и как процесс актуализации этой системы: «Но термин шифровальный код, конечно, слишком узок. Хромосомные структуры служат в то же время и инструментом, осуществляющим развитие, которое они же и предвещают. Они являются и кодексом законов, и исполнительной властью, или, употребляя другое сравнение, они являются одновременно и архитектором, и строителем» [Шредингер, 2002, с. 28]<sup>4</sup>.

---

<sup>2</sup> Сопоставление методологии и концептуальных принципов этих все еще находящихся в стадии становления дисциплин выходит за рамки данной статьи, тем более что из комплекса биосемиотики нас интересуют лишь аспекты, связанные с генетическим кодом.

<sup>3</sup> Ср.: «Можно сказать, что среди всех систем передачи информации только генетический код и языковой код базируются на использовании дискретных компонентов, которые сами по себе не имеют смысла, но служат для построения минимальных единиц, имеющих смысл, т. е. сущностей, наделенных собственным смыслом в данном коде» [Якобсон, 1985, с. 393].

<sup>4</sup> В данном случае не столь важно, что Шредингер ошибочно предполагал, что генетический код соотносится с хромосомой, а не с ДНК – РНК. Дальнейшие от-

В целом в генетике, проводя аналогии с языком, лишь в малой степени учитывали данные лингвистики. Для большинства генетиков язык состоит из... букв, т. е. из того, что, строго говоря, вовсе не является языковой единицей, а есть лишь средство графической фиксации единиц языка на письме. Но даже без должного учета данных лингвистики проделанные в рамках биоинформатики описания генетической информации как допускающих формализацию языков и языковых структур приводили к нетривиальным результатам [Searls, 1993; 1999; 2002; 2010; Ji, 1997; 1999]. Безусловно, учет принципов организации не только формальных, но и естественных языков и языковых структур мог бы существенно обогатить подобные модели.

Однако, несмотря на первоначальный оптимизм, в целом предпринятые в основном в 1980–1990-е гг., попытки применить лингвистические методы для описания информационных процессов в молекулярной генетике не дали существенных результатов. Характерное для первооткрывателей генетического кода цельное представление об общности семиотических и генетических процессов, так и не получив адекватной теоретической экспликации, с течением времени становится все менее четким, нередко возникает также и смешение языков описания этих процессов (например, семиотические структуры описываются средствами представления биохимических реакций и т. п.)<sup>5</sup>. Поэтому не случайно, что не привели к каким-либо значимым результатам попытки использовать при дешифровке белковых генов методы, при которых последовательности нуклеотидных оснований рассматриваются как последовательности букв в словах. Малая информативность такого рода декодирования была предсказуемой – по такой методике, в лучшем случае, можно составить словарь, но не описать язык: такими методами невозможно выявить определяющие для коммуникации структуры смысла и текста.

Связанные с кодированием / декодированием информационные процессы – лишь малая часть семантики генома. Семантические аспекты генетической организации не ограничиваются кодированием, – об этом свидетельствует то, что до сих пор недостаточно ясны функции порядка 95 %

---

крытия сделали это предвидение Шредингера даже еще точнее: используя ключевую для лингвистики дихотомию, можно идентифицировать триплеты ДНК с системой (языком, программой), в то время как цепочки РНК явятся речью, манифестацией программы и будут описаны как текст.

<sup>5</sup> Франсуа Жакоб имел основания сетовать: «В биологии существует масса обобщений, но всего несколько настоящих теорий... Когда в поле зрения оказывается какой-либо абстрактный элемент типа гена, биологи не успокаиваются до тех пор, пока не заменят абстрактную единицу материальными компонентами, частицами или молекулами, словно теория, чтобы играть в биологии какую-то роль, должна прежде всего оперировать какой-либо конкретной моделью» [Жакоб, 1990, с. 139].

некодирующих ДНК последовательностей генома человека. Между тем хотя бы часть тех нуклеотидных последовательностей, функции которых не связаны с кодированием или не совсем понятны (это так называемые «бессмысленные кодоны» или же, например, интроны), могут быть вполне естественно описаны как средства текстовой организации: они выступают либо как маркеры делимитации, обозначающие начало или конец сегмента, либо как так называемые дискурсивные частицы (слова-«паразиты»), которые исчезают из окончательного текста и т. п.

В последнее время чтобы расширить понимание языка генома [Материалы..., 2014], вводятся некоторые новые идеи, в частности язык модификации гетероциклических оснований нуклеотидов генома (метилирование, фосфорилирование) [Ванюшин, 2006], что уже привело к возникновению эпигенетики, или липидного кодирования генома [Zhdanov, Ibragimova, 2014]. Однако, наряду с освоением новых областей, необходимо вернуться к уточнению самих методологических основ. Сложность проблемы требует привлечения новых подходов и перехода от вероятностных моделей к структурным, от описания статичной системы к ее рассмотрению в процессе динамического функционирования, от комбинаторики к структурам смысла и текста. Как нам представляется, лишь основанные на понятиях смысла и текста теории могут стать интеграторами структурного описания систем, различных по своей природе и сферам функционирования. На наш взгляд, именно указанное методологическое ограничение, при котором ген и геном рассматриваются не как целостная структура, а только как линейная последовательность элементов, не позволило перейти от фиксации некоторых общностей к общезначимым фундаментальным результатам (см. следующую главу).

### **Геном как (гипер)текст**

Процесс передачи информации обычно представляют как передачу знаков или сигналов. Но это поверхностный взгляд, поскольку единицей коммуникации является текст. Кроме того, передаваемые тексты могут и не иметь знаковой природы (например, кинематограф или живопись не используют готовые знаки из заданного алфавита). Хотя Ф. Крик при изложении теории генетического кода пишет, что «генетический код – это небольшой словарь, который устанавливает связь между языком нуклеиновых кислот из четырех букв и языком белков из двадцати букв» [Crick, 1981, p. 171], необходимо представлять, что адекватный перевод есть не сопоставление слов из словаря одного языка со словами из словаря другого, а соотнесение контекстов, в которых они встречаются. Соответственно слово может быть переведено не как изолированная единица, а только

в контексте. Так называемое бесконтекстное значение слова – это, как правило, его значение при употреблении в наиболее характерном контексте<sup>6</sup>.

Только текст можно рассматривать как объект создания – передачи – сохранения – преобразования информации. Отдельный знак потенциально может выражать «и все, и ничего», но он остается безгласным пока не будет употреблен в некотором тексте и в определенном контексте. Можно вспомнить Людвига Витгенштейна: «3.3. Только предложение имеет смысл; только в контексте предложения имя обладает значением. 4.23. Имя выступает в предложении только в контексте элементарного предложения» («Трактат»). Эта идея значима и для биосемиотики: «Знак, хотя и является абсолютно необходимым элементом любой семиотической системы, тем не менее не может считаться основополагающей семиотической единицы, поскольку знак не может существовать в качестве отдельного знака – знак всегда является частью более крупной системы, знак всегда находится в сочетании с другим знаком (знаками). Основная семиотическая единица – это биотекст. Любой текст есть композиция знаков, однако знаки есть не более чем функциональные части текста, которые не могут существовать без текста или вне его» [Kull, 2002, p. 329–330].

Создание основанной на понятии текста новой лингвистики и семантики генома – исключительно актуальная задача для генетики. Можно предположить, что генетические тексты состоят не из четырех исходных единиц (для ДНК это нуклеотидные основания аденозин, гуанозин, цитидин, а также тимин, причем в РНК тимин заменяется урацилом), но из намного большего числа: за счет сочетаемости первичных единиц конструируются словосочетания, предложения и, в итоге, субтексты. В этом отображается фрактальная иерархическая организация генетических текстов как аналога естественно-языковых текстов. Аналогия с процессами фрактализации (контекстуальные и интертекстуальные связи) может способствовать пониманию процессов семантизации текстов (в их природе распадаться в зависимости от контекста на субтексты и в то же время формировать сверхтексты, в обоих случаях порождая новую информацию). А помимо «классического» генетического кода, благодаря которому происходит перевод единиц «языка нуклеотидов на язык белков» (Ф. Крик), за последние тридцать лет было выявлено еще одиннадцать «кодов»; все они были названы их открывателями «вторыми» [Trifonov, 2008; 2011]. Такая «поликодовость» также является серьезным аргументом в пользу текстуальной организации генетической информации. Как было продемонстрировано еще в 1980-х гг. Ю. М. Лотманом [1981; 1992] и М. Ю. Лотманом [1985],

---

<sup>6</sup> Ср.: «В абсолютной изоляции ни один знак не имеет какого-либо значения; любое знаковое значение возникает в контексте, под которым мы понимаем ситуационный или эксплицитный контекст, неважно какой, поскольку в неограниченном или продуктивном тексте (живом языке) мы всегда можем превратить *ситуационный контекст* в *эксплицитный контекст*» [Ельмслев, 1960, с. 303–304].

текст, в отличие от других единиц языка (слово, предложение), в семиотическом отношении принципиально гетерогенен, он является сообщением, которое многократно закодировано, причем посредством как минимум двух различных знаковых систем.

Не следует, однако, думать, что, заменив слово «язык» или «код» термином «текст», мы сразу окажемся в состоянии решить существующие проблемы. В самой генетике понятие «текст» хотя и употреблялось начиная с первооткрывателя ДНК Ф. Мишера (см. выше), но без должного соотношения с лингвистикой или семиотикой текста, как аналогия с более известным и даже очевидным.

Подход к генетической информации как тексту лег в основу, к сожалению, не доведенных в этом отношении до своего логического завершения перспективных системных исследований профессора В. Ратнера и его группы в Новосибирске. Его пионерские идеи о процессах управления генетической информацией [Ратнер, 1965] в дальнейшем закономерно приводят к поискам обнаружения механизмов ее лингвистического структурирования и выработке соответствующих методов описания: «Гены представляют собой не зачатки биологических структур, а напоминают линейные тексты (участки ДНК), записанные при помощи некоторых правил и несущие генетическую информацию о молекулярных структурах и функциях. Они ограничены некоторыми размерами, имеют разные последовательности символов-мономеров. Возможны “бессмысленные” участки, не кодирующие белок или кодирующие испорченные белки. При этом во всех случаях и гены, и не кодирующие участки – сегменты молекул ДНК, т. е. построены из того же алфавита четырех нуклеотидов. Поэтому различия между такими текстами не в их физической природе, а исключительно в последовательностях символов-мономеров. В этом ключ информационно-лингвистического подхода. Значит, гены – не физические, а информационные единицы наследственности, т. е. предположение Шредингера о “шифровальном коде” действительно подтверждается и материализуется» [Ратнер, 2000а].

«Информационно-лингвистический подход» был реализован в ряде работ [Ратнер, 1993а; 1993б; 2000б]: в них была предложена система уровней, и даже фигурирует уровень текста, но исключительно как сложная композиция символов максимальной длины – она называется «книгой» [Ратнер, 1993б]. Как и при других упоминаниях текста, генетиками текст понимался не столько как языковая структура высшего уровня, сколько как некоторая графическая фиксация определенным образом организованных знаков (т. е. на самом деле все сводилось к чему-то вроде книги в ее типографском понимании). Именно такое понимание текста зафиксировано в ценном исследовании, непосредственно посвященном описанию белковых молекул как генетических текстов: «Любая реальная генетическая макромолекула может быть представлена в виде последовательности конечной длины из эле-

ментов данного алфавита, которую можно назвать генетическим текстом» [Соловьев и др., 1988, с. 5]. Заметим, что подобное и даже еще более узкое «графическое» понимание текста все еще продолжает считаться эталонным, по крайней мере среди ведущих экспертов по биосемиотике<sup>7</sup>, а упомянутое понимание (био-)текстовой природы генома до сих пор не является общепринятым, определяющими продолжают оставаться процессы кодирования / декодирования. Впрочем, отметим, что в самой лингвистике вплоть до 1960-х гг. было принято именно такое «графическое» понимание текста, и хотя уже в 1970-е гг. существовали достаточно разработанные структурно-функциональные модели текста (см. [Николаева, 1977; 1978; Dijk, 1972; de Beaugrande, Dressler 1981]), однако вплоть до настоящего времени текст, рассматриваемый как единица коммуникации, остается на периферии теоретической лингвистики. Множественность пониманий текста требует, несколько отвлекаясь от основной темы, предложить некоторые разъяснения.

### Текст-как-продукт vs текст-как-процесс

Все возможные методологические подходы к тексту в конечном счете могут быть сведены к двум основным: текст можно рассматривать либо как заверченный продукт, как некоторую линейную последовательность знаков, либо как процесс, включающий его порождение, контекстуализацию, интертекстуализацию и интерпретации (перевод). В лингвистике все еще преобладает подход к тексту как к продукту, тогда как в семиотике он понимается скорее как процесс.

Начало этому разграничению можно найти в известной статье Ролана Барта «Смерть автора» (1966), где он предлагает рассматривать текст как процесс или даже деятельность: «Ныне мы знаем, что текст представляет собой не линейную цепочку слов, выражающих единственный, как бы теологический смысл (“сообщение” Автора-Бога), но многомерное пространство, где сочетаются и спорят друг с другом различные виды письма, ни один из которых не является исходным; текст соткан из цитат, отсылающих к тысячам культурных источников...» [Барт, 1989, с. 388]. Позднее (1971) Р. Барт предложит разграничение между двумя аспектами текста, в его терминах между произведением и текстом: «Различие здесь вот в чем:

---

<sup>7</sup> Как разъяснил нам анонимный рецензент из журнала «Biosemiotics», «What is understood under “text”? I suspect that the authors have in mind the concept as raised by Y. Lotman: for him a culture, a living being, a biosphere, a picture, even the universe may be conceive as a text. This is in opposition with the Western understanding, where text is a string of characters belonging to some alphabet. Not all such strings, however, have a status of text, only those that represent projections (reduction) from some natural language, i.e. string representing numerals or formal (programming) languages are not texts» (2015 г., личная коммуникация).

произведение есть вещественный фрагмент, занимающий определенную часть книжного пространства (например, в библиотеке), а Текст – поле методологических операций (*un champ methodologique*)... *Текст ощущается только в процессе работы, производства*» [Барт, 1989, с. 388].

Динамический подход к тексту был в дальнейшем развит в работах Юрия Лотмана и других представителей Тартуской школы. Такое понимание становится ключевым при объяснении процессов семиозиса и коммуникации: «Многослойный и семиотически неоднородный текст, способный вступать в сложные отношения как с окружающим культурным контекстом, так и с читательской аудиторией, перестает быть элементарным сообщением, направленным от адресанта к адресату. Обнаруживая способность конденсировать информацию, он приобретает память. Одновременно он обнаруживает качество, которое Гераклит определил как “самовозрастающий логос”. На такой стадии структурного усложнения текст обнаруживает свойства интеллектуального устройства: он не только передает вложенную в него извне информацию, но и трансформирует сообщения и вырабатывает новые... В свете сказанного текст предстает перед нами не как реализация сообщения на каком-либо одном языке, а как сложное устройство, хранящее многообразные коды, способное трансформировать получаемые сообщения и порождать новые, как информационный генератор, обладающий чертами интеллектуальной личности...» [Лотман, 1992, с. 131–132].

Возвращаясь к вопросу о генетической информации, очевидно, что ее структурирование и обработка не являются линейными, и только будучи «переписанным» посредством букв, геном может быть представлен в виде конечной линейной последовательности. Недостаточность вышеупомянутых подходов к представлению генетической информации как текста показывает, что необходимы более тонкие и, главное, методологически более продуманные методы, учитывающие, что текст – это не графическое или даже синтаксическое, а семантическое, структурное и функциональное единство. Это стимулирует пересмотр самой природы текстуальности гена: это не последовательность знаков алфавита, а механизм передачи генетических сообщений, обладающих смысловыми и функциональными характеристиками (как это было предположено сравнительно недавно: [Гаряев, 2009]). Уже сейчас становится ясно, что описание гена не следует сводить исключительно к последовательности нуклеотидов: при рассмотрении его организации прослеживается четкая иерархия формирования внутри текстовых структурных и функциональных единиц. На первом уровне нуклеотидные основания формируют триплеты, кодоны, которым соответствуют единицы второго уровня – аминокислоты. На третьем уровне последовательность аминокислот складывается в мотивы – генетические информационные блоки. Организация мотивов внутри молекулы белка имеет несколько измерений, например альфа-спирали и бета-структуры, что со-

ответствует парадигматическому и синтагматическому измерению языкового текста (в музыковедческих терминах – это управляемая законами гармонии соотнесенность мотивов и аккордов). Возможны также аналогии и с метрическими и ритмическими структурами поэтической речи. Четвертый уровень текстовой организации генома – уровень топологических поверхностей, на котором различные субтексты объединяются в целое. Такая интерпретация позволит осуществить переход от молекулярных структур к социальной геномике: семантико-информационные структуры (память, программа поведения, врожденные лингвистические структуры) могут быть описаны как реакция молекулярных структур на получаемую информацию, которая фиксируется как программа.

#### **Биосемиотика второго поколения – геном как текст и текст как организм**

Можно конкретизировать задачу, взяв за основу дисциплины, в которых наиболее развито понятие текста, хотя его формализация и осуществляется в противоположных направлениях и посредством различных исследовательских аппаратов: это – лингвистика текста и лингвистическая поэтика, с одной стороны, и молекулярная генетика, с другой. При таком подходе задачу можно сформулировать как построение такой теории текста, которая позволит описать, с одной стороны, языковые тексты и процесс их структурирования, функционирования и трансформации, а с другой – генетические, представив геном в виде системы текстов, или гипертекста – текста, состоящего из упорядоченного подмножества других текстов. Информационные процессы следует рассматривать динамически, в процессе их порождения и функционирования, как смысловое и функциональное единство, т. е. как систему, способную к целеполаганию, адаптации и трансформации в зависимости от изменяемых условий (контекста), а также обладающую способностью сохранять информацию (памятью), перерабатывать ее (креативно-когнитивными характеристиками) и передавать ее (коммуникативной силой). В свое время Л. Витгенштейн считал, что каждый язык (языковая игра) есть особая форма жизни (*Lebensformen*), имея в виду функциональное разнообразие социальной коммуникативной деятельности (поведения). Эта мысль может быть перенесена и на биологические формы жизни в следующей редакции: каждая форма жизни есть язык. Таким образом, каждый язык есть форма жизни, и каждая форма жизни есть язык.

Подобные идеи уже высказывались в семиотике. Во-первых, следует вспомнить идею Юрия Лотмана о том, что в процессе функционирования текст приобретает черты живого организма, обладающего памятью и способностью продуцировать новые смыслы [Лотман, 1981; 1984]. Можно, в развитие этой идеи и в определенной мере воскрешая некоторые из идей

лингвистического биологизма XIX в., рассматривать языки и тексты, в том числе и генетические, как подобие биосистем с их такими хорошо известными свойствами, как выживаемость, адаптируемость к новой среде (языковой, социокультурной и т. п.), способность к самовоспроизведению, мутациям (трансформациям), селекции мутаций, контактам и порождению гибридов и т. д.

Во-вторых, также и организм может быть описан как наделенный характеристиками гипертекста «самоинтерпретирующий биотекст» («a self-interpreting biotext» [Kull, 2002]): «Организм можно рассматривать как множество текстов, которые переводят друг друга и, как следствие, создают новые тексты. Все эти тексты являются компонентами в циклах замещения или репродукции, тем самым прямо или косвенно также и в циклах самозамещения и саморепродукции. Таким образом, организм можно рассматривать как совокупность процессов чтения и перевода, при которых некоторые компоненты организма читают и переводят другие компоненты того же организма. Следовательно, организм является текстом, читающим самого себя (self-reading text)» [Kull, 1998, p. 94].

В-третьих, геном человека (генетический гипертекст) также может быть описан посредством семантических и семиотических характеристик, присущих интеллектуальным объектам, что позволит прийти к новым теоретическим и практическим результатам. При таком подходе генетический гипертекст (геном) становится системой и структурой, изоморфной искусственному и коллективному интеллекту. В подтверждение этому, как фрагмент будущего полного описания семиотики генома, можно предложить описание белкового синтеза (генетической трансляции) как семиотический процесс, показав, что «чтение» клеткой генетической информации разделяет существенные характеристики человеческого чтения текста человеком (подробнее об этом см.: [Золян, 2016]). В силу определенных биохимических ограничений кодирующий нуклеотидный триплет из ДНК не может непосредственно быть перенесен в транспортную РНК, где и происходит выбор соответствующей аминокислоты. Привычный способ описания трансляции как сложного биохимического процесса синтеза мешает увидеть его относительно простую и изящную контекстно зависимую семиотику. Как мы попытались показать, определяющим фактором оказывается не биохимический состав триплета, а кодирующая функция, отношение между тройкой нуклеотидов, их позицией и аминокислотой. Иными словами, быть кодоном (или антикодоном) – это определяемая позицией некоторая функция, а не собственно биохимическая субстанция.

Если отойти от рассмотрения биохимического субстрата процесса белкового синтеза и сосредоточиться на его семиотических характеристиках, можно увидеть удивительное сходство между тем, как читает (обрабатывает) текст человек и как читает генетическую информацию рибосома (клетка, ответственная за процессы транскрипции и трансляции). Чтение

есть комплексный когнитивный процесс взаимодействия текста и читателя, который предполагает следующие способности: 1) распознавание букв; 2) объединение их в слоги, далее в слова и предложения; 3) приписывание полученным последовательностям смысла и, наконец, 4) понимание целостного текста. Все эти способности и операции основаны на выявлении и использовании семиотической связи между означаемым и означающим.

«Чтение» клеткой нуклеотидных последовательностей и выбор соответствующей аминокислоты явно выходит за пределы тривиального распознавания знаков и безальтернативного декодирования. Клетка (ее рибосома) способна соотносить биохимически совершенно различные тройки нуклеотидов (означающие) с одной и той же аминокислотой (одним и тем же означаемым), и в то же время она в состоянии соотнести омонимичные означающие с различными означаемыми. В данном случае имеет место основополагающее для любой знаковой системы разграничение между абстрактными элементами языка и их конкретной манифестацией в единицах речи. Так, клетка способна идентифицировать биохимические последовательности исходя из их функции в системе; ее способность читать не ограничивается распознаванием букв, но предполагает, по крайней мере, понимание слов и умение оперировать абстрактными сущностями (выделять инвариант при трансформациях). Это предполагает, что в определенном смысле клетка (рибосома) обладает (квази)интеллектом [Золян, 2016].

### Заключение

Будучи описаны как семиотические объекты, генетические тексты, которые сами по себе и являются биосистемами, могут оказаться не только областью приложения, но и моделью для общей теории текста. Однако для достижения подобных результатов необходимы и новые методы описания. Так, нам представляются неудовлетворительными попытки описать генетическую информацию посредством линейных контекстно свободных лингвистических моделей (подобно хорошо знакомой генетикам и биосемиотикам порождающей грамматике Н. Хомского). Более адекватной может оказаться новая версия семиотики, которая в качестве своего главного прототипа будет иметь не слово (и даже не предложение), а текст, причем, возможно, не предполагающий наличие некоего исходного набора. В такой мультимодальной и текстоцентричной версии семиотики исходным понятием явится текст, а смысл знака будет рассматриваться как контекстно зависимая переменная [Zolyan, 2012; Золян, 2013]. При этом, следуя принципу динамической семантики, следует исходить из того, что семантика единиц не только зависит от контекста, но, в свою очередь, сама видоизменяет контекст. Учитывая разнородность семиотических и генетических систем, исследование способов организации генетической информации будет стимулировать появление новых методов и в семиотической

теории текста. Достигнутые в лингвистике и поэтике результаты, несмотря на имеющие почти полувековую историю программные декларации Э. Бенвениста об ориентированной на исследование текстов «семиотике второго поколения» [Бенвенист, 1974, с. 88–89], еще не нашли своего обобщения в семиотике, которая продолжает исходить из приоритета знака, а не текста.

Исходя из того, что биохимические и генетические процессы есть особый тип информационных процессов, они должны разделять наиболее существенные характеристики, общие для всех видов коммуникации. Информативной единицей коммуникации, будь то социальная или биологическая коммуникация, является не отдельный сигнал, а некоторая целостная структура – текст. Именно текст есть форма организации информации, если исходить из ее качественных смысловых характеристик. Текст (а не знак) есть объект создания – передачи – сохранения – преобразования информации. Поэтому насущной проблемой представляется построение такой общей семиотической теории текста и такой текстоцентричной модальной семиотики, благодаря которым можно было бы описывать как тексты нелинейной структуры в целом, так и тексты, несущие генетическую информацию, в частности. Это создаст основу для решения столь важной задачи, как соотнесение биологической и социальной геномики – как переход от молекулярных структур к семантико-информационным (память, программа поведения, врожденные лингвистические структуры).

### Список литературы

- Барт Р.* Смерть автора // Барт Р. Избранные работы. Семиотика. Поэтика. М.: Прогресс, 1989. С. 384–391.
- Барт Р.* От произведения к тексту // Барт Р. Избранные работы. Семиотика. Поэтика. М.: Прогресс, 1989. С. 413–423.
- Бенвенист Э.* Общая лингвистика. М.: Прогресс, 1974. 446 с.
- Ванюшин Б. Ф.* Метилирование ДНК и эпигенетика // Генетика. 2006. Т. 42, № 9. С. 1–14.
- Витгенштейн Л.* Логико-философский трактат. М.: Иностран. лит., 1958. 131 с.
- Гамкрелидзе Т. В. Р. О.* Якобсон и проблема изоморфизма между генетическим кодом и семиотическими системами // Вопросы языкознания. 1988. № 3. С. 5–9.
- Гаряев П. П.* Лингвистико-волновой геном. Киев, 2009.
- Гельфанд М. С.* Коды генетического языка и естественный язык // Вопросы языкознания. 1990. № 6. С. 60–70.
- Ельмслев Л.* Прологомены к теории языка // Новое в лингвистике. М.: Прогресс, 1960. Вып. 1. С. 264–389.

Жакоб Ф. Лингвистическая модель в биологии // Вопросы языкознания. 1992. № 2. С. 135–141.

Золян С. Т. «Бесконечный лабиринт сцеплений»: семантика текста как многомерная структура // Критика и семиотика. 2013а. № 1/18. С. 18–44.

Золян С. Т. Вновь о соотносительности языка и генетического кода // Вопросы языкознания. 2016. № 1. С. 114–132.

Иванов Вяч. Вс. Границы семиотики: вопросы к предварительному обсуждению // Современная семиотика и гуманитарные науки. М.: Языки славянской культуры, 2010. С. 31–52.

Лотман М. Ю. Двойственная природа текста (связный текст как семиотическое и коммуникативное образование) // Текст и культура: общие и частные проблемы / Под ред. Ю. А. Сорокина. М.: Изд-во Ин-та языкознания АН СССР, 1985. С. 3–20.

Лотман Ю. М. Мозг – текст – культура – искусственный интеллект // Семиотика и информатика. М., 1981. Вып. 1. С. 13–17.

Лотман Ю. М. О семиосфере // Учен. зап. Тарт. гос. ун-та. 1984. Вып. 641. (Труды по знаковым системам. Т. 17.) С. 5–23.

Лотман Ю. М. Семиотика культуры и понятие текста // Лотман Ю. М. Избранные статьи. Таллинн, 1992. Т. 1. С. 129–132.

Материалы Международного симпозиума «Язык генома» / Казанский федеральный университет, Академия наук РТ; под ред. Р. И. Жданова. Казань, 2014. 32 с.

Николаева Т. М. Лингвистика текста и проблемы общей лингвистики // Изв. АН СССР. Серия литературы и языка. 1977. Т. 36, № 4. С. 304–313.

Николаева Т. М. Теория текста. Современное состояние и перспективы // Новое в зарубежной лингвистике. М., 1978. Вып. 8. С. 5–41.

Пропп В. Я. Морфология сказки // Вопросы поэтики. Л.: Academia, 1928. Вып. 12. 152 с.

Ратнер В. А. Генетические управляющие системы: Автореф. дис. ... канд. биол. наук. Новосибирск, 1965.

Ратнер В. Генетический язык: грамматика, предложения, эволюция // Генетика. 1993а. № 29. С. 709–719.

Ратнер В. Сравнительная иерархическая структура генетического языка // Генетика. 1993б. № 29. С. 720–739.

Ратнер В. А. Генетический код как система // Биология. Соросовский образовательный журнал. 2000. Т. 6, № 3. С. 17–22.

Ратнер В. А. Хроника великого открытия: идеи и лица // Природа. 2000. № 6. URL: [http://vivovoco.astronet.ru/VV/JOURNAL/NATURE/06\\_00/CODE/CODE.HTML](http://vivovoco.astronet.ru/VV/JOURNAL/NATURE/06_00/CODE/CODE.HTML) (дата обращения 18.02.2016).

Седов А. Е. Метафоры в генетике // Вестн. РАН. 2000. Т. 70, № 6. С. 526–534.

Седов А. Е. Иерархические концепции и междисциплинарные связи генетики, запечатленные в ее метафорах: количественный и структурный анализ терминов и высказываний // *Науковедение*. 2001. № 1. С. 135–154.

Соловьев В. В., Кель А. Э., Rogozin И. Б., Колчанов Н. А. Использование ЭВМ в молекулярной биологии. Введение в теорию генетических текстов. Новосибирск, 1988. 92 с.

Степанов Ю. С. Семиотика. М.: Наука, 1971.

Шредингер Э. Что такое жизнь? Физический аспект живой клетки. М.; Ижевск, 2002.

Чебанов С. Российская биосемиотика как школа: проблема институализации. Научные чтения – 2003. Материалы конф. СПб., 2004. С. 164–171.

Якобсон Р. Лингвистика в ее отношении к другим наукам // Якобсон Р. Избранные работы. М.: Прогресс, 1985. С. 387–404.

Augustyn P. What Connects Biolinguistics and Biosemiotics? // *Biolinguistics*. 2013. Vol. 7. P. 96–111.

Barbieri M. Is the Cell a Semiotic System? // Barbieri M. (ed.) *Introduction to Biosemiotics. The New Biological Synthesis*. Dordrecht: Springer, 2007.

Barbieri M. Biosemiotics: A New Understanding of Life // *Naturwissenschaften*. Springer Verlag, 2008a. P. 577–599.

Barbieri M. What is Biosemiotics? // *Biosemiotics* (1). Springer, 2008b. P. 1–3.

Barbieri M. A Short History of Biosemiotics // *Biosemiotics*. Berlin: Springer Science, 2009. No. 2. P. 221–245.

Bertalanffy L. *Theoretische Biologie*. Berlin, 1932.

Bertalanffy L. *General Systems Theory. Foundations, Development, Applications*. New York, 1968. 289 p.

Beaugrande R.-A. de, Dressler W. *Introduction to Text Linguistics*. London: Longman, 1994. 270 p.

Bickerton D. Some Problems for Biolinguistics // *Biolinguistics*. 2014. Vol. 8. P. 73–96.

Dijk T. van. *Some aspects of text grammars. A study in theoretical poetics and linguistics*. The Hague: Mouton, 1972.

Crick F. H. C. The Genetic Code – Yesterday, Today and Tomorrow // *Cold Spring Symposium on Quantitative Cold Spring Harb. Symp. Quant. Biol.* 1966. Vol. 31. P. 3–9.

Crick F. *Life Itself: Its Origin and Nature*. Simon and Schuster, 1981. 193 p.

Elumalai A., Eswaraiah M. Review on application of bioinformatics // *Journal of Science. Bioinformatics*. 2013. Vol. 3. Iss. 1. P. 21–27.

Emmeche C., Hoffmeyer J. From language to nature – the semiotic metaphor in biology // *Semiotica*. 1991. Vol. 84 (1/2). P. 1–42.

Favareau D. The evolutionary history of biosemiotics // Barbieri M. (ed.) *Introduction to Biosemiotics. The New Biological Synthesis*. Dordrecht: Springer, 2007. P. 1–69.

*Galik D.* Biosemiotics: A New Science of Biology? // *Filozofia*. 2013. Vol. 68. No. 10. P. 859–867.

*Gamkrelidze T.* «Paradigms» in Linguistics and the Problem of the Isomorphism between the Genetic Code & Semiotic Systems // *Bulletin of the Georgian National Academy of Sciences*. 2009. Vol. 3. No. 2.

*Gimona M.* Protein Linguistics and the Modular Code of the Cytoskeleton // *Barbieri M. (ed.) The Codes of Life: The Rules of Macroevolution*. Berlin: Springer, 2008. P. 189–206.

*Jakobson R.* Linguistics. Relationship between the science of language and other sciences // *Main trends of research in the social and human sciences*. The Hague: Mouton, 1970. P. 419–453.

*Ji S.* Isomorphism between Cell and Human Languages: Molecular Biological, Bioinformatic and Linguistic Implications // *Biosystems*. 1997. Vol. 44. Iss. 1. P. 17–39.

*Ji S.* The Linguistics of DNA: Words, Sentences, Grammar, Phonetics, and Semantics // *Molecular Strategies in Biological Evolution*. *Annals of the New York Academy of Science*. 1999. Vol. 870. P. 411–417.

*Kull K.* Organism as a Self-Reading Text: Anticipation and Semiosis // *International Journal of Computing Anticipatory Systems*. 1998. Vol. 1. P. 93–104.

*Kull K.* A Sign Is not Alive – a Text Is // *Sign Systems Studies*. 2002. Vol. 30.1.

*Kull K., Deacon K., Emmeche C., Hoffmeyer J., Stjernfelt F.* Theses on Biosemiotics: Prolegomena to a Theoretical Biology // *Biological Theory*. 2009. Vol. 4 (2). P. 167–173.

*Luhmann N.* *Essays on Self-Reference*. New York: Columbia University Press, 1990.

*Luhmann N.* *Social Systems*. Stanford: Stanford University Press, 1995.

*Maturana H. R., Varela F. J.* *Autopoiesis and cognition: The realization of the living*. Boston: D. Reidel, 1980.

*Pattee H. H.* The Necessity of Biosemiotics: Matter-Symbol Complementarity the Physics and Metaphysics of Biosemiotics // *Journal of Biosemiotics*. 2005. Vol. 1 (1). P. 223–238.

*Raible W.* Linguistics and Genetics: Systematic parallels // *Language Typology and Language Universals. An International Handbook* / Eds. M. Haspelmath, E. König, W. Oesterreicher, W. Raible. Berlin; New York: Walter de Gruyter, 2001. P. 103–123.

*Rotshild F. S.* Laws of symbolic mediation in the dynamics of self and personality // *Annals of the New York Academy of Sciences*. 1962. Vol. 96 (3). P. 774–784.

*Searls D.* The Computational Linguistics of Biological Sequences // *Artificial Intelligence and Molecular Biology*. 1993. P. 47–120.

Searls D. B. Formal Language Theory and Biological Macromolecules // Mathematical Support for Molecular Biology. American Mathematical Society Press, 1999. P. 117–140.

Searls D. B. The Language of Genes // Nature. 2002. Vol. 420 (6912). P. 211–217.

Searls D. B. Molecules, Languages and Automata // Grammatical Inference: Theoretical Results and Applications Lecture Notes in Computer Science. 2010. Vol. 6339. P. 5–10.

Shannon C. E. A mathematical theory of communication // Bell System Technical Journal. 1948. Vol. 27 (379–423). P. 623–656.

Stjernfelt F. Tractatus Hoffmeyerensis: Biosemiotics as expressed in 22 basic hypotheses // Sign Systems Studies. 2002. Vol. 30.1. P. 337–345.

Trifonov E. N. Earliest pages of bioinformatics // Bioinformatics. 2000. Vol. 16. No. 1. P. 5–9.

Trifonov E. N. Codes of Biosequences // Barbieri M. (ed.) The Codes of Life. The Rules of Macroevolution. Springer, 2008. P. 3–14

Trifonov E. N. Thirty Years of Multiple Sequence Codes // Genomics. Proteomics. Bioinformatics. 2011. Vol. 9 (1–2). P. 1–6.

Wiener N. Cybernetics: Or Control and Communication in the Animal and the Machine. Paris: Hermann & Cie; Cambridge: MIT Press, 1948.

Witzany G. Review: Marcello Barbieri (ed.) Introduction to Biosemiotics. The New Biological Synthesis. Dordrecht: Springer, 2007. Triple C 5 (3). P. 104–109.

Zhdanov R. I., Ibragimova M. Y. New informational level at genomic DNA: lipids specifically bound to DNA // Genetics and Chemistry Sharing a Language of Discovery: Cell Symposium series, May 23–25, 2012. Boston; Marriott; Cambridge, 2014. P. 29.

Zolyan S. Text as a Multisemantic Entity – a Prolegomenon to Formalization // International Congress Cultural Poliglottism, to the anniversary of Juri Lotman's 90<sup>th</sup> birthday. Tartu, 2012. P. 57–62.

#### Article metadata

*Title:* Genome as a (hyper) text: from metaphor to theory

*Author:* S. T. Zolyan

*Author's e-mail:* surenzolyan@gmail.com

*Author affiliation:* National Academy of Sciences of Armenia, Institute of Philosophy, Sociology and Law

*Author:* R. I. Zhdanov

*Author's e-mail:* zrenad@gmail.com

*Author affiliation:* Russian Institute for Advanced Study, Moscow Pedagogical State University and Institute of Fundamental Medicine and Biology, Kazan Federal University, 18, Kremlin St., Kazan 420008 Russian Federation

*Abstract.* The similarity between language and genetic information transmission processing was recognized from the very foundation of genetics. A lot of attempts have been made to use linguistics techniques for deciphering protein genes. A modest informative impact of various kinds of decoding of «protein language» was predictable: this kind of technique is limited by the processes of encoding and at best it helps to compile a dictionary of units, but is not able to describe a language and identify the semantic and textual structures which are decisive for communication. Thus, functions of the huge amount of non-coding DNA still remain unclear. A text but not a sign should be regarded as an object of creation – transmission – conservation – transformation of information. The general text theory should be capable to describe linguistic texts and the process of their structuring, functioning and transformation, and to represent a biochemical structure of genome as a hypertext consisting of ordered subset of other texts. The text can be considered as a quasi-organism having memory, creative-cognitive characteristics and communicative force; the cell (ribosome) – as a quasi-intelligence capable to operate with abstract semiotic entities. Such an approach will require the construction of a new, multimodal and textocentric version of semiotics (the semiotics of the next generation), where text and context will be the basic concepts, and the signs and meanings will be considered as context-dependent variables, and at the same time as context-forming operators.

*Key terms:* genome, genetic code, biosemiotics, general theory of text, genome as a text.

*Reference literature (in translation):*

Augustyn P. What Connects Biolinguistics and Biosemiotics? *Biolinguistics*, 2013, vol. 7, p. 96–111.

Barbieri M. A Short History of Biosemiotics. *Biosemiotics*. Berlin, Springer Science, 2009, no. 2, p. 221–245.

Barbieri M. Biosemiotics: A New Understanding of Life. *Naturwissenschaften*. Springer Verlag, 2008a, p. 577–599.

Barbieri M. Is the Cell a Semiotic System? *Introduction to Biosemiotics. The New Biological Synthesis*. M. Barbieri (ed.). Dordrecht, Springer, 2007.

Barbieri M. What is Biosemiotics? *Biosemiotics* (1). Springer, 2008b, p. 1–3.

Barthes R. From work to text. *Barthes Rolan. Izbrannyye raboty. Semiotika. Poetika*. Moscow, Progress, 1989, p. 413–423.

Barthes R. The death of the Author. *Barthes Rolan. Izbrannyye raboty. Semiotika. Poetika*. Moscow, Progress, 1989, p. 384–391.

Beaugrande R.-A. de, Dressler W. *Introduction to Text Linguistics*. London, Longman, 1994, 270 p.

Benveniste E. *Obshchaya lingvistika [General linguistics]*. Moscow, Progress, 1974.

Bertalanffy L. *General Systems Theory. Foundations, Development, Applications*. New York, G. Braziller, 1968, 289 p.

- Bertalanffy L. *Theoretische Biologie*. Berlin, 1932.
- Bickerton D. Some Problems for Biolinguistics. *Biolinguistics*, 2014, vol. 8, p. 73–96.
- Chebanov S. Russian biosemiotics as a school: The problem of institutionalization. *Nauchnye chteniya – 2003. Materialy konferentsii*. St. Petersburg, Philological Department of St. Petersburg State Univ., 2004, p. 164–171.
- Crick F. H. C. The Genetic Code – Yesterday, Today and Tomorrow. *Cold Spring Symposium on Quantitative Cold Spring Harb. Symp. Quant. Biol.*, 1966, vol. 31, p. 3–9.
- Crick F. *Life Itself: Its Origin and Nature*. Simon and Schuster, 1981, 193 p.
- Dijk T. van. *Some aspects of text grammars. A study in theoretical poetics and linguistics*. The Hague, Mouton, 1972.
- Elumalai A., Eswaraiah M. Review on application of bioinformatics. *Journal of Science. Bioinformatics*, 2013, vol. 3, iss. 1, p. 21–27.
- Emmeche C., Hoffmeyer J. From language to nature – the semiotic metaphor in biology. *Semiotica*, 1991, vol. 84 (1/2), p. 1–42.
- Favareau D. The evolutionary history of biosemiotics. *Introduction to Biosemiotics. The New Biological Synthesis*. M. Barbieri (ed.). Dordrecht, Springer, 2007, p. 1–69.
- Galik D. Biosemiotics: A New Science of Biology? *Filozofia*, 2013, vol. 68, no. 10, p. 859–867.
- Gamkrelidze T. «Paradigms» in Linguistics and the Problem of the Isomorphism between the Genetic Code & Semiotic Systems. *Bulletin of the Georgian National Academy of Sciences*, 2009, vol. 3, no. 2.
- Gamkrelidze T. V. R. O. Jakobson and the problem of isomorphism between the genetic code and semiotic systems. *Voprosy Jazykoznanija*, 1988, no. 3, p. 5–9.
- Garyaev P. P. *Lingvistiko-volnovoï genom [Linguistic wave genome]*. Kiev, The Institute of quantum genetics, 2009.
- Gelfand M. S. Codes of genetic language and natural language. *Voprosy jazykoznanija*, 1990, no. 6, p. 60–70.
- Gimona M. Protein Linguistics and the Modular Code of the Cytoskeleton. *The Codes of Life: The Rules of Macroevolution*. M. Barbieri (ed.). Berlin, Springer, 2008, p. 189–206.
- Hjelmslev L. Prolegomena to the theory of language. *Novoe v lingvistike*. Moscow, Progress, 1960, no. 1, p. 264–389.
- Ivanov V. V. «The frontiers of semiotics»: Questions for a preliminary discussion. *Sovremennaya semiotika i gumanitarnye nauki*. Moscow, Yazyki Slavyanskikh Kul'tur, 2010, p. 31–50.
- Jakob F. The linguistic model in biology. *Voprosy jazykoznanija*, 1992, no. 2, p. 135–141.
- Jakobson R. Linguistics in its relation to other sciences. *Jakobson R. Izbrannye raboty*. Moscow, Progress, 1985, p. 387–404.

Jakobson R. *Linguistics. Relationship between the science of language and other sciences. Main trends of research in the social and human sciences.* Mouton, The Hague, 1970, p. 419–453.

Ji S. Isomorphism between Cell and Human Languages: Molecular Biological, Bioinformatic and Linguistic Implications. *Biosystems*, 1997, vol. 44, iss. 1, p. 17–39.

Ji S. The Linguistics of DNA: Words, Sentences, Grammar, Phonetics, and Semantics. *Molecular Strategies in Biological Evolution. Annals of the New York Academy of Science*, 1999, vol. 870, p. 411–417.

Kull K. A Sign Is not Alive – a Text Is. *Sign Systems Studies*, 2002, vol. 30.1.

Kull K. Organism as a Self-Reading Text: Anticipation and Semiosis. *International Journal of Computing Anticipatory Systems*, 1998, vol. 1, p. 93–104.

Kull K., Deacon K., Emmeche C., Hoffmeyer J., Stjernfelt F. Theses on Bio-semiotics: Prolegomena to a Theoretical Biology. *Biological Theory*, 2009, vol. 4 (2), p. 167–173.

Lotman M. Yu. The dualistic nature of the text (discourse as a semiotic and communicative entity). *Tekst i kul'tura: obshchie i chastnye problemy.* Moscow, Institute of Linguistics, Academy of Sciences of the USSR, 1985, p. 3–20.

Lotman Yu. M. The semiotics of culture and the notion of text. *Lotman Yu. M. Izbrannyye stat'i.* Tallinn, 1992, vol. 1, p. 129–132.

Lotman Yu. M. Brain – text – culture – artificial intelligence. *Semiotika i informatika.* Moscow, All-Union Institute of Scientific and Technical Information, 1981, no. 17, p. 13–17.

Lotman Yu. M. On semiosphere. *Uchenye zapiski Tartuskogo gosudarstvennogo uiversiteta*, 1984, no. 641. (Trudy po znakovym sistemam. Vol. 17.), p. 5–23.

Luhmann N. *Essays on Self-Reference.* New York, Columbia University Press, 1990.

Luhmann N. *Social Systems.* Stanford, Stanford University Press, 1995.

Maturana H. R., Varela F. J. *Autopoiesis and cognition: The realization of the living.* Boston, D. Reidel, 1980.

Nikolaeva T. M. Text linguistics: Current status and prospects. *Novoe v zarubezhnoi lingvistike.* Moscow, Progress, 1978, no. 8, p. 5–42.

Nikolaeva T. M. Text linguistics and the problems of general linguistics. *Izvestiya Akademii nauk SSSR. Seriya literatury i yazyka*, 1977, vol. 36, no. 4, p. 304–313.

Pattee H. H. The Necessity of Biosemiotics: Matter-Symbol Complementarity the Physics and Metaphysics of Biosemiotics. *Journal of Biosemiotics*, 2005, vol. 1 (1), p. 223–238.

Propp Vladimir. Morphology of the Folktale. (Morfologiya skazki). *Voprosy poetiki.* Leningrad, Academia, 1928, iss. 12, p. 152.

Raible W. Linguistics and Genetics: Systematic parallels. *Language Typology and Language Universals. An International Handbook*. M. Haspelmath, E. König, W. Oesterreicher, W. Raible (eds.). Berlin, New York, Walter de Gruyter, 2001, p. 103–123.

Ratner V. A. Genetic code as a system. *Sorosovskii obrazovatel'nyi zhurnal. Biologiya*, 2000, vol. 6, no. 3, p. 17–22.

Ratner V. A. Genetic language: Grammar, sentences, evolution. *Genetika*, 1993, no. 29, p. 709–719.

Ratner V. A. *Geneticheskie upravlyayushchie sistemy. Avtoref. kand. diss. [Genetic control systems. Author's abstract of the cand. diss.]* Academy of Sciences of the USSR, Siberian Branch, Joint Council for Biological Sciences, Novosibirsk, 1965.

Ratner V. A. The Chronicle of the great discovery: the ideas and persons. *Priroda*, 2000, no. 6.

Ratner V. A. The comparative hierarchic structure of the genetic language. *Genetika*, 1993, no. 29, p. 720–739.

Rotshild F. S. Laws of symbolic mediation in the dynamics of self and personality. *Annals of the New York Academy of Sciences*, 1962, vol. 96 (3), p. 774–784.

Schrödinger E. *Chto takoe zhizn'?* *Fizicheskii aspekt zhivoi kletki [What is life? Physical aspect of the living cell]*. Moscow, Izhevsk, Scientific Research Center «Regular and chaotic dynamics», 2002.

Searls D. B. Formal Language Theory and Biological Macromolecules. *Mathematical Support for Molecular Biology*. American Mathematical Society Press, 1999, p. 117–140.

Searls D. B. Molecules, Languages and Automata. *Grammatical Inference: Theoretical Results and Applications Lecture Notes in Computer Science*, 2010, vol. 6339, p. 5–10.

Searls D. B. The Language of Genes. *Nature*, 2002. vol. 420 (6912), p. 211–217.

Searls D. The Computational Linguistics of Biological Sequences. *Artificial Intelligence and Molecular Biology*, 1993, p. 47–120.

Sedov A. E. Hierarchic conceptions and cross-disciplinary connections of genetics embedded in its metaphors: Quantitative and structural analysis of terms and utterances. *Naukovedenie*, 2001, no. 1, p. 135–154.

Sedov A. E. Metaphors in genetics. *Vestnik Rossiiskoi akademii nauk*, 2000, vol. 70, no. 6, p. 526–534.

Shannon C. E. A mathematical theory of communication. *Bell System Technical Journal*, 1948, vol. 27 (379–423), p. 623–656.

Solov'ev V. V., Kel' A. E., Rogozin I. B., Kolchanov N. A. *Ispol'zovanie EVM v molekulyarnoi biologii. Vvedenie v teoriyu geneticheskikh tekstov [The use of computers in molecular biology. Introduction to the theory of genetic texts]*. Novosibirsk, Novosibirsk Univ., 1988.

- Stepanov Yu. S. *Semiotika* [*Semiotics*]. Moscow, Nauka, 1971.
- Stjernfelt F. Tractatus Hoffmeyerensis: Biosemiotics as expressed in 22 basic hypotheses. *Sign Systems Studies*, 2002, vol. 30.1, p. 337–345.
- Trifonov E. N. Codes of Biosequences. *The Codes of Life. The Rules of Macroevolution*. M. Barbieri (ed.). Springer, 2008, p. 3–14
- Trifonov E. N. Earliest pages of bioinformatics. *Bioinformatics*, 2000, vol. 16, no. 1, p. 5–9.
- Trifonov E. N. Thirty Years of Multiple Sequence Codes. *Genomics. Proteomics. Bioinformatics*, 2011, vol. 9 (1–2), p. 1–6.
- Vanyushin B. F. Metilirovaniye DNK i epigenetika (DNA methylation and epigenetics). *Genetika*, 2006, vol. 42, no. 9, p. 1–14.
- Wiener N. *Cybernetics: Or Control and Communication in the Animal and the Machine*. Paris, Hermann & Cie, Camb. Mass., MIT Press, 1948.
- Wittgenstein L. *Tractatus Logico-Philosophicus*. Moscow, Izd. inostrannoy literatury, 1958, 131 p.
- Witzany G. *Review: Marcello Barbieri (ed.) Introduction to Biosemiotics. The New Biological Synthesis*. Dordrecht, Springer, 2007, triple C 5 (3), p. 104–109.
- Zhdanov R. I., Ibragimova M. Y. New informational level at genomic DNA: lipids specifically bound to DNA. *Genetics and Chemistry Sharing a Language of Discovery: Cell Symposium series*, May 23–25, 2012. Boston, Marriott, Cambridge, 2014, p. 29.
- Zhdanov Renad I. (ed.) Abstracts book of International Symposium «Language of Genome», Kazan Federal University. Kazan, 2014, 32 p.
- Zolyan S. T. «The endless labyrinth of connections»: Text semantics as a multidimensional structure. *Kritika i semiotika*, 2013, no. 1 (18), p. 18–44.
- Zolyan S. T. Rethinking the correlations between language and genetic code. *Voprosy jazykoznanija*, 2016, no. 1, p. 114–132.
- Zolyan S. Text as a Multisemantic Entity – a Prolegomenon to Formalization. *International Congress Cultural Poliglotism, to the anniversary of Juri Lotman's 90<sup>th</sup> birthday*. Tartu, 2012, p. 57–62.