

## ЯЗЫКИ РАЗВЕРТЫВАНИЯ НАСЛЕДСТВЕННОЙ ИНФОРМАЦИИ В ЭМБРИОГЕНЕЗЕ: ЛИНГВОСЕМИОТИЧЕСКИЕ АНАЛОГИ И АНАЛОГИИ

*А. В. Спилов*<sup>1, 2</sup>

<sup>1</sup> Институт научной информации по общественным наукам РАН,  
Россия, 117418, Москва, Нахимовский проспект, 51/21

<sup>2</sup> Институт эволюционной физиологии и биохимии им. И. М. Сеченова РАН  
Россия, 194223, Санкт-Петербург, просп. Тореза, 44

Поступила в редакцию 12.04.2024 г.

Принята к публикации 15.06.2024 г.

doi: 10.5922/2225-5346-2024-4-2

*Известно, что практически вся наследственная информация о неисчислимых характеристиках многоклеточного организма, в частности организма человека, определенным образом закодирована в ядре оплодотворенной яйцеклетки. Принципы разворачивания генетической информации в развитии многоклеточного зародыша давно привлекают внимание как биологов, так и представителей других наук. Если молекулярные биологи концентрируются на информационных и кибернетических аспектах хранения и передачи генетической информации, авторы биосемиотических исследований настаивают на специфике биологических знаков в этих процессах и на особой природе биологических текстов. Мы сосредоточились на информационных аспектах этих процессов с целью продемонстрировать аналогии между ними и лингвосемиотическими понятиями. Описаны проблемы сложности (комплексности) и иерархичности (многоуровневости) генетических механизмов реализации генетической информации в эмбриогенезе, как она видится в системной биологии.*

**Ключевые слова:** биосемиотика, лингвистика, наследственная информация, разворотка информации в индивидуальном развитии, языки регуляции генов

### 1. Введение

Нами будут рассмотрены принципы молекулярных процессов и механизмов разворачивания (unfolding) генетической информации в многоклеточном эмбрионе. Основное внимание уделяется информационно-лингвистическим и семиотическим аспектам регуляции активности генов. Принято считать, что большая часть информации о том, как устроен и функционирует организм (включая человеческий), хранится в ядре оплодотворенной яйцеклетки. В ходе эмбрионального развития в считанные недели или месяцы эта информация «разворачивается» во все нарастающую сложность будущего многоклеточного организма. Так, тело взрослого человека включает порядка тридцати триллионов клеток. Все это организовано во многочисленные органы и ткани, интегрируется нервной и гормональной системами, защищается иммунитетом, поддерживается, обновляется и «чинится» соответствующими системами. Нас будут интересовать базовые информационные и семиотические аспекты этих биологических процессов и явлений.



Попытки рассматривать процессы реализации (развертывания) генетической информации в эмбриогенезе в семиотических аспектах и по аналогии с языками предпринимаются уже давно (Pattee, 1968; 1969; Jakobson, 1970; Соловьев и др., 1988). При этом с каждым десятилетием появляются новые результаты, демонстрирующие, насколько эти процессы сложнее, чем мы полагали еще совсем недавно. Если молекулярные биологи концентрировались на информационных и кибернетических аспектах хранения и передачи генетической информации, то авторы биосемиотических публикаций настаивали на специфике биологических знаков в этих процессах и на уникальной природе биологических текстов (Pattee, 2001; Hoffmeyer, 2014). Мы сосредоточимся на вопросах комплексности и иерархичности (многоуровневости) генетических (и эпигенетических механизмов) реализации генетической информации в эмбриогенезе, как она видится системным биологам в наши дни. Достигнутый уровень детализации механистической трактовки процессов реализации генетической информации свидетельствует о ее имманентной сложности, что должно быть отражено и с семиотических позиций.

Лингвистическое видение базовых процессов молекулярной биологии и молекулярной генетики формировалось со времен открытия универсального генетического кода и продолжается так или иначе все последующие десятилетия (см., напр., обзор: (Raible, 2001)). Именно в этой области были сформулированы начальные лингвистические аналогии, которые некоторыми авторами успешно развиваются и по сей день (Zolyan, 2022; 2023). Вместе с тем нынешняя область исследования кодов и «языков» в биологии невообразимо расширилась. С одной стороны, код-биология, ставшая новым направлением биосемиотики, собрала и начала систематизировать более двух сотен различных биологических кодов (Prinz, 2023). С другой — в пределах современной биологии расширяются исследования «языков» реализации генетической информации в эмбриональном развитии и в функционировании зрелого организма (Weingarten-Gabbay, Segal, 2014; Avsec et al., 2021; Lal, 2022; Bentsen et al., 2022). Биологи выделяют как минимум слова и грамматические правила для этих слов, что помогает понимать и сопоставлять их. Однако практически без внимания остается проблематика, связанная с лингвосемиотическими аспектами развертывания генетической информации в индивидуальном развитии и возможными аналогиями с языковыми процессами.

## 2. Как биомолекулы становятся знаками?

При изучении живой природы наука обнаружила процессы хранения, передачи, интерпретации и реализации информации, аналогичные тем, что реализованы в современных технологиях. Но в биологии эти системы появились и развились самопроизвольно. Поэтому вполне резонно возникают такие вопросы: кто посылает генетическую информацию (сообщения) и кто является получателем этих сообщений?



Как отмечает Дж. Хоффмайер, под генетической информацией в молекулярной биологии обычно понимают просто набор «инструкций» (Hoffmeyer, 2014). Поскольку «инструкция» является малоосмысленным термином, если не адресована «кому-то», кто может в конечном итоге ее выполнить, то «информация» остается весьма неопределенным термином применительно к генам: кто именно должен понимать и выполнять предполагаемую «инструкцию»? Инструкция должна быть интерпретирована некоторым агентом, в противном случае она не будет иметь какого-либо эффекта (Ibid.).

С более общих позиций мы можем вслед за Т. Диконом задаться следующим вопросом: «какая форма молекулярного процесса необходима и достаточна для интерпретации некоторого свойства молекулы как предоставления информации о других молекулярных свойствах» (Deason, 2021, p. 540)? По Дикону, в терминах Чарльза Пирса это равносильно вопросу о том, какая молекулярная система способна создавать интерпретанты, которые могут привести это репрезентированное свойство в полезную связь с этой системой. Интегрируя семиотику Пирса с биологическими процессами, Дикон вслед за Хоффмайером создает основу для понимания того, как макромолекулы переходят из простых химических объектов в носителей биологически значимой информации, становясь, таким образом, знаками в семиотическом смысле. Иначе говоря, переход от химических процессов к семиотическим предполагает появление значения. Так что макромолекулы, такие как РНК, ДНК и белки, начинают функционировать не просто как химические образования, а как носители информации, которая может влиять на биологические функции и процессы.

Следуя триадической модели знака Пирса, которая включает в себя знак (репрезентамен, *representamen*), объект и интерпретант, Дикон предполагает, что макромолекулы можно рассматривать как знаки, когда они вступают в триадические отношения (Ibid.). Например, последовательность ДНК (знак) представляет собой определенную аминокислотную последовательность (объект) и интерпретируется клеточным механизмом (интерпретант) для производства белка.

Дикон (Ibid.) вслед за Г. Патти (Pattee, 1968; 1969; 2001) подчеркивает роль «ограничений» (constraints) в возникновении информации. Макромолекулы становятся носителями информации, накладывая ограничения на возможные результаты биохимических процессов. Например, структура фермента ограничивает возможные реакции, эффективно направляя биохимический процесс в определенном направлении и неся таким образом информацию о потенциальных реакциях. Именно с позиций роли ограничений Дикон описывает семиотические аспекты процессов генной регуляции в индивидуальном развитии и эволюции.

С более общих семиотических позиций существенна прежде всего иерархичность языков (и / или кодов) реализации генетической информации в зародышевом развитии (Deason, 2021). Как формулирует



Дикон, на базовом уровне регуляции гена локальная нуклеотидная последовательность молекулы ДНК влияет на организацию двойной спирали ДНК в этом локусе. Это облегчает избирательное связывание белков, способных изменять экспрессию соседних генов. Итак, белковая структура, определяемая последовательностями ДНК, может стимулировать, ингибировать или регулировать транскрипцию многих других последовательностей ДНК и определяемых ими белковых структур. Таким образом, смещение функциональных ограничений на другую молекулярную среду (то есть от нуклеотидной последовательности к структуре белка) открывает путь к рекурсивной информационной динамике. Обеспечивая рекурсивную регуляцию больших наборов генов из одного локуса, эта регуляторная логика создает основу для семиотического построения и появления все более высоких уровней интерпретационной компетентности. Следовательно, семиотическое ограничение постепенно переносится от молекул к клеткам, тканям и структуре тела (Ibid.). С каждым более высоким уровнем перемещения на новый уровень субстрата возникает форма рекурсии более высокого порядка.

### 3. Иерархия языков регуляции генов

Вышеизложенные соображения семиотического характера хорошо иллюстрируются конкретными биологическими данными о многоуровневом регулировании процессов развертки генетической информации. Поэтому далее мы рассмотрим в минимально необходимых деталях иерархию языков регуляции генов в процессах развертывания генетической (и эпигенетической) информации.

Охарактеризуем принципы реализации информации в развитии многоклеточного зародыша. В основной части нашего изложения мы будем оставаться в пределах известных позиций геноцентризма, рельефно сформулированных еще Докинзом (Dawkins, 1976), и сосредоточимся на роли генов в организации процессов зародышевого развития. Гены поэтапно активируются с самого начала этого процесса, формируя все более многочисленные и сложные ансамбли, основанные на регуляторных взаимодействиях и именуемые генными регуляторными сетями (ГРС). В основе их функциональной архитектуры заложено то базовое свойство, что данный ген сети кодирует белок, основная функция которого находить, узнавать и связываться с другими генами данной сети. Когда данный белок специфически связывается с данным геном, то он оказывает на этот ген-мишень регулирующее действие: или активирует его, или наоборот подавляет его активность.

Регуляция генов осуществляется «языками» (или кодами) регуляции. Можно выделить по крайней мере четыре языка регуляции генов (табл.). Это язык межгенного взаимодействия (1); межклеточный сигнальный язык (2); язык биохимических регуляций клеток на расстоянии (3); языки регуляции иммунной и нервной систем (4).



### Основные языки реализации (развертывания) генетической информации в развитии многоклеточного организма

«Языки» (коды)	Факторы-сигналы	Специфика	Конечная мишень
Язык межгенного взаимодействия	Белки – факторы транскрипции	Функционирует в пределах одной клетки	Гены
Межклеточный сигнальный язык	Биохимические факторы	Обеспечивает коммуникацию между соседними клетками	
Язык биохимических регуляций клеток на расстоянии	Биохимические факторы (факторы роста, гормоны)	Основа на химических сигналах, способных преодолевать значительные расстояния в пределах организма	
Языки регуляции иммунной и нервной систем	Биохимические факторы (нейромедиаторы)	«Дальнодействующие» сигналы передаются кровотоком	

#### 4. Язык межгенного взаимодействия

*Гены несут в себе наборы инструкций по их считыванию:* ген не только является «носителем генетической информации», как это зачастую понимается и декларируется. Разумеется, он кодирует последовательность аминокислот в белке. Однако ген типично включает себя обширный набор регуляторных элементов, вовлекаемых в контроль генной активности (рис. 1). Регуляции активности генов определяют, какие именно гены, в какое время, в каких эмбриональных клетках, сколь интенсивно и как долго будут активны. Базовым элементом генных регуляторных областей резонно рассматривать специфические сайты связывания. Это относительно короткие последовательности ДНК, узнаваемые белками, участвующими в контроле активности генов. Прежде всего это белки – транскрипционные факторы. Помимо них в генной регуляции участвуют другие белки, способные, в частности, взаимодействовать с самими транскрипционными факторами или действовать на организацию хроматина. Сайты связывания объединяются в кластеры – регуляторные модули. В итоге интеграции всех регуляторных импульсов (прямых и опосредованных) ген активируется в данное время в данном месте и с определенной силой (рис. 1). Наличие меняющейся совокупности опосредованных регуляторных связей естественно понимать как контекст. Это наводит на аналогии с лингвистикой.

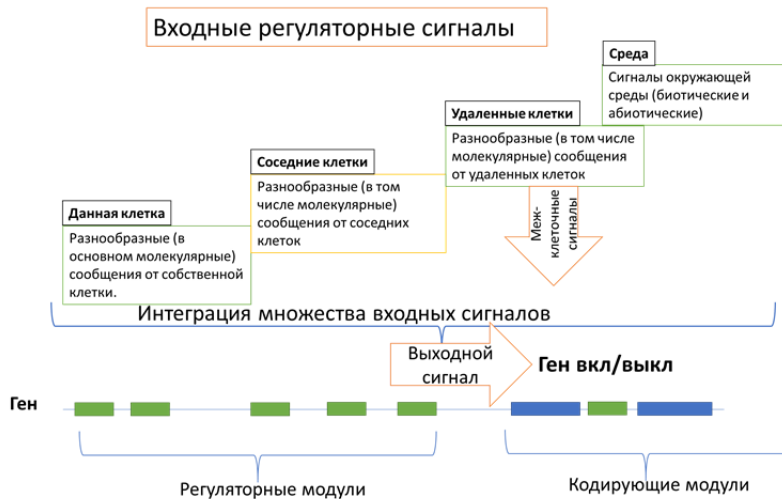


Рис. 1. Общая схема функциональной организации гена, ответственной за регуляторные взаимодействия

*Язык межгенных взаимодействий.* Рассмотрим язык, функционирующий на уровне ГРС, и в качестве образца возьмем только один тип регуляторного модуля. Это эволюционно наиболее древний и базовый компонент, имеющийся у всех генов и называемый промотером (рис. 2).

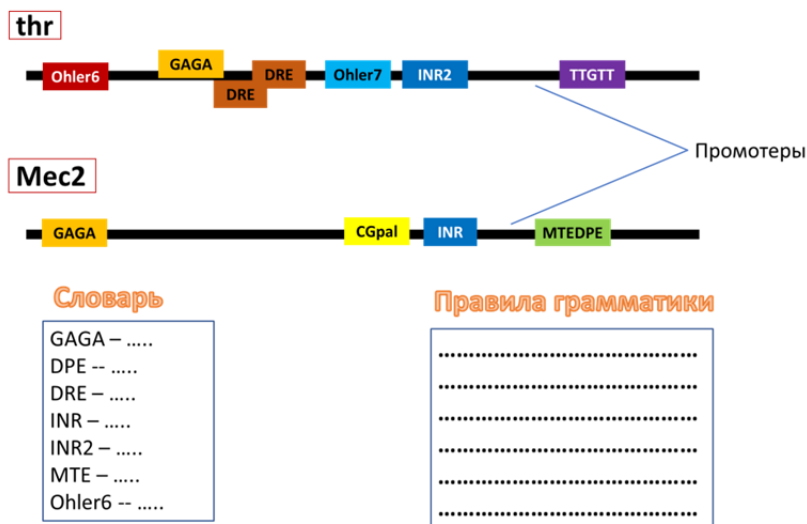


Рис. 2. Примеры генных регуляторных модулей (промоторов) генов многоклеточных ядерных организмов (дрозофила). Модули включают «сигнальные слова», для которых имеется словарь и набор грамматических правил

Именно промотер включает механизм кодирования генов: после появления промотера запускается механизм декодирования гена и на основе этого генерация белка. Промотер многоклеточных организмов



состоит из небольшого набора «слов» (примерно 4–10), составляющих словарь регуляторных элементов (сайтов связывания транскрипционных факторов) промоторов (подробнее см.: (Спи́ров, 2023)). Под «словом» мы понимаем следующий в определенном порядке набор «букв» (нуклеотидов). Словарь включает два-три десятка таких слов. Помимо словаря формулируется набор грамматических правил их использования (рис. 2). Например, имеются обязательные «слова», без которых промотор не будет функционировать; ряд «слов» демонстрирует жесткий или же предпочтительный порядок в типичном промоторе, многие не имеют жесткой позиции. В целом промотор можно уподобить фразе (предложению) на языке промоторов. Точнее, его можно интерпретировать как инструкцию для функционирования промотора, как набор команд, актуализируемых при различных условиях. Промотор во взаимодействии с другими регуляторными модулями (которые организованы в целом сходно) интегрирует входящие сигналы (как на рис. 1) и контролирует активность гена во времени индивидуального развития и в пространстве эмбриона / организма.

Отметим, что языки более высокого уровня также можно охарактеризовать, описав их с помощью словарей и правил грамматики (набор ограничений на употребление слов; см., напр.: (Antebi et al., 2017; Daniels et al., 2022)).

## 5. Гены функционируют в пространстве организма

С увеличением количества клеток в эмбрионе появляется новый уровень контроля активности генов. Это клеточная сигнализация: эмбриональные клетки начинают обмениваться специфическими биохимическими «сигналами» (рис. 3) (Armingol et al., 2021). Эти сигналы, продуцируемые одними клетками, специфически действуют на рецепторы других клеток. В результате рецепторы активируют каскады передачи сигналов, приводящих к включению или выключению определенных генов клетки-мишени. Соответственно контекст, определяющий реакцию гена на регуляторные молекулы, выходит за пределы одной клетки, и регуляторные связи становятся еще более распределенными в пространстве эмбриона. Ген получает сигналы от своей клетки, от соседних клеток, от удаленных клеток, из внешней среды. Ген функционирует как молекулярная «машина», которая обрабатывает непрерывные потоки информации, принимает решения, отдает команды и сам же их исполняет (как показано на рис. 1).

Охарактеризуем передачу информации в ходе зародышевого развития. Существенно, что процессы генной регуляции развертываются не только во времени (по стадиям развития эмбриона), но и в пространстве эмбриона. Пространственная размерность чрезвычайно важна. Регуляция генов сначала происходит в единственной клетке зиготе и затем в наборе функционально и физически связанных друг с другом клеток раннего эмбриона. Далее процессы развертываются во все более обособляющихся эмбриональных закладках (часто определяемых как морфогенетические поля).

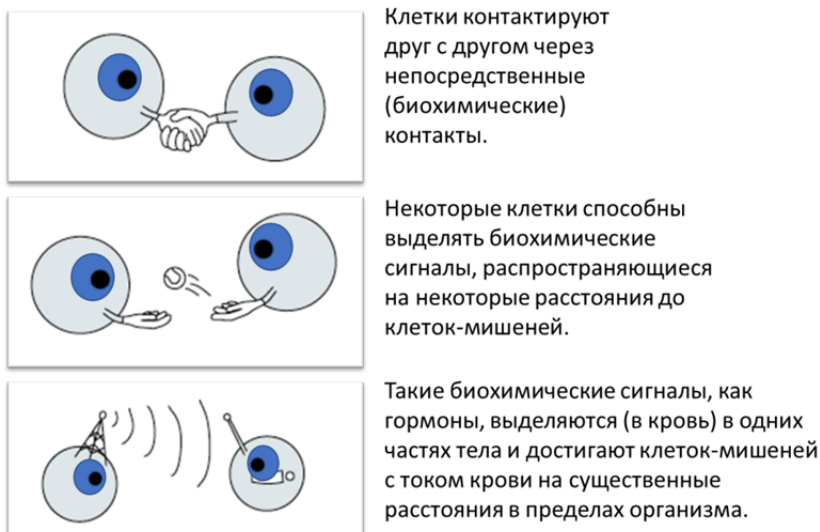


Рис. 3. Биохимические языки общения клеток (близкодействующие и далекодействующие)

Биологи-экспериментаторы начала прошлого века определяли свою область знаний как механику развития. В этой области весьма популярна была концепция карт эмбриональной судьбы (*fate maps*). В современной биологии достаточно полно описаны механизмы и процессы, определяющие судьбы клеток в зародышевом развитии, так что карты судьбы имеют под собой вполне конкретные биохимические механизмы и определяются прежде всего языками (наборами сигналов) межклеточного взаимодействия (как показано на рис. 3; см.: (Matsuda et al., 2016; Спиров, Мясникова, 2019)). В этой связи нам представляется вполне работающей *метафора карт политической географии*, причем на основе языков общения как в пределах географических территорий, так и на границах между территориями (рис. 4).

Как иллюстрирует рисунок 4, *а*, в эмбриональном развитии происходит обособление (структурное и функциональное) территорий – зачатков будущих органов, так что в пределах территорий клетки обмениваются собственным набором биохимических сигналов. При этом на границах территорий работают уже другие наборы сигналов, специально предназначенные для формирования структурных и функциональных границ между территориями. Это, на наш взгляд, весьма напоминает процессы в политической географии, когда в пределах территорий используются свои языки, тогда как на границах могут формироваться или использоваться другие языки для трансграничных коммуникаций. Именно так функционируют биохимические языки клеточных коммуникаций в пределах морфогенетических территорий и на границах территорий (рис. 4, *б*).



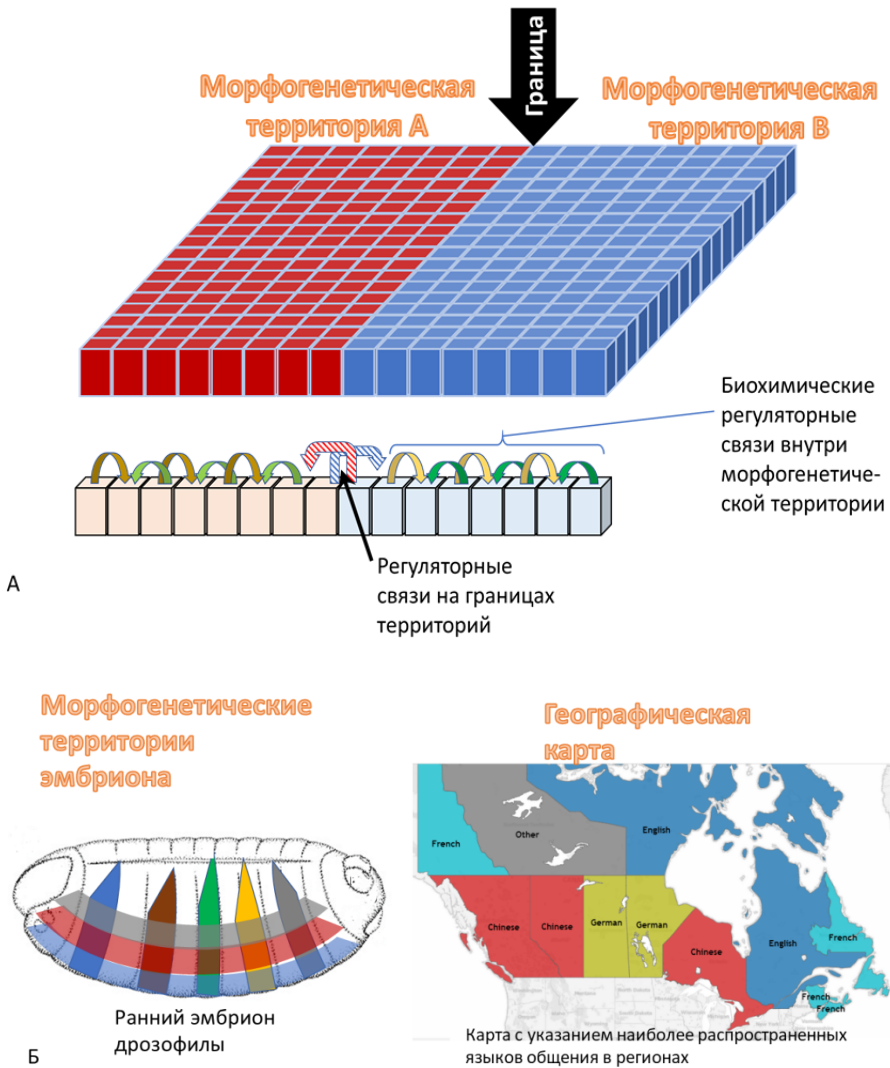


Рис. 4. Метафора карт политической географии<sup>1</sup>

Мы убеждены, что нарастающая пространственная распределенность регуляторных процессов, нарастающая контекстная зависимость, нарастающая множественность и иерархичность «языков» (кодов) инструкций нуждается в продолжении углубленного изучения с позиций семиотики. Это особо важно в той связи, что нейрональная регуляция,

<sup>1</sup> В процессе зародышевого развития эмбрион разбивается на все более детализированные многоклеточные территории (морфогенетические «поля»; а). В пределах каждой территории функционируют свои языки межклеточного общения. На границах территории действуют свои биохимические языки. Все это подводит к политико-географической метафоре (б).



включая формирование и дальнейшее функционирование центральной нервной системы, также во многом основана на очерченных здесь комплексных процессах.

Отметим, что в этих кратко охарактеризованных здесь процессах задействован ряд кодов, привлечших пристальное внимание исследователей из области код-биологии (Prinz, 2023). Здесь используются транскрипционные коды по классификации Принца, включая НОХ-коды (кода на основе генов семейства НОХ), коды сигналинга, pathway-коды, микроРНК коды, мембранные коды и, наконец, эпигенетические коды, включая гистоновые коды (Ibid.).

## **6. Контекст-зависимые и эпигенетические механизмы регуляции генов**

Контекст в процессах и механизмах контроля активности определенного гена определяется как совокупность всех опосредованных (точнее: не непосредственных) воздействий. Их нелегко выявить, и их может быть очень много. На уровне экспериментов и наблюдений это проявляется в том, что одни и те же белки-регуляторы по-разному действуют на ген-мишень в разных клетках / тканях в сходных экспериментах. На молекулярном уровне генной регуляции выделяются, в частности, факторы-коактиваторы (корепрессоры), факторы-посредники (медиаторы).

Параллельно с умножением клеток и формированием морфогенетических полей (зачатков, территорий) происходит развертывание эпигенетических механизмов регуляции, когда ядра зачастую соседних (или даже дочерних) клеток по-разному биохимически модифицируют белки хроматина, но при этом обычно не меняется сама последовательность нуклеотидов (букв) оснований в ДНК, то есть текст, кодирующий данный ген. Это создает еще один уровень в процессах реализации генетической информации. При этом процессы генерации и передачи инструкций на этих иерархических уровнях становятся все более распределенными в пространстве эмбриональных зачатков.

## **7. Ограничения и скаффолдинг в молекулярно-генетических процессах**

Высказанные положения имеет смысл развивать далее на основе построений Дикона, Патти и Хоффмайера о природе биологической (молекулярно-генетической и молекулярно-эпигенетической) информации (Pattee, 2001; Hoffmeyer, 2014; Deacon, 2021) используя концепцию биосемиотического скаффолда (основа, каркас, строительные леса, см.: (Hoffmeyer, 2014)) и ограничений (constraints). В самой общей формулировке биосемиотический скаффолдинг и ограничения описываются как набор средств, способствующих направлению процессов по определенным траекториям. Такой семиотический каркас в регуляции активности генов включает биохимические структуры и процессы, которые облегчают интерпретацию и использование генетической информации посредством знаковых процессов (семиозис).



Такие средства включают в себя несколько уровней организации — от простых химических взаимодействий до сложных биологических процессов. Каждый уровень опирается на уровни ниже себя и контролирует их, создавая структурированную иерархию, в которой более высокие уровни могут влиять на более низкие уровни. Это вполне соответствует охарактеризованной здесь нами иерархии языков регулирования активности генов.

Различные уровни скаффолда взаимозависимы, а это означает, что изменения или ограничения на одном уровне могут иметь каскадные последствия по всей иерархии (Deacon, 2021). Такая взаимозависимость имеет решающее значение для возникновения семиотических процессов, в которых значение и информация распределяются и интерпретируются на различных уровнях. В рамках наших построений это можно трактовать как иерархию от самого нижнего уровня генных регуляторных сетей в самом раннем эмбриогенезе до пространственно распределенных систем защиты, регулирования и интеграции всего многоклеточного организма при участии языков иммунной и нервной регуляции.

Семиотические функции, такие как передача сигналов, репрезентация и интерпретация, возникают в результате динамических взаимодействий внутри скаффолда. Дикон показывает, что эти функции не присущи отдельным молекулам, а возникают из сети взаимодействий, которые создают смысловой контекст (Ibid.). Исследователь убедительно продемонстрировал это на серии примеров начиная от базовых регуляторных биохимических взаимодействия и заканчивая механизмами регуляции генов.

## 8. Заключение

*Комплексность, пространственная распределенность и иерархичность* — это фундаментальные характеристики процессов реализации генетической информации.

Сложность (комплексность):

- В любом организме в регуляцию рассматриваемого процесса обычно вовлечено более одного языка (кода). В частности, это справедливо для процессов развертывания наследственной информации (табл.).

- Эти языки обычно формируют более сложную систему, меняющуюся в пространстве и во времени организма.

Иерархичность:

- Системы языков регуляции процессов развертывания наследственной информации типично являются иерархическими. Мы можем выделить как язык нижнего уровня, так и языки более высокого уровня.

- Распределенность в пространстве:

- Системы языков и их иерархичности меняются не только во времени организма, но и в организменном пространстве (рис. 4).

- В каждом малом участке развивающегося многоклеточного организма функционирует своя версия иерархии языков регуляции.

Под *комплексностью* мы понимаем то, что на гены как молекулярные машины хранения и реализации генетической информации действует, зачастую одновременно, множество управляющих сигналов (часто противоположных по смыслу — например, повысить и понизить



генную активность). Молекулярная машина любого гена в частности и генома многоклеточного организма вообще организована так, что все входящие, управляющие сигналы интегрируются и их итоговое действие определяет уровень активности данного гена в данное время (рис. 1).

*Иерархичность* мы понимаем в том смысле, что сигналы (коды, или, образно говоря, «языки» управления генами) имеют несколько взаимосвязанных уровней. Нижний уровень — это прежде всего непосредственная активация или репрессия данного гена специфическими факторами транскрипции, кодируемыми его соседями по ГРС (это обычно уровень конкретного ядра данной клетки). На следующем уровне мы включаем в рассмотрение обмен химическими сигналами между клетками. Здесь имеются другие механизмы и другие мессенджеры. Однако конечная цель — регулирование активности данного гена в данной клетке по сигналам от других клеток. На еще более общем уровне интеграции генной регуляции резонно описывать действие интегрирующих «системных» регуляторов как морфогены, гормоны, нейротрансмиттеры, цитокины.

*Пространственная распределенность* — это часто упускаемое свойство, оно характеризует то, что эмбрион, как и организм после рождения, составлен из многоклеточных интегрированных ансамблей — органов и тканей, в свою очередь составленных из клеток так, что в каждой клетке в любой отрезок времени происходят свои процессы реализации генетической информации. Даже две соседние клетки, соединенные межклеточными контактами, могут отличаться по профилям генной активности, и именно эти отличия часто критичны для дальнейшего развития или функционирования многоклеточного организма. Мы могли бы образно представить клетки как члены сообществ, населяющих различные соседствующие территории, причем территории обычно имеют четкие границы (рис. 4). При этом каждая клетка имеет свои функции, они обмениваются сигналами с соседними клетками и получают сигналы из удаленных центров или даже из внешней среды (рис. 3).

В целом же наш анализ приводит к общему заключению, что регуляция генов включает семиотические процессы, в которых различные молекулярные сигналы (факторы транскрипции, энхансеры и репрессоры) действуют как знаки, несущие инструкции для экспрессии генов. При этом экспрессия генов зависит от контекста. Это означает, что одна и та же генетическая последовательность может интерпретироваться по-разному в зависимости от клеточной среды и присутствия конкретных регуляторных молекул. Физические и функциональные каркасы (скаффолды), такие как белковые комплексы и архитектура хроматина, обеспечивают необходимую основу для семиотических взаимодействий. Эти структуры организуют и стабилизируют компоненты, участвующие в регуляции генов.

Далее активность генов регулируется посредством динамических сетей взаимодействий между ДНК, РНК, белками и другими молекулами. Эти взаимодействия образуют сложные регуляторные цепи, которые реагируют на внутренние и внешние сигналы. При этом генная регуляция интегрирована с метаболическими состояниями клеток. Внешние сигналы (например, гормоны, питательные вещества) пере-



даются в клетку и интерпретируются через сигнальные пути, которые в конечном итоге влияют на активность генов. Эти пути действуют как семиотические процессы, преобразующие внешнюю информацию в клеточные реакции.

**Благодарности.** Исследование выполнено при финансовой поддержке РФФ, проект «Междисциплинарные методологические основания расширенного эволюционного синтеза в науках о жизни и обществе» №22-18-00383, в ИНИОН РАН, [https://rscf.ru/prjcard\\_int?22-18-0038](https://rscf.ru/prjcard_int?22-18-0038).

### Список литературы

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

Спиров А.В. «Языки» регуляции генов на примере сигнальных элементов промоторов // МЕТОД: Московский ежегодник трудов из обществоведческих дисциплин. 2023. Вып. 4 (в печати).

Спиров А.В., Мясникова Е.М. Эволюционный консерватизм генных регуляторных сетей временной спецификации нейробластов // Молекулярная биология. 2019. Т. 53, №2. С. 225–239. doi: 10.1134/S0026898419020150.

Antebi Y.E., Linton J.M., Klumpe H. et al. Combinatorial Signal Perception in the BMP Pathway. *Cell*. 2017. Vol. 170, №6. P. 1184–1196.e24. doi: 10.1016/j.cell.2017.08.015.

Armingol E., Officer A., Harismendy O. et al. Deciphering cell–cell interactions and communication from gene expression // *Nature Reviews Genetics*. 2021. Vol. 22. P. 71–88. doi: 10.1038/s41576-020-00292-x.

Aovsec Ž., Weilert M., Shrikumar A. et al. Base-resolution models of transcription-factor binding reveal soft motif syntax // *Nature Genetics*. 2021. Vol. 53, №3. P. 354–366. doi: 10.1038/s41588-021-00782-6.

Bentsen M., Heger V., Schultheis H. et al. TF-COMB – Discovering grammar of transcription factor binding sites // *Computational and Structural Biotechnology Journal*. 2022. Vol. 20. P. 4040–4051. doi: 10.1016/j.csbj.2022.07.025.

Daniels K.G., Wang S., Simic M.S. et al. Decoding CAR T cell phenotype using combinatorial signaling motif libraries and machine learning // *Science*. 2022. Vol. 378. P. 1194–1200. doi: 10.1126/science.abq0225.

Dawkins R. *The selfish gene*. N. Y., 1976.

Deacon T.W. How molecules became signs // *Biosemiotics*. 2021. Vol. 14. P. 537–559. doi: 10.1007/s12304-021-09453-9.

Hoffmeyer J. The semiome: From genetic to semiotic scaffolding // *Semiotica*. 2014. Vol. 198. P. 11–31. doi: 10.1515/sem-2013-0099.

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

Lal A. Deciphering the regulatory syntax of genomic DNA with deep learning // *Journal of Biosciences*. 2022. Vol. 47: 47. doi: 10.1007/s12038-022-00291-6.

Matsuda S., Harmansa S., Affolter M. BMP morphogen gradients in flies // *Cytokine & Growth Factor Reviews*. 2016. Vol. 27. P. 119–127. doi: 10.1016/j.cytogfr.2015.11.003.

Pattee H.H. *The Physical Basis of Coding and Reliability in Biological Evolution* // *Towards a Theoretical Biology* / ed. by C.H. Waddington. Edinburgh, 1968. Vol. 1: *Prolegomena*. P. 67–93.



Pattee H.H. How Does a Molecule Become a Message? // Communication in Development, Developmental Biology Supplement / ed. by A. Lang. Cambridge, 1969. P. 116.

Pattee H.H. The physics of symbols: Bridging the epistemic cut // Biosystems. 2001. Vol. 60. P. 5–21. doi: 10.1016/S0303-2647(01)00104-6.

Prinz R. Nothing in evolution makes sense except in the light of code biology // Biosystems. 2023. Vol. 229: 104907. doi: 10.1016/j.biosystems.2023.104907.

Raible W. Linguistics and genetics: systematic parallels // Language typology and language universals = Sprachtypologie und sprachliche Universalien = La typologie des langues et les universaux linguistiques / ed. M. Haspelmath, E. König, W. Oesterreicher, W. Raible. Berlin, 2001. Vol. 1. P. 103–123.

Weingarten-Gabbay S., Segal E. The grammar of transcriptional regulation // Human Genetics. 2014. Vol. 133, №6. P. 701–711. doi: 10.1007/s00439-013-1413-1.

Zolyan S. From matter to form: the evolution of the genetic code as semio-poiesis // Semiotica. 2022. №245. P. 17–61. doi: 10.1515/sem-2020-0088.

Zolyan S. On the minimal elements of the genetic code and their semiotic functions (degeneracy, complementarity, wobbling) // Biosystems. 2023. Vol. 231:104962. doi: 10.1016/j.biosystems.2023.104962.

### Об авторе

Александр Владимирович Спиров, кандидат биологических наук, старший научный сотрудник, Институт научной информации по общественным наукам РАН; Институт эволюционной физиологии и биохимии им. И. М. Сеченова РАН, Россия.

E-mail: alexander.spirov55@gmail.com

#### Для цитирования:

Спиров А. В. Языки развертывания наследственной информации в эмбриогенезе: лингвосемиотические аналоги и аналогии // Слово.ру: балтийский акцент. 2024. Т. 15, №4. С. 25–40. doi: 10.5922/2225-5346-2024-4-2.



ПРЕДСТАВЛЕНО ДЛЯ ВОЗМОЖНОЙ ПУБЛИКАЦИИ В ОТКРЫТОМ ДОСТУПЕ В СООТВЕТСТВИИ С УСЛОВИЯМИ ЛИЦЕНЗИИ CREATIVE COMMONS ATTRIBUTION (CC BY) ([HTTP://CREATIVECOMMONS.ORG/LICENSES/BY/4.0/](http://creativecommons.org/licenses/by/4.0/))

## LANGUAGES OF UNFOLDING HEREDITARY INFORMATION IN EMBRYOGENESIS: LINGUO-SEMIOTIC ANALOGUES AND ANALOGIES

*Alexander V. Spirov*<sup>1, 2</sup>

<sup>1</sup> Institute of Information on Social Sciences of the Russian Academy of Sciences,  
51/21, Nakhimovskii Prospekt, Moscow, 117418, Russia

<sup>2</sup> Sechenov Institute of Evolutionary Physiology and Biochemistry  
of the Russian Academy of Sciences,

44, Thorez Prospekt, Saint Petersburg, 194223, Russia

Submitted on 12.04.2024

Accepted on 15.06.2024

doi: 10.5922/2225-5346-2024-4-2

*It is known that almost all hereditary information about the innumerable characteristics of a multicellular organism, including the human body, is encoded in a certain way in the nucleus of a fertilized egg. The principles of the unfolding of genetic information in the devel-*



opment of a multicellular embryo have long attracted the attention of both biologists and representatives of various sciences. While molecular biologists concentrate on the informational and cybernetic aspects of the storage and transmission of genetic information, the authors of biosemiotic studies insist on the specificity of biological signs in these processes and on the special nature of biological texts. We focused on the information aspects of these processes and, in order to demonstrate the analogies between them and with linguistic-semiotic concepts, we describe the problems of complexity and hierarchy of genetic mechanisms for the implementation of genetic information in embryogenesis, as it is seen in systems biology.

**Keywords:** biosemiotics, linguistics, hereditary information, information unfolding in individual development, languages of gene regulation

**Acknowledgement.** This research is supported by the grant of the Russian Science Foundation (project № 22-18-00383 “Methodological design of extended evolutionary synthesis”) carried out at the Institute of Scientific Information on Social Sciences of the Russian Academy of Sciences (Moscow).

### References

- Antebi, Y.E., Linton, J.M., Klumpe, H. et al., 2017. Combinatorial Signal Perception in the BMP Pathway. *Cell*, 170 (6), pp. 1184–1196, <https://doi.org/10.1016/j.cell.2017.08.015>.
- Armingol, E., Officer, A., Harismendy, O. et al., 2021. Deciphering cell–cell interactions and communication from gene expression. *Nature Reviews Genetics*, 22, pp. 71–88, <https://doi.org/10.1038/s41576-020-00292-x>.
- Avsec, Ž., Weilert, M., Shrikumar, A. et al., 2021. Base-resolution models of transcription-factor binding reveal soft motif syntax. *Nature Genetics*, 53 (3), pp. 354–366, <https://doi.org/10.1038/s41588-021-00782-6>.
- Bentsen, M., Heger, V., Schultheis, H. et al., 2022. TF-COMB – Discovering grammar of transcription factor binding sites. *Computational and Structural Biotechnology Journal*, 20, pp. 4040–4051, <https://doi.org/10.1016/j.csbj.2022.07.025>.
- Daniels, K.G., Wang, S., Simic, M.S. et al., 2022. Decoding CAR T cell phenotype using combinatorial signaling motif libraries and machine learning. *Science*, 378, pp. 1194–1200, <https://doi.org/10.1126/science.abq0225>.
- Dawkins, R., 1976. *The selfish gene*. New York, 224 p.
- Deacon, T.W., 2021. How Molecules Became Signs. *Biosemiotics*, 14, pp. 537–559, <https://doi.org/10.1007/s12304-021-09453-9>.
- Hoffmeyer, J., 2014. The semiome: From genetic to semiotic scaffolding. *Semiotica*, 198, pp. 11–31, <https://doi.org/10.1515/sem-2013-0099>.
- Jakobson, R., 1970. Linguistics. Relationship between the science of language and other sciences. In: *Main Trends of Research in the Social and Human Sciences*. Mouton, pp. 419–453.
- Lal, A., 2022. Deciphering the regulatory syntax of genomic DNA with deep learning. *Journal of Biosciences*, 47, <https://doi.org/10.1007/s12038-022-00291-6>.
- Matsuda, S., Harmansa, S. and Affolter, M., 2016. BMP morphogen gradients in flies. *Cytokine & Growth Factor Reviews*, 27, pp. 119–127, <https://doi.org/10.1016/j.cytogfr.2015.11.003>.
- Pattee, H.H., 1968. The Physical Basis of Coding and Reliability in Biological Evolution. In: C.H. Waddington, ed. *Towards a Theoretical Biology*. Vol. 1, Prolegomena. Edinburgh, pp. 67–93.
- Pattee, H.H., 1969. How Does a Molecule Become a Message? In: A. Lang, ed. *Communication in Development, Developmental Biology Supplement*. Cambridge, pp. 1–16.
- Pattee, H.H., 2001. The physics of symbols: Bridging the epistemic cut. *Biosystems*, 60, pp. 5–21, [https://doi.org/10.1016/S0303-2647\(01\)00104-6](https://doi.org/10.1016/S0303-2647(01)00104-6).



Prinz, R., 2023. Nothing in evolution makes sense except in the light of code biology. *Biosystems*, 229:104907, <https://doi.org/10.1016/j.biosystems.2023.104907>.

Raible, W., 2001. Linguistics and genetics: systematic parallels. In M. Haspelmath, E. König, W. Oesterreicher and W. Raible, eds. *Language typology and language universals/ Sprachtypologie und sprachliche Universalien/ La typologie des langues et les universaux linguistiques*. Vol. 1. Berlin, pp. 103 – 123.

Solov'ev, V. V., Kel', A. E., Rogozin, I. B. and Kolchanov, N. A., 1988. *Ispol'zovanie EVM v molekulyarnoi biologii. Vvedenie v teoriyu geneticheskikh tekstov* [Use of computers in molecular biology. Introduction to the theory of genetic texts]. Novosibirsk, 92 p. (in Russ.).

Spirov, A. V. and Myasnikova, E. M., 2019. Evolutionary Stability of Gene Regulatory Networks That Define the Temporal Identity of Neuroblasts. *Molekulyarnaya biologiya* [Molecular Biology], 53 (2), pp. 225 – 239, <https://doi.org/10.1134/S0026898419020150> (in Russ.).

Spirov, A. V., 2023. "Languages" of gene regulation using the example of signal elements of promoters. In: *Metod: moskovskii ezhegodnik trudov iz obshchestvovedcheskikh distsiplin* [METHOD: Moscow Yearbook of Works from Social Science Disciplines], 4 (in print) (in Russ.).

Weingarten-Gabbay, S. and Segal, E., 2014. The grammar of transcriptional regulation. *Human Genetics*, 133 (6), pp. 701 – 711, <https://doi.org/10.1007/s00439-013-1413-1>.

Zolyan, S., 2022. From matter to form: the evolution of the genetic code as semio-  
poiesis. *Semiotica*, 245, pp. 17 – 61, <https://doi.org/10.1515/sem-2020-0088>.

Zolyan, S., 2023. On the minimal elements of the genetic code and their semiotic functions (degeneracy, complementarity, wobbling). *Biosystems*, 231:104962, <https://doi.org/10.1016/j.biosystems.2023.104962>.

### The author

Dr. Alexander V. Spirov, Senior Researcher, Institute of Information on Social Sciences of the Russian Academy of Sciences; Senior Researcher, Sechenov Institute of Evolutionary Physiology and Biochemistry of the Russian Academy of Sciences, Russia.

E-mail: alexander.spirov55@gmail.com

#### To cite this article:

Spirov, A. V., 2024, Languages of unfolding hereditary information in embryogenesis: linguo-semiotic analogues and analogies, *Slovo.ru: Baltic accent*, Vol. 15, no. 4, pp. 25 – 40. doi: 10.5922/2225-5346-2024-4-2.

