



«ВОДЯНЫЕ ДНК-ЗНАКИ» В ВИДЕ ВСТРОЕННЫХ В ГЕНОМЫ РАЗЛИЧНЫХ ОРГАНИЗМОВ ЧУЖЕРОДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДНК И НЕ ТОЛЬКО ВСТРОЕННЫХ

¹Сахабутдинова А.Р., ²Михайленко К.И., ¹Гарафутдинов Р.Р., ³Кирьянова О.Ю., ⁴Сагитова М.А.,
²Привалов Л.Ю., ¹Гималов Ф.Р., ¹Баймиев Ан.Х., ¹Кулуев Б.Р., ¹Баймиев Ал.Х., ¹Чемерис А.В.

¹Институт биохимии и генетики – обособленное структурное подразделение Федерального государственного бюджетного научного учреждения Уфимского федерального исследовательского центра Российской академии наук, Россия, 450054, Уфа, пр. Октября, 71, E-mail: chemeris@anrb.ru

²Институт механики – обособленное структурное подразделение Федерального государственного бюджетного научного учреждения Уфимского федерального исследовательского центра Российской академии наук, Россия, 450054, Уфа, пр. Октября, 71

³Уфимский государственный нефтяной технический университет, Россия, 450062, Уфа, ул. Космонавтов, 1

⁴Московский физико-технический институт, Россия, 141701, г. Долгопрудный, Институтский пер., 9

Резюме

Водяные ДНК-знаки являются надежной технологией маркировки рекомбинантных штаммов бактерий, рас дрожжей, сортов и линий растений, включая ГМО. Водяной ДНК-знак представляет собой относительно короткую последовательность нуклеотидов, встроенную в ДНК какого-либо организма, с целью доказать его генно-инженерное или полностью синтетическое происхождение. В водяных ДНК-знаках может быть закодирована фактически любая небιοлогическая информация в виде фамилий авторов, мест их работы, неких уникальных текстов и пр. Причем разными авторами применяются различные способы превращения такой негенетической информации в нуклеотидные последовательности, кратко рассмотренные в данном обзоре. Помимо встраиваемых в геномы маркерных участков ДНК, водяными ДНК-знаками можно считать и добавляемые в различные среды и субстанции (подземные воды, нефть, молоко, чернила, пестициды и пр.) некие уникальные нуклеотидные последовательности, представляющие собой, в том числе химически синтезированные протяженные олигонуклеотиды, которые можно затем детектировать с помощью ПЦР. Уделено некоторое внимание в этом обзоре и проблеме ГМО, которой, на самом деле, существовать не должно в силу множества причин, главная из которых заключается в том, что ГМ-растения ничуть не опаснее обычных сортов сельскохозяйственных растений.

Ключевые слова: ДНК, геном, водяные ДНК-знаки, хранение в ДНК, прокариоты, эукариоты, ГМО, нефть, подземные воды

Цитирование: Сахабутдинова А.Р., Михайленко К.И., Гарафутдинов Р.Р., Кирьянова О.Ю., Сагитова М.А., Привалов Л.Ю., Гималов Ф.Р., Баймиев Ан.Х., Кулуев Б.Р., Баймиев Ал.Х., Чемерис А.В. «Водяные ДНК-знаки» в виде встроенных в геномы различных организмов чужеродных последовательностей ДНК и не только встроенных // *Biomics*. 2021. Т.13(4). С. 368-386. DOI: 10.31301/2221-6197.bmcs.2021-26

© Автор(ы)

"DNA WATERMARKS" IN THE FORM OF ALIEN DNA SEQUENCES INSERTED IN THE GENOMES OF VARIOUS ORGANISMS AND NOT ONLY INSERTED ONES

¹Sakhabutdinova A.R., ²Mikhailenko K.I., ¹Garafutdinov R.R., ³Kiryanova O.Yu., ⁴Sagitova M.A.,
²Privalov L.Yu., ¹Baymiev An.Kh., ¹Kuluev B.R., ¹Gimalov F.R., ¹Baymiev Al.Kh., ¹Chemeris A.V.

¹Institute of Biochemistry and Genetics, Ufa Federal Research Center, Russian Academy of Sciences,
71 Prospekt Oktyabrya, 450054, Ufa, Russia, E-mail: chemeris@anrb.ru

²Institute of Mechanics, Ufa Federal Research Center, Russian Academy of Sciences,
71 Prospekt Oktyabrya, 450054, Ufa, Russia, E-mail: const@mail.ru

³Ufa State Petroleum Technological University, 1 Kosmonavtov str., Ufa, 450062, Russia

⁴The Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow Region, 141701, Russia

Resume

DNA watermarking is a reliable technology for marking recombinant strains of bacteria, yeast races, plant varieties and lines, including GMOs. A DNA watermark is a rather short sequence of nucleotides inserted into the DNA of any organisms in order to prove its genetically engineered or fully synthetic origin. Virtually any non-biological information in the form of authors' surnames, places of their work, some unique texts, etc. can be encoded in DNA watermarks. Moreover, different authors use various methods of converting such non-genetic information into nucleotide sequences, briefly discussed in this review. In addition to DNA marker sites embedded in genomes, water DNA signs can also be considered as some unique nucleotide sequences added to various media and substances (groundwater, oil, milk, ink, pesticides, etc.), including chemically synthesized long oligonucleotides, which can then be detected using PCR. Some attention is paid in this review to the problem of GMO, which, in fact, should not exist due to many reasons, the main of which is that GM plants are no more dangerous than ordinary varieties of agricultural plants.

Keywords: DNA, genome, DNA watermarks, storage in DNA, prokaryotes, eukaryotes, GMO, petroleum, underground waters

Citation: Sakhabutdinova A.R., Mikhailenko K.I., Garafutdinov R.R., Kiryanova O.Yu., Sagitova M.A., Privalov L.Yu., Gimalov F.R., Baymiev An.Kh., Kuluev B.R., Baymiev Al.Kh., Chemeris A.V. "DNA Watermarks" in the form of alien DNA sequences inserted in the genomes of various organisms and not only inserted ones. *Biomics*. 2021. V.13(4). P.368-386. DOI: 10.31301/2221-6197.bmcs.2021-26 (In Russian)

© Authors

Введение

В последние годы молекулы ДНК все чаще рассматриваются как весьма удобные биополимеры для их небιологического применения. В частности, они теоретически могут использоваться для молекулярных вычислений или иначе ДНК-компьютинга; для скрытой передачи данных путем ДНК-криптографии и ДНК-стеганографии; для долговременного хранения негенетических данных в ДНК взамен кремния особой чистоты, которого через некоторое время может ощущаться нехватка. Причем подобное хранение подразделяется на хранение цифровых данных в системах *in vivo* и *in vitro*, что в последнем случае предполагает еще и длительное хранение самой ДНК, что довольно подробно рассмотрено нами ранее [Сахабу̀тдинова и др. (Sakhabutdinova et al.), 2019; Гарафутдинов и др. (Garafutdinov et al.), 2020]. Однако хранение небιологической информации в ДНК в живых организмах или иначе *in vivo* очень тесно соприкасается с отдельным направлением применения молекул ДНК в виде их использования в качестве своеобразных меток, позволяющих устанавливать принадлежность биологических объектов, например, рекомбинантных штаммов бактерий или генетически-

модифицированных растений, в том числе их владельцам (создателям) с целью получить возможность доказать при необходимости их право собственности. Подобный подход уже получил название «водяных ДНК-знаков» по аналогии с таковыми, используемых на некоторых бумагах (документах) и банкнотах. При этом нужно заметить, что в англоязычной литературе, помимо термина 'DNA watermarking' для обозначения подобных меток используются также 'hiding DNA', 'tagging' или 'taggant', 'PCRTag', 'signature', либо 'barcode' или 'barcoding'. Хотя последние два термина чаще применяются при генотипической характеристике отдельных видов организмов, а также в полногеномном секвенировании для разграничения типов секвенируемых молекул ДНК, наряду с еще одним термином 'indexing'. Причем даже в одной отдельно взятой статье можно встретить разные термины, например watermark и tag или их герундии. В русскоязычной литературе лишь в единичных работах применительно к ДНК можно встретить упоминание только термина «водяные знаки», что все же на наш взгляд требует уточнения их вида, тем более, что сфера применения именно водяных ДНК-знаков неизбежно будет расширяться.

Рекомбинантные молекулы ДНК и первые прокариотические организмы, несущие чужеродные фрагменты ДНК, начали создавать еще в первой половине 1970-х гг. Эукариотические организмы со встроенным в их геномы чужеродными участками ДНК стали получать лишь десятилетие спустя¹. Но в те годы ни в те, ни в другие организмы не помещали некую небіологическую информацию, поскольку для их трансформации и трансгеноза в подавляющем большинстве случаев использовали лишь конкретные гены, кодирующие какие-либо белки, определяющие важные признаки. Кроме непосредственного встраивания участков ДНК, несущих небіологическую информацию, в геном того или иного организма, ДНК-метки, вносимые в какие-либо жидкости, либо наносимые на поверхности отдельных предметов, служат для контроля и слежения за этими объектами, что также найдет отражение в этом обзоре. При этом излагаемый материал будет подаваться насколько возможно с соблюдением хронологической последовательности событий. Насколько можем судить данная обзорная статья является фактически единственной в мировой литературе, если не считать публикацию 2011 г. [Heider, Barnekow, 2011], но далеко не все вопросы в той работе были охвачены, поскольку она имела определенную направленность, и к тому же с момента ее выхода появились новые важные сведения. Есть еще недавний небольшой обзор [Zhao et al., 2021], где упоминаются водяные ДНК-знаки, но в нем основной акцент сделан на синтетической биологии.

Водяные ДНК-знаки в геномах прокариот

Пожалуй, первый опыт внедрения в бактерию *E.coli* некой текстовой информации был осуществлен в рамках проекта Genesis (<https://www.ekac.org/geninfo.html>) при его выполнении в 1998-1999 гг., когда в последовательности ДНК была закодирована несколько сокращенная фраза из Библии “Let man have dominion over the fish of the sea and over the fowl of the air, and over every living thing that moves upon the earth” путем создания некоего искусственного гена.

¹ Справедливости ради, следует заметить, что о регенерации трансгенного растения табака из волосовидных корней, полученных с помощью бактерии *Agrobacterium rhizogenes*, было сообщено еще в 1977 г., а сами эксперименты были осуществлены еще в 1973 г. практически одновременно с получением рекомбинантных прокариот, но не будем здесь на этом останавливаться, поскольку ранее мы уже уделили достаточное большое внимание этому вопросу [Кулуев и др. (Kuluev et al.), 2015].

Для этого данный текст сначала был переведен в азбуку Морзе, которая была преобразована уже в нуклеотидную последовательность. В этом случае при кодировании нуклеотидами азбуки Морзе «точке» соответствовал цитозин, «тире» – тимин, а гуанин и аденин кодировали пробелы между буквами и между словами соответственно. В итоге данный «ген» длиной 358 нуклеотидов был синтезирован и клонирован в кишечной палочке [Кас, 2000].

Затем последовали работы, в которых в системе *in vitro* в последовательностях ДНК были закодированы короткие фразы [Clelland et al., 1999; Bancroft et al., 2001], оказавшие серьезное влияние на дальнейшее развитие событий в этой сфере кодирования азотистыми основаниями небіологической информации. В теоретической статье 2001 г. ее автор [Сох, 2001a] коснулся вопроса долговременного хранения негенетической информации в дрожжах и бациллах, в том числе в спорах, которые, как известно, способны сохранять жизнеспособность очень долго. Это, впрочем, имеет лишь опосредованное отношение к водяным ДНК-знакам, но в той же статье приведен пример кодирования короткого фрагмента одного из высказываний Нобелевского лауреата Р.Фейнмана, произнесенного им в своей знаменитой речи в конце 1959 г. [Feynman, 1960]. При этом Сох для кодирования применил преимущественно три нуклеотида лишь с крайне редким включением гуанина, сославшись на других авторов, использовавших для тэгов при масштабном секвенировании именно три нуклеотида (без G) [Brenner et al., 2000], что позволяет сглаживать температуры плавления таких олигонуклеотидов и исключает образование из них гетеродимеров. Чуть позже такой подход с соответствием неких мотивов азотистых оснований цифровым данным в виде трех нуклеотидов с периодическим включением гуанина тем же Сох с соавторами был предложен среди других вариантов кодирования небіологической информации [Smith et al., 2003]. Также в одной из теоретических статей тех лет, посвященных использованию водяных ДНК-знаков в геномах разных организмов для их мечения, рассматриваются пути сокрытия такой негенетической информации или в не кодирующих участках, не нарушающих функционирование геномов, или за счет изменения благодаря вырожденности генетического кода отдельных нуклеотидов в триплетах некоторых генов, что может придать уникальность последовательности ДНК [Shimanovsky et al., 2003].

Весьма интересна статья, где описывается долговременное хранение в ДНК небіологической информации с помощью триплетного кодирования (табл. 1) [Wong et al., 2003].

Таблица 1

Кодировка букв, цифр, знаков (по Wong et al., 2003)
 Table 1 - Encoding of letters, numbers, signs (according to Wong et al., 2003)

AAA – 0	AAC – 1	AAG – 2	AAT – 3	ACA – 4	ACC – 5	ACG – 6	ACT – 7
AGA – 8	AGC – 9	AGG – A	AGT – B	ATA – C	ATC – D	ATG – E	ATT – F
CAA – G	CAC – H	CAG – I	CAT – J	CCA – K	CCC – L	CCG – M	CCT – N
CGA – O	CGC – P	CGG – Q	CGT – R	CTA – S	CTC – T	CTG – U	CTT – V
GAA – W	GAC – X	GAG – Y	GAT – Z	GCA – SP	GCC – :	GCG – ,	GCT – -
GGA – .	GGC – !	GGG – (	GGT –)	GTA – `	GTC – ‘	GTG – “	GTT – ”
TAA – ?	TAC – ;	TAG – /	TAT – [	TCA –]	TCC –	TCG –	TCT –
TGA –	TGC –	TGG –	TGT –	TTA –	TTC –	TTG –	TTT –

Особенностью данной работы было то, что закодированные отрывки детской песенки «It's a Small World» были клонированы в бактерии *E.coli* и затем перенесены в геном микроорганизма *Deinococcus radiodurans*, взятого как вид, устойчивый к высоким дозам радиации и теоретически способного пережить любые катаклизмы. С этой целью были синтезированы семь фрагментов ДНК размерами от 57 до 99 п.н., фланкированных по краям участками для отжига праймеров, что позволило амплифицировать клонированные фрагменты ДНК и затем их секвенировать. При этом 11 триплетов из 64 возможных не были задействованы.

Несколько иная (шестибитная) кодировка арабских цифр, букв английского алфавита и ряда знаков (табл. 2) была предложена в работе [Arita, Ohashi, 2004], целью которой было создать водяные ДНК-знаки в геноме *Bacillus subtilis* путем сайт-направленного мутагенеза ряда кодонов в гене *ftsZ*, благодаря чему в нем оказалось закодировано слово KEIO по названию места работы авторов – Keio University.

Таблица 2

Шестибитная кодировка букв английского алфавита и некоторых символов по [Arita, Ohashi, 2004]

Table 2 - Six-bit encoding of letters of the English alphabet and some characters by [Arita, Ohashi, 2004]

000001 “	010011 R	001110 F	110100 J
000010 E	100011 H	010110 G	111000 Q
000100 T	001101 D	011010 W	011111 Z
001000 A	010101 L	011100 Y	101111 ’
010000 O	100101 C	100110 B	110111 ,
100000 S	011001 M	101010 V	111011 &
000111 N	101001 U	101100 K	
001011 I	110001 P	110010 X	

Не указанный в таблице код 000000 соответствовал пробелу.

The code 000000 not specified in the table corresponded to a space.

Другими авторами для внедрения в геном *B.subtilis* короткого текста он был сначала переведен в скан-код клавиатуры, затем в шестнадцатиричный код, потом в двоичный. В итоге шестнадцати возможным комбинациям динуклеотидов были присвоены двоичные четырехразрядные коды (табл. 3) [Yachie et al., 2007].

Таблица 3

Кодирование динуклеотидов двоичными четырехразрядными числами

Table 3 - Encoding of dinucleotides with binary four-digit numbers

AA = 0000	AG = 1000
CA = 0001	CG = 1001
GA = 0010	GG = 1010
TA = 0011	TG = 1011
AC = 0100	AT = 1100
CC = 0101	CT = 1101
GC = 0110	GT = 1110
TC = 0111	TT = 1111

В результате такого перевода текста «E=mc² 1905!» потребовалось создать 4 кассеты из олигонуклеотидов размерами 64, 62, 62, 62 звена, которые в составе подходящих конструкций были внедрены в геном бациллы *B.subtilis*. Последующее секвенирование подтвердило возможность извлечения из генома бактерии данной текстовой информации [Yachie et al., 2007]. С учетом того, что в живых организмах хранящаяся в них чужеродная ДНК (как, впрочем, и их собственная) может подвергаться мутациям, рекомбинациям, то чуть позже этими же авторами [Yachie et al., 2008] было предложено внедрять сразу несколько копий таких последовательностей в ряде мест с тем, чтобы при извлечении данных путем секвенирования проводить выравнивание нуклеотидных последовательностей и восстанавливать исходную информацию, считая, что во всех копиях изменения если и произойдут, то не будут одинаковыми.

В 2009 г. в журнале Nature была опубликована заметка [Gustafsson, 2009], в которой сообщалось, что еще в 2005 г. в ДНК была закодирована поэма Tomtem (автор V.Rydberg) из 50 слов, которая сначала была переведена в последовательность аминокислот, используя их однобуквенное кодирование (при этом заменив букву О на визуально схожую букву Q, соответствующую глутаминовой кислоте), и затем в последовательность нуклеотидов, после чего этот «ген» (нуклеотидная последовательность которого, точнее поэмы, хранится в GenBank под номером EU600200.1 – Synthetic construct Tomten gene, complete cds, 776 bp), был клонирован с помощью плазмидного вектора в *E.coli*.

Был использован также и подход к созданию водяных ДНК-знаков в виде замен нуклеотидов в отдельных генах, не приводящих благодаря вырожденности генетического кода к замене аминокислот и соответственно не нарушающих конформацию белков. Так, было показано, что использование замен нуклеотидов в шести позициях у трех триплетов позволяет обеспечить двадцать вариаций нуклеотидных последовательностей, детектируемых с помощью ПЦР со специфичными праймерами, что было продемонстрировано для генов *bphC* и *tomA3* микроорганизма *Burkholderia cepacia*, внедренных в *E.coli*. [Morono et al., 2007]. Также вводя молчащие мутации в 26 триплетов из 70 в гене *tatAD* *Bacillus subtilis* с учетом ASCII кодов, в них было закодировано послание ICSP08² при этом нуклеотиды А, С, G, Т соответствовали двоичным числам 00, 01, 10 и 11 [Jiao, Goutte, 2008].

Нельзя не обратить внимание на статью, в которой в одной молекуле ДНК оказались закодированными сразу текст и мелодия детской песенки «Mary Had a Little Lamb», а также графическое изображение овечки в виде простенького детского рисунка [Ailenberg, Rotstein, 2009]. Для этого были использованы соответственно три варианта кодировок нуклеотидами текста, музыки и графического изображения, после чего был создан фрагмент ДНК размером 844 п.н., который был клонирован в плазмидном векторе. Для его секвенирования (с целью извлечения текстовых, музыкальных и графических данных) использовались две пары праймеров, образующие ампликоны из 500 и 344 п.н. При этом авторы отметили, что хранить плазмиды отдельно в виде ДНК, а не в живом организме более предпочтительно, поскольку можно не опасаться мутаций. Но такое внедрение нуклеотидов с небиологической информацией за водяные ДНК-знаки считать не совсем правомерно.

Однако в той же в работе, в Приложении к ней упоминается о создании еще дополнительной индексной плазмиды (“index plasmid”), в которую была помещена следующая информация в виде нуклеотидной последовательности из 194 звеньев – “DNA archive” (заголовок), “1.1.2008” (дата), “Ailenberg, Rotstein et al.” (авторы), “SMH” (место их работы), “p3” (число страниц), “pl1” (номер плазмиды в клонотеке).

Неоднозначное предложение прозвучало в одной работе [Jupiter et al., 2010] маркировать водяными ДНК-знаками патогенные микроорганизмы с тем расчетом, чтобы можно было в случае их утечки из стен лаборатории отслеживать дальнейшее распространение в окружающей среде. С этой затеей не все так просто. Во-первых, нужно сделать все возможное для исключения подобных утечек. Во-вторых, теоретически можно выявлять принадлежность таких микроорганизмов с помощью природного полиморфизма их ДНК, который не хуже водяных ДНК-знаков позволит установить циркулирующий штамм. Для этого подобная информация должна сопровождать патогенные штаммы в коллекциях, включая данные об их полных геномах, установление которых уже и в 2010 г. не представляло неразрешимой проблемы. Однако нужно иметь в виду, что такой подход позволит определить именно штамм, но не место его утечки при хранении одинаковых штаммов в разных коллекциях. В этом случае действительно маркирование водяными ДНК-знаками хранящихся патогенных бактерий (причем каждой коллекции может быть присвоен конкретный ДНК-знак) может возыметь определенное действие и в реальности установить источник, из которого произошла утечка. С другой стороны – нужно ли такое вмешательство в геном патогенных микроорганизмов и к чему оно может привести?!

При создании под руководством К.Вентера первого полностью синтетического организма – бактерии *Mycoplasma genitalium* с размером генома 582970 п.н. в него были внедрены в межгенные участки несколько коротких водяных ДНК-знаков, свидетельствующих, что это искусственный микроорганизм [Gibson et al., 2008]. В дополнительном материале к этой статье приведены геномные координаты этих водяных знаков, из которых видно, что они расположены относительно равномерно по всему геному: WM1\CRAIGVENTER – 084823...084879; WM2\VENTERINSTITVTE – 169040...169108; WM3\HAMSMITH – 231638...231685; WM4\CINDIANDCLYDE – 314130...314272; WM5\GLASSANDCLYDE – 350422...350484. В русском переводе книги К.Вентера «Жизнь на скорости света. От двойной спирали к

² 9th International Conference on Signal Processing, состоявшаяся осенью 2008 года в Пекине

рождению цифровой биологии [Вентер, 2016; original - Venter, 2013] при описании экспериментов по конструированию *M.genitalum* “DNA watermarks” упомянуты просто как «водяные знаки».

В своей следующей аналогичной работе [Gibson et al., 2010] при синтезе полного генома другой бактерии *M.mycoides* эти авторы поместили в него четыре водяных ДНК-знака, также служащих доказательствами синтетической природы генома, и они имели уже длину 1246, 1081, 1109 и 1222 п.н., в которых закодированы их фамилии, электронные адреса, а также три цитаты - “TO LIVE, TO ERR, TO FALL, TO TRIUMPH, TO RECREATE LIFE OUT OF LIFE.”- JAMES JOYCE; “SEE THINGS NOT AS THEY ARE, BUT AS THEY MIGHT BE.” – из книги “AMERICAN PROMETHEUS”; “WHAT I CANNOT BUILD, I CANNOT UNDERSTAND.” – RICHARD FEYNMAN.

Продемонстрирована также возможность хранения *in vivo* небиологических данных, осуществленная с помощью CRISPR/Cas-технологии, когда в популяции бактерий *E.coli* были размещены ряд рисунков, включающих изображения человеческой руки и кадры, изображающие галопирующую лошадь из знаменитой хронофотографической работы 1870-х гг. Э.Мейбриджа, явившейся, по сути, первой GIF-анимацией и послужившей прообразом синематографа [Shipman et al., 2017]. Для этого авторы воспользовались тем, что бактерии, несущие CRISPR-локусы, внедряют в свою ДНК некие участки геномов атакующих их бактериофагов в виде так называемых протоспейсеров и помещают их в особое место, формируя CRISPR-кассеты. В этом случае вместо фагов с помощью электропорации внутрь бактериальных клеток *E.coli*, относящихся к типу I-E, представленных специальным штаммом со сверхэкспрессией генов белков Cas1 и Cas2, выполняющих роль интегразного комплекса, доставлялись олигонуклеотиды, несущие в себе закодированную в ДНК небиологическую информацию. Таковой послужили два рисунка руки человека: один - размером 56×56 пикселей, представленный 4 оттенками серого, каждый из которых кодировался одиночным нуклеотидом, и второй размером 30×30 пикселей, представленный уже 21 оттенком серого, кодируемых 21 триплетом. Особенностью последнего можно считать то, что такому кодированию, совпадающему со стандартным 64-буквенным генетическим кодом (в котором один триплет в их случае не кодировал ничего) придана дополнительная избыточность, поскольку любой оттенок из 21 кодируется трижды, например, номеру 2 соответствуют триплеты TCT, CAC и AGA. Для первого рисунка потребовалось использовать 784

байта в 112 протоспейсерах, а для второго хватило 100 протоспейсеров и 494 байта, но кодировался каждый цвет тремя азотистыми основаниями. Третьим файлом, выбранным для хранения, была анимационная картинка скачущей лошади из пяти кадров размером 36×26 пикселей, нарисованная 21 оттенком серого. Для нее потребовалось синтезировать 104 олигонуклеотида (протоспейсера) с триплетным кодированием цветов общим объемом 2,6 килобайта. При этом для указания в какой части изображений должен располагаться тот или иной пиксель из конкретного набора пиксетов, они сопровождалась неким штрих-кодом размером в один байт в виде четверки нуклеотидов, согласно следующего принципа кодировки: С – 00; Т – 01; А – 10; G – 11. Секвенирование пула бактерий позволило воссоздать рисунки и анимацию, полнота которых несколько варьировала и в одном из случаев достигла 96%. Таким образом, объединив принцип хранения данных в ДНК с функционирующей в бактериальных клетках системой захвата участков ДНК, можно создавать микроорганизмы, которые будут хранить и передавать информацию во времени. При этом получающиеся микроорганизмы могут считаться помеченными такими нуклеотидными последовательностями, поскольку не несут в них биологической информации и, по сути, маркированными водяными ДНК-знаками.

Водяные ДНК-знаки в геномах эукариот

Хотя выше заодно вскользь уже упоминались дрожжи, но поскольку они являются эукариотическими организмами, то уделим им некоторое внимание и в данном разделе. Помимо них, из эукариот остановимся только на растениях, с одной стороны, потому что они нам ближе с профессиональной точки зрения, а с другой – потому, что с ними получено наибольшее количество результатов по маркированию их геномов, неизбежно приводя к созданию трансгенных или генетически-модифицированных (ГМ) растений, даже если какого-либо другого целевого гена в них не внедрено. А от трансгенного растения до вышедшего на поля ГМО, распространение которого нужно контролировать, хотя и дистанция огромного размера, тем не менее, первое, по сути, этап второго. Несмотря на то, что появившаяся в последнее десятилетие технология CRISPR/Cas редактирования геномов позволяет при желании не вносить никакой чужеродной ДНК, наличие которой как раз является главной причиной боязни ГМ-растений, добиваясь при этом реальных изменений метаболизма таких растений по типу сайт-направленного мутагенеза, формально напоминающего дозволенный при создании новых сортов химический или радиационный мутагенез, во

многих странах на это не желают обращать внимание, апеллируя к задействованной генно-инженерной технологии. Более подробно мы коснулись этих вопросов в ряде наших статей, посвященных CRISPR/Cas геномному редактированию [Баймиев и др. (Baymiev et al.), 2017; Кулуев и др. (Kuluev et al.), 2019] и посему здесь на этом больше останавливаться не будем. При этом не можем удержаться, чтобы не заметить, что контролировать ГМО просто бессмысленно, поскольку таковое уже невозможно, детальное объяснение чему также можно найти в серии наших публикаций на эту тему [Чемерис и др. (Chemeris et al.), 2014; 2015; Вершинина и др. (Vershinina et al.), 2020].

До 1999 г. отношение к трансгенным растениям было вполне благодушное. Никто и не думал как-то контролировать ГМО, и в той же Великобритании (ставшей отправной точкой всемирной «ГМО-боязни») люди даже с большей охотой покупали томатную пасту из генно-инженерных помидоров сорта Flavr Savr, о чем прямо сообщалось на упаковке, поскольку она была дешевле обычной, благодаря несколько упрощенной технологии обращения с такими плодами. Но в результате неумной статьи шотландских авторов (которую не хочется цитировать, но заинтересовавшиеся читатели могут найти ее здесь - [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(98\)05860-7/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(98)05860-7/fulltext)) и выступлений тех авторов по британскому телевидению вокруг ГМО началась самая настоящая вакханалия, продолжающаяся во многих странах и по сей день. К тому же Европейский союз в 2001 г. принял директиву EU Directive 2001/18/EC (называемую еще GMO Directive) и вопреки здравому смыслу пока не хочет ее пересматривать, даже после появления нокаутного геномного CRISPR/Cas редактирования, о котором говорилось выше. При этом верится, что рано или поздно ГМО станут настолько обыденными³, что на их генно-инженерный статус перестанут обращать внимание и тогда как раз водяные ДНК-знаки, маркирующие конкретные культуры с целью доказать авторство их создателей и получать полагающееся им роялти окажутся востребованными. Но то, что было сделано в этом плане в начале 2000-х гг. с одной стороны опоздало на десяток лет, а с другой опередило время, по крайней мере, наверное, на четверть столетия. Если судить по ниже приведенным датам, то полного привыкания к растительным ГМО

человечеству по нашим оценкам придется ждать еще около десятка лет. При этом хотелось бы ошибиться здесь в большую сторону, то есть ГМО все же должны перестать бояться раньше⁴. По крайней мере можно надеяться на это⁵. Причинами такого нашего благожелательного отношения к ГМО является понимание как их безопасности (ГМО ничуть не опаснее нынешних сельскохозяйственных растений-мутантов), невозможность на практике количественной детекции генетически-модифицированных ингредиентов в продуктах питания⁶, в том числе из-за недостоверности данных и чрезвычайной дороговизны анализов⁷, если подходить к этому очень серьезно, а не заниматься профанацией, и наконец, неоправданное сдерживание научных исследований в этом направлении, которые могли бы существенно повысить урожайность сельскохозяйственных культур, их послеуборочную сохранность и принести реальную пользу всему человечеству.

Так, в 2003 г. в одной из работ было предложено маркировать ГМО путем внедрения в геном данных организмов, помимо целевого трансгена, еще и сведений о фирме-создателе и некоторую другую техническую информацию [Marillonnet et al., 2003]. В качестве примера такой возможности этими авторами была создана трансгенная резуховидка Таля (*Arabidopsis thaliana*), в геноме которой был размещен следующий текст на латыни «NEC VERO TERRAE FERRE OMNES OMNIA POSSUNT» (Не может любая земля родить любое растение), для чего было применено триплетное кодирование латинского алфавита, арабских цифр, а также пробела, причем с учетом меньшего количества символов чем триплетов имела место дополнительная вырожденность такого кодирования, что присуще и для стандартного генетического (аминокислотного) кода. Эта публикация вызвала серьезную критику из целых семи

³ как и многочисленные сорта растений, созданных путем абсолютно бесконтрольного мутагенеза с совершенно непредсказуемыми последствиями в виде плейотропного действия генов, именно которого предлагают так бояться противники ГМО

⁴ хотя «мир» сейчас становится настолько непредсказуемым по многим направлениям, что делать какие-либо прогнозы - дело точно неблагоприятное

⁵ В другой нашей статье, посвященной столетию термина «геном» [Кулуев и др. (Kuluev et al.), 2020], мы позволили себе заглянуть на 30 лет вперед и выразили уверенность, что к этому времени аббревиатура «ГМО» просто исчезнет из употребления.

⁶ что требуется по нынешним нормативным документам

⁷ резко возрастающей при попытках улучшить достоверность анализов, требуемую точность которых в любом случае невозможно достичь

пунктов, за главный из которых, помимо увеличивающегося риска, возникающего вследствие еще неких последовательностей, дополнительно внедряемых в геном, можно считать, что фирмой-создателем ГМО и так предоставляется информация о месте интеграции трансгена, включая прилегающие последовательности нуклеотидов, что, по существу, служит «фирменной» меткой [Pauli, 2003]. В своем ответе на эту критику авторы предлагаемого маркирования нашли что ответить, заметив, что место внедрения трансгена, являющееся каждый раз действительно разным, может в будущем стать одинаковым и что такие подходы уже разрабатываются и тогда исчезнет уникальность [Marillonnet et al., 2003a]. Были даны ответы и на остальные пункты, включая упоминание о том, что системы получения генно-модифицированных организмов постоянно совершенствуются, причем это было сказано за десятилетие до появления CRISPR/Cas-технологии редактирования геномов.

Целую серию статей подготовила группа авторов [Heider, Barnekow, 2007; 2008; Heider et al., 2009], в которых описывается использование специальной программы DNA-crypt (http://heiderlab.de/?page_id=909), предназначенной для формирования водяных знаков в геномах разных организмов, в частности дрожжей *Saccharomyces cerevisiae*. К важной особенности DNA-crypt следует отнести благодаря коду Хэмминга возможность исправления ошибок при извлечении данных, при этом четыре азотистых основания кодировались следующими двоичными числами: А – 11; С – 10; G – 01; Т – 00. Также уделили серьезное внимание исправлению ошибок, вызываемых тандемными дупликациями при внедрении в геномы чужеродной ДНК для хранения небиологической информации другие авторы [Jain et al., 2017]. При этом они заметили, что ими игнорируются потенциальные биологические последствия нахождения такой ДНК в живых организмах. Ряд схожих с DNA-crypt алгоритмов для генерации водяных ДНК-знаков разработаны и другими авторами [Naughton, Balado, 2013; Lee, 2014; 2014a; Lee et al., 2018].

Возвращаясь к программе DNA-crypt нужно упомянуть еще одну публикацию [Heider et al., 2008], в которой авторы задумались о маркировании водяными ДНК-знаками организмов, размножающихся половым путем, где будет иметь место расщепление признаков, в том числе водяных ДНК-знаков, инициировав проект Project Mito, суть которого заключается в избирательном маркировании митохондриального генома, наследуемого по материнской линии. Авторы сообщили, что ими *in silico* найден участок в митохондриальной ДНК человека, пригодный для внедрения в него водяных

ДНК-знаков, но хочется верить, что до этого никогда человечество не дойдет. Также ими предлагается маркировать избирательно Y-хромосому, присущую мужчинам и наследуемую патрилинейно, что с одной стороны очень непросто сделать, а с другой - также бесчеловечно.

Целый ряд текстовых посланий длиной от 18 до 83 знаков были переведены с помощью укороченных шестибитных ASCII кодов для уплотнения информации в нуклеотидные последовательности и внедрены в несколько генов разного происхождения, среди которых большинство было представлено зеленым флуоресцентным белком GFP, которыми затем трансформировались *S.cerevisiae*, *Nicotiana benthamiana*, культуры клеток человека, а также прокариотический организм *E.coli* [Liss et al., 2012]. Для этого подбирались аминокислоты, кодируемые 4 или 6 триплетами, что давало большую свободу при использовании вырожденности генетического кода, не допуская изменения белковых последовательностей при формировании внутри выбранных генов водяных ДНК-знаков.

Другие авторы несколько английских букв (H, I, T, Z, J; A) и двоичных чисел (0, 1) закодировали восемью бинарными триплетами: H = 100; I = 001; T = 110; Z = 010; J = 000; A = 011; 0 = 111; 1 = 101, которые затем перевели в конкретные триплеты нуклеотидов [Yamamoto et al., 2014]. В итоге у них получился водяной ДНК-знак HITZJ001A0⁸, который они, снабдив «инициирующим и «терминирующим» триплетами в виде AAA и TTT, используя вырожденность «генетического» кода, представили следующей последовательностью нуклеотидов – AAA CAG AAT TCG GCG AAG CTT TCT TAT GCT CTT TTT. Данный водяной ДНК-знак был включен в состав одного из праймеров, рассчитанных на амплификацию гена пренилтрансферазы, который в составе вектора pCambia2301 был внедрен в геном *Nicotiana tabacum*. Обнаружение водяного знака HITZJ0010 велось путем амплификации с помощью ПЦР с соответствующими праймерами участка генома табака, который затем подвергался секвенированию и дополнительному дот-матрикс анализу. Причем было отмечено, что подобный водяной ДНК-знак в виде последовательности ДНК может быть представлен на разных языках, перечислив при этом греческий, арабский, японский.

⁸ HITZ – владелец водяного ДНК-знака, в данном случае Hitz Research Alliance Laboratory; J001 – обозначение вида, сорта, линии и т.д. трансгенного растения; A0 – обозначение генно-инженерной конструкции.

В одной из работ [Ljubic, Fister, 2014] была предположена возможность хранения англоязычной Wikipedia объемом на тот момент около 42 гигабайт в каком-либо дереве, растущем в парке, и чтобы получить из него нужную информацию каждый человек должен будет иметь портативный ДНК секвенатор. Все же считаем нужным заметить, что в одном дереве такой объем не уместить, поскольку самые крупные геномы из известных у растений составляют величины приблизительно того же порядка (например, у *Paris quadrifolia* – почти 60

млрд.п.н.) и второй такой «набор хромосом» никакому растению просто не выдержать.

Для маркирования ГМО водяными ДНК-знаками был предложен особый алгоритм, основанный на триплетах из двоичных чисел и соответствующих им нуклеотидных триплетах, приведенный в таблице 4, где p_1, p_2, p_3 и p_4 соответствуют нуклеотидам А, С, G и Т соответственно [Mueller et al., 2016]. При этом число комбинаций триплетов из двоичных чисел равно $8 (2^3)$, тогда как для четверки нуклеотидов оно значительно больше и составляет $64 (4^3)$.

Таблица 4

Соответствие бинарных триплетов нуклеотидным триплетам
Table 4 - Correspondence of binary triplets to nucleotide triplets

binary text triplets, B_i	codons $\{n_1n_1n_1, n_1n_1n_2, n_1n_2n_1 \dots n_4n_4n_4\}, C_i$
$B_1 = 000$	AAA, AAC, ACA, ACC, CAA, ACA, CCA, CCC
$B_2 = 001$	AAG, AAT, ACG, ACT, CAG, CAT, CCG, CCT
$B_3 = 010$	AGA, AGC, ATA, ATC, CGA, CGC, CTA, CTC
$B_4 = 100$	GAA, GAC, GCA, GCC, TAA, TAC, TCA, TCC
$B_5 = 011$	AGG, AGT, ATG, ATT, CGG, CGT, CTG, CTT
$B_6 = 101$	GAG, GAT, GCG, GCT, TAG, TAT, TCG, TTC
$B_7 = 110$	GGA, GGC, GTA, GTC, TGA, TGC, TTA, TTC
$B_8 = 111$	GGG, GGT, GTG, GTT, TGG, TGT, TTG, TTT

Недавно было продемонстрировано внесение водяных ДНК-знаков в геном дрожжей *S.cerevisiae* сразу в 13 генов, кодирующих белки, участвующие в основном пути утилизации сахара (т.е. гликолизе и алкогольной ферментации) [Boonekamp et al., 2020]. При этом сравнивались подходы кластерного и распределенного расположения водяных ДНК-знаков, причем последнее показало определенные преимущества. Кроме того, наличие водяных знаков позволило выборочно редактировать геном дрожжей с помощью CRISPR/Cas технологии, специально нацеливаясь на нативную копию гена, оставляя синтетический вариант с водяными знаками нетронутым.

Путем агробактериальной трансформации было предложено вводить в геномы созданных сортов растений нейтральные фрагменты ДНК со специально подобранной последовательностью, которые с помощью ПЦР можно выявить и секвенировать, расшифровав некий текст, позволяющий доказать права на конкретную интеллектуальную собственность [Fister et al., 2017; 2017a]. Для этого авторы разработали специальную программу (<http://www.storing-data-into-living-plant.net>), переводящую короткий текст (всего допускается 300 символов), написанный с использованием английского алфавита, включая арабские цифры, а также пару знаков # и ' в последовательность ДНК (1200 нуклеотидов), в том числе в сжатом виде по алгоритму Хаффмана, и наоборот -

последовательность ДНК в соответствующий закодированный в ней текст. Так, ими была закодирована фраза 'Hello world' и после ее перевода в нуклеотидную последовательность было получено трансгенное растение *N.benthamiana*, несущее данное послание. При этом, похоже, авторы забыли указать, что их программа также кодирует пробелы между словами, и это обнаружено нами при преобразовании для пробы с помощью их on-line конвертера нашего текста 'Hello 2022' в последовательность ДНК и обратно, что видно из представленных ниже последовательностей нуклеотидов –

ACGT GACA GCGG GCTA GCTA GCTT ACAA ATAC ATAA ATAC ATAC ACGT
' H E L L O 2 0 2 2 ' ,

где ACAA как раз указывают на пробел между Hello и 2022, показанный здесь точками. Фактически при таком маркировании селекционных достижений предлагается получать настоящее трансгенное растение или ГМО, пометив чужеродной ДНК, в том числе даже сорт обычной селекции.

Для организации хранения в рекомбинантных митотрофных дрожжах *Pichia pastoris* и в трансгенном арабидопсисе поэмы Мао Цзэдуна «Снег» был разработан алгоритм и созданы две версии программы "DNA Studio", переводящей текстовую информацию в последовательность ДНК [Sun et al., 2019]. В итоге 153 иероглифа этой поэмы с помощью первой версии были закодированы в 2448 нуклеотидах, тогда как вторая оказалась более экономной и дала возможность уложить их в 1232

нуклеотида. Имеется on-line версия DNA Studio, генерирующая из английских текстов последовательности азотистых оснований (<http://www.synbio-tech.com/dna-studio/>). Мы не преминули этим воспользоваться и преобразовали тот же текст Hello 2022 (но без апострофов, поскольку они не допускаются) в последовательность нуклеотидов –

CCTT TGCA TCTC TCTC TTAG ATCG ATAC AAGT ATAC ATAC

Н Е Л Л О 2 0 2 2 ,

где как можно видеть используются иные тетраплеты, в том числе и для пробела между словами, также обозначенного нами точками. К сожалению, на указанном сайте не предусмотрена обратная конвертация.

Целая серия публикаций посвящена проблеме ГМО и биотерроризма на их основе для чего предлагается внедрять в их геномы специфичные последовательности ДНК, называемые авторами сигнатурами [Mueller et al., 2019; 2019a], что является синонимом водяным ДНК-знакам. При этом отмечается, что такое маркирование ГМ-культур и обнаружение этих сигнатур с помощью ПЦР имеет своей целью также защитить производителей от подделок, которые могут предлагаться недобросовестными участниками рынка. Здесь можно еще добавить, что в критичных случаях, связанных с реальными угрозами благополучию людей, ПЦР и обычным секвенированием уже не обойтись, благо современные технологии позволяют секвенировать полные геномы растений, включая неизвестные ГМО, что уже описано в ряде статей [Arulandhu et al., 2016; Holst-Hensen et al., 2016].

В связи с предлагаемой интеграцией в геном растений чужеродной человеческой ДНК нужно упомянуть английскую фирму Biopresence, организованную двумя выпускниками Королевского колледжа искусств и ставящую целью вместо установки привычных надгробных памятников создание трансгенных яблонь, несущих фрагменты ДНК умерших людей в качестве некой памяти о них. Британская пресса окрестила их как «трансгенные надгробия». В настоящее время сайт этой фирмы <https://www.biopresence.com> полноценно не функционирует, хотя приведены следующие слова – “Human DNA Trees as Living Memorials” и говорится, что готовится его повторный запуск. Но еще при организации этой фирмы в 2004 г. возникло множество проблем, как природоохранных, поскольку создаваемые яблони являлись бы ничем иным как трансгенными растениями или ставшими теми же ГМО, будучи выпущенными в окружающую среду, так и этических моментов, задаваясь вопросом – А готовы ли вы кушать яблоко, где пусть и чисто символически, но присутствует, например, ДНК вашей бабушки? При этом сообщалось, что

изначально предполагалось в некодирующих участках генома яблонь размещать также некодирующие фрагменты ДНК людей. Но после обсуждения в мае 2003 г. этой задачи с известным биохудожником J.Davis⁹ он предложил вносить изменения в кодируемые триплеты в виде молчащих мутаций, создавая transhuman plant genes в виде неких manifolds. Но считать ли их водяными ДНК-знаками, устанавливающими некую связь живущего дерева и ушедшего в мир иной человека, каждый должен решать для себя сам. Возможно, этот проект так и не перейдет в практическую плоскость, оставшись лишь проектом. Остается только добавить, что стоимость создания такой трансгенной яблони с символической человеческой ДНК была оценена в 20 тысяч фунтов стерлингов. Вероятно, это и была самая главная цель проекта. При этом мы не встретили упоминания - дерево какого размера должно предоставляться родственникам, и какое время должно пройти от момента заказа и его оплаты. Похоже треба было просто (ж)дать.

Водяные ДНК-знаки без их встраивания в геномы каких-либо организмов

Помимо живых организмов, с помощью водяных ДНК-знаков можно метить и другие объекты. Например, молекулы ДНК можно использовать в прямом и переносном смысле в виде водяных ДНК-знаков при слежении за потоками воды в гидрогеологии, что нами рассмотрено довольно подробно отдельно [Гималов и др. (Gimalov et al.), 2021], но здесь также уделим этому вопросу некоторое внимание. Так, бывает необходимо установить движение подземных вод для чего используются различные метчики, включая флуоресцентные красители, растворы солей, радиоактивные элементы. Но это оказывается или довольно дорого, или небезопасно для окружающей среды и самого человека. Что касается ДНК, будь это синтезированные олигонуклеотиды или плазмидная ДНК, либо инкапсулированные в кремнеземе молекулы ДНК, то вносимое количество этого биополимера с одной стороны не может нанести никакого вреда экологии, с другой – вполне доступно по финансовым затратам, включая проведение процедур по амплификации и

⁹ Joe Davis еще в 1988 г. первым преобразовал графическое изображение Microvenus в рамках одноименного проекта в бинарный код, а затем и в нуклеотидную последовательность, которая была клонирована в *E.coli* [Davis, 1996], но считать ту последовательность из 35 нуклеотидов водяным ДНК-знаком нельзя, поскольку его такая кодировка предназначалась для контактов с внеземными цивилизациями вкупе с радиосигналами.

детекции специфических нуклеотидных последовательностей с помощью ПЦР. Так, в частности, введя в две скважины глубиной около 10 метров каждая инкапсулированные молекулы ДНК и отбирая затем с разных глубин пробы воды на площадке в 100 кв.м, удалось провести фактически томографию водотоков этого участка [Kong et al., 2018]. Предложено метить наночастицами кремнезема с инкапсулированными в них молекулами ДНК даже молоко для того, чтобы отслеживать его перемещение, включая производные от него продукты в виде йогурта, сыра и пр. [Bloch et al., 2014]. При этом подсчитана стоимость такого мечения при использовании 10 мкг ДНК-наночастиц на 1 литр молока, составляющая менее 10 центов на тонну. В своих дальнейших работах авторы расширили спектр применения таких наночастиц кремнезема с молекулами ДНК внутри в качестве метчиков и использовали их на слежение за распространением в окружающей среде пестицидов, маркированных выявляемыми с помощью ПЦР молекулами ДНК [Mora et al., 2015a], для выявления трофических связей, метя такой ДНК определенные цветки, опыляемые мухами [Mora et al., 2015], потоками воздуха счет аэрозолированной формы таких кремнеземных частиц с ДНК [Luescher et al., 2021], а также для слежения за путями распространения внутрибольничных инфекций [Scotoni et al., 2020; Ullrich et al., 2022].

Молекулами ДНК можно метить и иные субстанции, нефть в частности, что позволит легко установить в случае необходимости ее происхождение путем детекции тем или иным методом высокочувствительной амплификации специфических фрагментов ДНК добавленных конкретных¹⁰ молекул ДНК в нефть на определенных стадиях ее добычи и транспортировки с помощью соответствующих праймеров (которые в данном случае могут быть названы шифровальными ключами и о коих должно быть известно только производителю и возможно также конечному получателю). Причем вполне достаточно небольшого количества подходящей ДНК, которая может быть синтезирована химическим путем или наработана в рекомбинантных бактериях. Так, например, для такой детекции достаточно нахождения в 1 мл нефти всего 10^3 копий молекул соответствующей ДНК или олигонуклеотидов. В миллиарде тонн нефти таких миллилитров всего 10^{15} . Таким образом, на несколько миллиардов тонн нефти, добываемой в мире ежегодно, потребуется приблизительно 10^{18} молекул ДНК. Такое количество может быть синтезировано за одну загрузку ДНК-синтезатора, обеспечивающего масштаб синтеза в 500

микромоль. Но поскольку разную нефть нужно метить разными последовательностями ДНК, то и обычно используемые масштабы синтеза олигонуклеотидов в 20 микромоль будут вполне способны обеспечить такое мечение нефти (перед транспортировкой после ее обезвоживания и обессоливания) из различных месторождений и даже из отдельных скважин, если это будет необходимо и позволено технологическими возможностями. Какого-либо изменения качества нефти при внесении молекул ДНК произойти не может ввиду их крайне малого количества, к тому же остаточное количество воды в нефти, присутствующей в трех формах – свободной, растворенной и диспергированной, в любом случае имеется, и варьирует в зависимости от групп нефти от 0,5 до 1%. При смешении нефтей в трубопроводе где-нибудь на выходе (или в пробах из танкеров) можно будет с помощью ПЦР детектировать присутствие той или иной нефти (то есть выяснять ее происхождение), в том числе количественно в режиме реального времени, используя конкретные праймеры. По грубым прикидкам вся ежегодно добываемая в мире нефть может быть помечена олигонуклеотидами, стоимость которых не превысит миллиона рублей и это по сравнению со стоимостью самой нефти составляет ничтожную сумму. Здесь можем лишь добавить, что некоторое время назад нами проводился многолетний эксперимент по оценке сохранения в нефти добавленных в нее молекул ДНК в виде рекомбинантной плазмиды с клонированным геном лектина гороха, которые после их экстракции из нефти, легко детектировались с помощью ПЦР, показавший применимость и воспроизводимость такого подхода. Возможно, стоит такие водяные ДНК-знаки инкапсулировать в наночастицы кремнезема для их лучшей сохранности, что несколько удорожит стоимость маркирования и детекции, но на качество нефти также не сможет оказать никакого влияния.

Можно метить молекулами ДНК и жидкости небольшого объема, например, чернила, как обычные, так и водостойкие, что позволяет с помощью ПЦР со специфическими праймерами провести оценку подлинности подписанных важных документов [Hashiyada, 2004; Oshima, Yusa, 2015]. Показано, что для этих целей ДНК может быть синтетической, бактериальной, растительной, животной, включая некую производную ДНК от ДНК конкретных людей, делая в последнем случае такие чернила по существу именными.

Нельзя исключать использование водяных ДНК-знаков для маркирования и природного газа, например транспортируемого по трубам путем внесения, скорее всего, не инкапсулированных, а чистых молекул ДНК в виде аэрозолей, что не должно составлять серьезной технической проблемы. Однако

¹⁰ фактически каждое месторождение может иметь свою ДНК-метку

для обнаружения ДНК в газе потребуется уже использовать особые устройства, что может оказаться непростым делом. В случае с газом все же вероятно легче собирать наночастицы с инкапсулированной в них ДНК, на самом деле имеющих субмикрометровые размеры, благодаря чему они могут задерживаться на фильтрах, через которые следует прокачивать анализируемый газ. Но поведение довольно тяжелых кремнеземных частиц в газовых потоках может быть не приемлемым для подобной диагностики. Что касается сжиженного газа, то его подобное маркирование теоретически вполне может быть достижимо при условии предварительного мечения метана до его сжижения чистой ДНК.

В одной из работ сообщается о нанесении ДНК-метки на обычную писчую бумагу, отправленную затем по почте и по получению письма из известного участка бумажного листа, куда до этого было нанесено 0,5 мкл водно-спиртового раствора ДНК в виде смеси из 10 матриц, имеющих различное происхождение и размеры от 267 до 804 п.н., проводили экстракцию этого биополимера и детектировали его наличие с помощью ПЦР со специфичными праймерами для подтверждения его подлинности [Cox, 2001]. В следующей аналогичной работе [Cook, Cox, 2003] с нанесением ДНК-меток на бумагу и их последующим извлечением было принято решение увеличить полиморфизм молекул ДНК за счет использования двух метилаз вкупе с соответствующими рестрикционными эндонуклеазами, что действительно произошло, но излишние сложности всех этих процедур сводят на нет эффект от увеличения числа полиморфных фрагментов ДНК.

В Соединенном Королевстве действует фирма SelectaDNA (<https://www.selectadna.co.uk>), специализирующая на мечении с помощью уникальных молекул ДНК различных дорогостоящих предметов путем нанесения некоего клея с ДНК, который практически невозможно удалить полностью, и это вместе с предупреждающей наклейкой предотвращает воровство, поскольку такие вещи становится гораздо труднее сбыть. Использованные для мечения последовательности ДНК вместе с информацией о клиентах хранятся в защищенной базе данных, что позволяет установить законного владельца после анализа даже крошечной частички такого клея. В США фирма DNA Technologies (<https://dnatechnologies.com>) предоставляет сходные услуги по мечению с помощью ДНК различных ценных предметов, включая произведения искусства.

Выше мы упоминали о готовности или неготовности людей кушать трансгенные яблоки с ДНК их бабушки и можно было бы считать, что о том уже забыли, но получается, что здесь опять необходимо вернуться к яблонам. К оригинальному способу защиты ГМО прибегла английская художница S.Jarvis (<https://cjarvis.com/blighted-by-kenning/>), инициировавшая в феврале 2012 г. кодирование в ДНК первой статьи Всеобщей декларации прав человека и для этого ей тоже понадобились яблоки, как одни из самых потребляемых фруктов в мире, к тому же несущие различную смысловую нагрузку в виде библейского запретного плода, но при этом также символизирующего и человеческие знания, имея в виду открытие силы всемирного тяготения И.Ньютоном.

```

GCT CGT ACT ATT TGT TTA GAA AGA ATA AAT GAA AGA GCT TTA TTA AGA CAT GGG ATG GCT AAT AGA
A R T I C L E O N E A L L H U M A N
TTG GAA ATT AAT GGT TCT AGA GCT CGT GAA AGA TTG ATA CGT AAT AGA TTT CGT GAA GAA AGA
B E I N G S A R E B O R N F R E E
GCT AAT GAT AGA GAA CAA GGG GCT TTA AGA ATT AAT AGA GAT ATT GGT AAT ATT ACT TAT AGA
A N D E Q U A L I N D I G N I T Y
GCT AAT GAT AGA CGT ATT GGT CAT ACT TCT AGA ACT CAT GAA TAT AGA GCT CGT GAA AGA
A N D R I G H T S T H E Y A R E
GAA AAT GAT ATA TGG GAA GAT AGA TGG ATT ACT CAT AGA CGT GAA GCT TCT ATA AAT AGA GCT AAT GAT
E N D O W E D W I T H R E A S O N A N D
AGA TGT ATA AAT TCT TGT ATT GAA AAT TGT GAA AGA GCT AAT GAT AGA TCT CAT ATA GGG TTA GAT AGA
C O N S C I E N C E A N D S H O U L D
GCT TGT ACT AGA ACT ATA TGG GCT CGT GAT TCT AGA ATA AAT GAA AGA GCT AAT ATA ACT CAT GAA CGT
A C T T O W A R D S O N E A N O T H E R
AGA ATT AAT AGA GCT AGA TCT CCT ATT CGT ATT ACT AGA ATA TTT AGA
I N A S P I R I T O F
TTG CGT ATA ACT CAT GAA CGT CAT ATA ATA GAT
B R O T H E R H O O D

```

Рис. 1. Нуклеотидная последовательность, в которой записана первая статья Декларации прав человека (синий шрифт) и ее отображение в «аминокислотном» тексте (красный шрифт).

Кодон AGA был выбран для кодирования пробела между словами.

Fig. 1. The nucleotide sequence in which the first article of the Declaration of Human Rights is written (blue font) and its display in the "amino acid" text (red font). The AGA codon was chosen to encode a space between words.

Ранее мы уже цитировали работу [Gustafsson, 2009], в которой поэма Tomtem через аминокислоты, используя их однобуквенное кодирование, была закодирована последовательностью нуклеотидов. Схожий подход был использован при записи в виде ДНК первой статьи Декларации прав человека, по инициативе Jarvis выполненной Netherland Proteomics Centre - **“All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood”**. Как известно в английском алфавите 26 букв, но в этой фразе их задействовано всего 21 и при этом ни разу не встречаются буквы J, X, Z, K и V, причем две последние соответствуют аминокислотам лизину и валину и как бы в перспективе заняты. Поэтому для записи первой статьи Декларации нуклеотидными последовательностями через однобуквенный аминокислотный код законных триплетов для всех букв в тексте не хватило и пришлось три буквы B, O и U (аминокислот, обозначаемых которыми, не имеется) кодировать подошедшими триплетами, что не должно удивлять тех, кто решит проверить на соответствие настоящему генетическому коду нуклеотидную последовательность, приготовленную для клонирования и приведенную выше.

В рамках этого проекта трансгенных яблоны получать не стали и ограничились созданием рекомбинантных бактерий *E.coli*, несущих фрагмент ДНК, в котором была закодирована приведенная выше первая статья Всеобщей декларации прав человека. Яблоки, выращенные недалеко от Гааги (где находится Международный суд по правам человека), были сознательно загрязнены (опрысканы) раствором этой синтетической ДНК, выделенной из бактерий, превратив обработанные таким образом яблоки в некое подобие ГМО. Затем эти как бы дважды с учетом наличия чужеродной ДНК, отождествляющей их с ГМО, запретные плоды были отправлены в лаборатории геномики по всему миру с предложением просеквенировать находящуюся на поверхности яблок ДНК и также съесть присланный им плод. По крайней мере, некоторые лаборатории выполнили эти установки. Сама Jarvis также не побрезговала таким яблоком, и на представленной серии фотографий на ее вышеупомянутом персональном сайте видно, как она такое запретное, загрязненное чужеродной ДНК, квазиГМО яблоко с удовольствием поедает.

Заключение

Отдельные буквы¹¹ в любом алфавите ничего не значат, но складываемые в слова, фразы, предложения, абзацы и целые литературные

произведения (включая данную статью) обретают разнообразный смысл, которое им придают автор(ы), но определяемый в первую очередь последовательностями этих самых букв. То же самое имеет место и для генетического алфавита - четверки нуклеотидов, которые, перемежаясь по-разному, обеспечивают гигантское разнообразие всего Живого. ДНК - это созданный самой Природой уникальнейший биополимер, являющейся к тому же молекулой наследственности, для чего хватило алфавита из всего четырех букв-нуклеотидов.

В ДНК заложен целый ряд особенностей, благодаря которым ей отведена такая важная миссия на нашей Планете. Во-первых, это несусветное количество теоретически возможных комбинаций нуклеотидных последовательностей, определяемое как число размещений с повторениями из 4 элементов и подчиняющееся формуле 4^n , где 4 – четыре разных азотистых основания, а n – длина некоего участка ДНК. Так, для короткого декануклеотида таковых будет за миллион (4^{10}), для олигонуклеотида длиной 15 звеньев – миллиард (биллион)¹², для 20 звеньев – триллион, для олигонуклеотида из 30 звеньев – квинтиллион (10^{18}), для участков ДНК из 100 азотистых оснований количества комбинаций превысит нонадециллион (10^{60}). Гугол¹³ комбинаций (10^{100}) будет превышен, если иметь дело с ДНК длиной всего 167 нуклеотидов. И это притом, что считается, что вся наша Вселенная состоит из всего 10^{80} элементарных частиц. А небольшие геномы относительно свободноживущих организмов состоят из приблизительно миллиона пар нуклеотидов. Вот и считайте - сколько может быть в них комбинаций перестановок азотистых оснований. Выше мы упомянули одно растение, у которого размер генома приблизительно оценен в 60 млрд. п.н., хотя, скорее

¹² Здесь мы привели наименование больших чисел одновременно по двум системам, действующим в разных странах. По так называемой длинной шкале после миллиона идет миллиард, затем чередуя «...лионы» и «...лиарды», следуют биллион, биллиард и т.д. По короткой шкале за миллионом следует биллион, затем триллион, квадриллион, квинтиллион и так далее. Но на самом деле нечасто возникает необходимость оперировать столь большими числами и в быту уже произошло смешение этих систем, так как в России, например, после миллиарда идет триллион, тогда как должен был быть сначала биллион, потом биллиард и только затем триллион.

¹³ Число «гугол» (googol), равное 10^{100} не нужно путать со сходным по произношению названием поисковика Google, что представляет собой «игру слов» и подразумевает претензию последнего на огромную информативность.

¹¹ за исключением аббревиатур и местоимения «я»

всего, около 99,99% его генома это повторяющиеся последовательности, снижающие число возможных комбинаций перестановок четверки нуклеотидов. Но в любом случае ДНК это биополимер с поистине безграничными возможностями.

Другой важной особенностью ДНК является двухцепочечная природа этой молекулы, причем одна цепь при энзиматическом копировании при репликации определяет другую за счет принципа комплементарности азотистых оснований. Но, конечно, ДНК это только текст генетического послания потомкам и требуются еще молекулы РНК, переводящие ДНКовый текст в белки, выполняющие уже строительные и прочие функции в организме, обеспечивая его функционирование¹⁴. Наконец, то, что копирование цепей ДНК, осуществляемое при размножении в живых организмах в системе *in vivo*, можно воспроизводить с помощью специальных выделенных и очищенных ферментов ДНК полимераз в системе *in vitro*, дает возможность экспериментаторам использовать этот биополимер для разнообразнейших целей.

Когда в 1968 г. Х.Г.Коране вместе с Р.Холли и М.Ниренбергом была присуждена Нобелевская премия по физиологии и медицине «за расшифровку генетического кода и его роли в синтезе белков», то в поздравительной речи представитель Каролинского института П.Рейхард сравнил нуклеиновые кислоты и белки с языками, а их составные элементы – с буквами алфавита. Он отметил: «Химическая структура нуклеиновых кислот определяет химическую структуру белка, а алфавит нуклеиновых кислот – алфавит белков. Генетический код – это словарь, благодаря которому возможен переход с одного алфавита на другой». С позиций дня сегодняшнего можно считать, что это скорее не словарь, а специальная, обеспечивающая транслитерацию программа-переводчик, причем нашедшая применение уже не только при обращении с генетической информацией, но и при кодировании всевозможных текстовых посланий. Но ДНК и сама по себе, являясь, по сути, цифровой молекулой способна кодировать своими нуклеотидами напрямую любую информацию, что уже в ходе изложения материала данной статьи неоднократно показано. Поэтому можно не сомневаться в том, что благодаря уникальным особенностям ДНК ее применение, как биополимера, для небиологических целей, включая маркирование водяными ДНК-знаками всего чего

только можно и нужно, будет непременно расширяться.

Благодарности

Работа выполнена при финансовой поддержке гранта РФФИ № 20-07-00222.

Литература

1. Баймиев Ан.Х., Кулуев Б.Р., Вершинина З.Р., Князев А.В., Чемерис Д.А., Рожнова Н.А., Геращенко Г.А., Михайлова Е.В., Баймиев Ал.Х., Чемерис А.В. CRISPR/Cas редактирование геномов (растений) и общество // Биомика. 2017. Т. 9. С. 183-202.
2. Вентер К. Жизнь на скорости света. От двойной спирали к рождению цифровой биологии М.: АСТ. 2018. 320 с.
3. Вершинина З.Р., Кулуев Б.Р., Максимов И.В., Михайлова Е.В., Гумерова Г.Р., Малеев Г.В., Князев А.В., Баймиев Ан.Х., Баймиев Ал.Х., Чемерис А.В. ГМО запретить невозможно разрешить! // Биомика. 2020. Т.12(1). С. 80-120. doi: 10.31301/2221-6197.bmcs.2020-6.
4. Гарафутдинов Р.Р., Сахабутдинова А.Р., Чемерис А.В. Долговременное хранение молекул ДНК при комнатной температуре // Биомика. 2020. Т.12(4). С. 552-563. DOI: 10.31301/2221-6197.bmcs.2020-49
5. Гималов Ф.Р., Гарафутдинов Р.Р., Кулуев Б.Р., Привалов Л.Ю., Кирьянова О.Ю., Михайленко К.И., Урманчеев С.Ф., Чемерис А.В. Трассирование в гидрогеологии подземных водотоков с помощью ДНК-метчиков // *Biomics*. 2021. Т.13(3). С.244-253. DOI: 10.31301/2221-6197.bmcs.2021-16
6. Сахабутдинова А.Р., Михайленко К.И., Гарафутдинов Р.Р., Кирьянова О.Ю., Сагитова М.А., Сагитов А.М., Чемерис А.В. Небиологическое применение молекул ДНК // Биомика. 2019. Т.11(3). С. 344-377. DOI: 10.31301/2221-6197.bmcs.2019-28
7. Чемерис А.В., Бикбулатова С.М., Чемерис Д.А., Баймиев Ал.Х., Князев А.В., Кулуев Б.Р., Максимов И.В. Надо ли опасаться ГМО? Взгляд несторонних наблюдателей на истерию вокруг // Биомика. 2014. Т. 6. С. 77-138.
8. Чемерис А.В., Чемерис Д.А., Баймиев Ал.Х., Князев А.В., Кулуев Б.Р., Максимов И.В. Борьба с ГМО как неолысенковщина // Биомика. 2015. Т.7. С. 1-39.
9. Ailenberg M., Rotstein O. An improved Huffman coding method for archiving text, images, and music characters in DNA // *Biotechniques*. 2009. V. 47(3). P. 747-754. doi: 10.2144/000113218.
10. Arulandhu A.J., van Dijk J.P., Dobnik D., Holst-Jensen A., Shi J., Zel J., Kok E.J. DNA enrichment

¹⁴ Помимо этих веществ первичного происхождения для обеспечения жизнедеятельности организма нужно еще и множество веществ вторичного происхождения, но они выходят за рамки данной статьи.

- approaches to identify unauthorized genetically modified organisms (GMOs) // *Anal. Bioanal. Chem.* 2016. V. 408(17). P. 4575-93. doi: 10.1007/s00216-016-9513-0.
11. Bancroft C., Bowler T., Bloom B., Clelland C.T. Long-term storage of information in DNA // *Science*. 2001. V. 293(5536). P.1763-1765. doi: 10.1126/science.293.5536.1763c.
 12. Bloch M.S., Paunescu D., Stoessel P.R., Mora C.A., Stark W.J., Grass R.N. Labeling milk along its production chain with DNA encapsulated in silica // *J. Agric. Food Chem.* 2014. V. 62(43). P. 10615-10620. doi: 10.1021/jf503413f.
 13. Boonekamp F.J., Dashko S., Duiker D., Gehrmann T., van den Broek M., den Ridder M., Pabst M., Robert V., Abeel T., Postma E.D., Daran J.M., Daran-Lapujade P. Design and Experimental Evaluation of a Minimal, Innocuous Watermarking Strategy to Distinguish Near-Identical DNA and RNA Sequences // *ACS Synth. Biol.* 2020. V. 9(6). P. 1361-1375. doi: 10.1021/acssynbio.0c00045.
 14. Brenner S., Williams S.R., Vermaas E.H., Storck T., Moon K., McCollum C., Mao J.I., Luo S., Kirchner J.J., Eletr S., DuBridge R.B., Burcham T., Albrecht G. In vitro cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs // *Proc. Natl. Acad. Sci. USA*. 2000. V. 97(4). P. 1665-1670. doi: 10.1073/pnas.97.4.1665.
 15. Clelland C.T., Risca V., Bancroft C. Hiding messages in DNA microdots // *Nature*. 1999. V. 399(6736). P. 533-534. doi: 10.1038/21092.
 16. Cook L.J., Cox J.P. Methylated DNA labels for marking objects // *Biotechnol. Lett.* 2003. V. 25(1). P. 89-94. doi: 10.1023/a:1021713625176.
 17. Cox J.P. Bar coding objects with DNA // *Analyst*. 2001. V. 126(5). P. 545-547. doi: 10.1039/b101850g.
 18. Cox J.P. Long-term data storage in DNA // *Trends Biotechnol.* 2001a. V. 19(7). P. 247-250. doi: 10.1016/s0167-7799(01)01671-7.
 19. Feynman R.P. There's Plenty of Room at the Bottom: An Invitation to Enter a New Field of Physics // *Engineering and Science (California Institute of Technology)*. 1960. V. 23. P. 22-36.
 20. Fister K., Fister I. Jr., Murovec J., Bohanec B. DNA labelling of varieties covered by patent protection: a new solution for managing intellectual property rights in the seed industry // *Transgenic Res.* 2017. V. 26. P. 87-95.
 21. Fister K., Iztok Fister J.R. How to Store Wikipedia into a Forest Tree: Initial Idea // *MSIVISM*. 2014. P. 45-52.
 22. Fister K., Iztok Fister J.R., Murovec J. The Potential of Plants and Seeds in DNA-Based Information Storage // *Understanding Information*. 2017. P. 69-81. doi: 10.1007/978-3-319-59090-5_4.
 23. Gibson D.G., Benders G.A., Andrews-Pfannkoch C., Denisova E.A., Baden-Tillson H., Zaveri J., Stockwell T.B., Brownley A., Thomas D.W., Algire M.A., Merryman C., Young L., Noskov V.N., Glass J.I., Venter J.C., Hutchison C.A. 3rd, Smith H.O. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome // *Science*. 2008. V. 319(5867). P. 1215-20. doi: 10.1126/science.1151721.
 24. Gibson D.G., Glass J.I., Lartigue C., Noskov V.N., Chuang R.Y., Algire M.A., Benders G.A., Montague M.G., Ma L., Moodie M.M., Merryman C., Vashee S., Krishnakumar R., Assad-Garcia N., Andrews-Pfannkoch C., Denisova E.A., Young L., Qi Z.Q., Segall-Shapiro T.H., Calvey C.H., Parmar P.P., Hutchison C.A. 3rd, Smith H.O., Venter J.C. Creation of a bacterial cell controlled by a chemically synthesized genome // *Science*. 2010. V. 329(5987). P. 52-6. doi: 10.1126/science.1190719.
 25. Gustafsson C. For anyone who ever said there's no such thing as poetic gene // *Nature*. 2009. V. 458. P. 703. doi: 10.1038/458703a.
 26. Hao M., Qiao H., Gao Y., Wang Z., Qiao X., Chen X., Qi H. A mixed culture of bacterial cells enables an economic DNA storage on a large scale // *Commun. Biol.* 2020. V. 3(1). P. 416. doi: 10.1038/s42003-020-01141-7.
 27. Hashiyada M. Development of biometric DNA ink for authentication security // *Tohoku J. Exp. Med.* 2004. V. 204. P. 109-117.
 28. Haughton D., Balado F. BioCode: two biologically compatible Algorithms for embedding data in non-coding and coding regions of DNA // *BMC Bioinformatics*. 2013. V. 14. P. 121. doi: 10.1186/1471-2105-14-121.
 29. Heider D., Barnekow A. DNA Watermarking: Challenging Perspectives for Biotechnological Applications // *Curr. Bionf.* 2011. V. 6. P. 375-382. doi: 10.2174/157489311796904646.
 30. Heider D., Barnekow A. DNA watermarks: a proof of concept // *BMC Mol. Biol.* 2008. V. 9. P. 40. doi: 10.1186/1471-2199-9-40.
 31. Heider D., Barnekow A. DNA-based watermarks using the DNA-Crypt algorithm // *BMC Bioinformatics*. 2007. V. 8. P. 176. doi: 10.1186/1471-2105-8-176.
 32. Heider D., Kessler D., Barnekow A. Watermarking sexually reproducing diploid organisms // *Bioinformatics*. 2008. V. 24(17). P. 1961-1962. doi: 10.1093/bioinformatics/btn342.
 33. Heider D., Pyka M., Barnekow A. DNA watermarks in non-coding regulatory sequences // *BMC Res.*

- Notes. 2009. V. 2. P. 125. doi: 10.1186/1756-0500-2-125.
34. Holst-Jensen A., Spilberg B., Arulandhu A.J., Kok E., Shi J., Zel J. Application of whole genome shotgun sequencing for detection and characterization of genetically modified organisms and derived products // *Anal. Bioanal. Chem.* 2016. V. 408(17). P. 595-614. doi: 10.1007/s00216-016-9549-1.
35. Jain S., Farnoud F., Schwartz M., Bruck J. Duplication-Correcting Codes for Data Storage in the DNA of Living Organisms // *IEEE Trans. Inform. Theory.* 2017. V. 63(8). P. 4996-5010. doi: 10.1109/TIT.2017.2688361.
36. Jiao S., Goutte R. Code for encryption hiding data into genomic DNA of living organisms // 9th International Conference on Signal Processing. 2008. doi: 10.1109/ICOSP.2008.4697576.
37. Jupiter D.C., Ficht T.A., Samuel J., Qin Q.M., de Figueiredo P. DNA watermarking of infectious agents: progress and prospects // *PLoS Pathog.* 2010. V. 6(6). e1000950. doi: 10.1371/journal.ppat.1000950.
38. Kac E. Genesis: a transgenic artwork // *Art, Technology, Consciousness mind@large*. R. Ascott (ed.). 2000. P. 17-19.
39. Kong X.Z., Deuber C.A., Kittilä A., Somogyvári M., Mikutis G., Bayer P., Stark W.J., Saar M.O. Tomographic Reservoir Imaging with DNA-Labeled Silica Nanotracers: The First Field Validation // *Environ. Sci. Technol.* 2018. V. 52(23). P. 13681-13689. doi: 10.1021/acs.est.8b04367.
40. Lee S-H. DNA sequence watermarking based on random circular angle // *Digital Signal Processing.* 2014. V. 25. P. 173-189. doi: 10.1016/j.dsp.2013.11.010.
41. Lee S-H. DWT based coding DNA watermarking for DNA copyright protection // *Inform. Sci.* 2014. V. 273. P. 263-286. doi: 10.1016/j.ins.2014.03.039.
42. Lee S-H., Lee E-J., Hwang W-J., Kwon K-R. Reversible DNA data hiding using multiple difference expansions for DNA authentication and storage // *Multimedia Tools & Applications.* 2018. V. 77. P. 19499-19526. 28p. doi: 10.1007/s11042-017-5379-1.
43. Lindsey J Cook, Jonathan P L Cox Methylated DNA labels for marking objects // *Biotechnol. Lett.* 2003. V. 25(1). P. 89-94. doi: 10.1023/a:1021713625176.
44. Liss M., Daubert D., Brunner K., Kliche K., Hammes U., Leiberer A., Wagner R. Embedding permanent watermarks in synthetic genes // *PLoS One.* 2012. V. 7(8). e42465. doi: 10.1371/journal.pone.0042465.
45. Luescher A.M., Koch J., Stark W.J., Grass R.N. Silica-encapsulated DNA tracers for measuring aerosol distribution dynamics in real-world settings // *Indoor Air.* 2021. Oct 21. doi: 10.1111/ina.12945.
46. Marillonnet S., Klimyuk V., Gleba Y. Encoding technical information in GM organisms // *Nat. Biotechnol.* 2003. V. 21(3). P. 224-6. doi: 10.1038/nbt0303-224.
47. Marillonnet S., Klimyuk V., Gleba Y. Identity tags revisited // *Nat. Biotechnol.* 2003. V. 21(9). P. 974-975. doi: 10.1038/nbt0903-974.
48. Mora C.A., Paunescu D., Grass R.N., Stark W.J. Silica particles with encapsulated DNA as trophic tracers // *Mol. Ecol. Resour.* 2015. V. 15(2). P. 231-41. doi: 10.1111/1755-0998.12299.
49. Morono Y., Kitagawa W., Kimura N., Noda N., Nakamura K., Kamagata Y. "Mark the gene": a method for nondestructive introduction of marker sequences inside the gene frame of transgenes // *Appl. Environ. Microbiol.* 2007. V. 73(15). P. 4915-4921. doi: 10.1128/AEM.00068-07.
50. Mueller S. Are Market GM Plants an Unrecognized Platform for Bioterrorism and Biocrime? // *Front. Bioeng. Biotechnol.* 2019. V. 7. P. 121. doi: 10.3389/fbioe.2019.00121.
51. Mueller S. On DNA Signatures, Their Dual-Use Potential for GMO Counterfeiting, and a Cyber-Based Security Solution // *Front. Bioeng. Biotechnol.* 2019a. V. 7. P. 189. doi: 10.3389/fbioe.2019.00189. eCollection 2019.
52. Mueller S., Jafari F., Roth D. A covert authentication and security solution for GMOs // *BMC Bioinformatics.* 2016. V. 17(1). P. 389. doi: 10.1186/s12859-016-1256-6.
53. Mueller S., Jafari F., Roth D. A covert authentication and security solution for GMOs // *BMC Bioinformatics.* 2016. V. 17(1). P. 389. doi: 10.1186/s12859-016-1256-6.
54. Pauli U. Identity tags // *Nat. Biotechnol.* 2003. V. 21(6). P. 598. doi: 10.1038/nbt0603-598.
55. Scotoni M., Koch J., Julian T.R., Clack L., Pitot A.K., Wolfensberger A., Grass R.N., Sax H. Silica nanoparticles with encapsulated DNA (SPED) - a novel surrogate tracer for microbial transmission in healthcare // *Antimicrob. Resist. Infect. Control.* 2020. V. 9(1). P. 152. doi: 10.1186/s13756-020-00813-7.
56. Shimanovsky B., Feng J., Potkonjak M. Hiding data in DNA // *LNCS* 2578. 2003. P. 373-386. doi: 10.1007/3-540-36415-3_24.
57. Shipman S.L., Nivala J., Macklis J.D., Church G.M. CRISPR-Cas encoding of a digital movie into the genomes of a population of living bacteria // *Nature.* 2017. V. 547(7663). P. 345-349. doi: 10.1038/nature23017.
58. Smith G.C., Fiddes C.C., Hawkins J.P., Cox J.P. Some possible codes for encrypting data in DNA // *Biotechnol. Lett.* 2003. V. 25(14). P. 1125-1130. doi: 10.1023/A:1024539608706.

59. Tai W.-L., Wang C.C.N., P.C.Y. Sheu, Tsai J.J.P. Data Hiding in DNA for Authentication of Plant Variety Rights // *J. Electr. Sci. Technol.* 2013. V. 11(1). P. 38-43. doi: 10.3969/j.issn.1674-862X.2013.01.008
60. Ullrich C., Luescher A.M., Koch J., Grass R.N., Sax H. Silica nanoparticles with encapsulated DNA (SPED) to trace the spread of pathogens in healthcare // *Antimicrob. Resist. Infect. Control.* 2022. V. 11(1). P. 4. doi: 10.1186/s13756-021-01041-3.
61. Wong P.C., Wong K.-K., Foote H. Organic data memory using the DNA approach // *Commun. ACM.* 2003. V. 46(1). P.95-98. doi: 10.1145/602421.602426
62. Yachie N., Ohashi Y., Tomita M. Stabilizing synthetic data in the DNA of living organisms // *Syst. Synth. Biol.* 2008. V. 2(1-2). P. 19-25. doi: 10.1007/s11693-008-9020-5.
63. Yamamoto N., Kajiura H., Takeno S., Suzuki N., Nakazawa Y. A watermarking system for labeling genomic DNA // *Plant Biotechnol.* 2014. V. 31. P. 241-248. doi: 10.5511/plantbiotechnology.14.0609b.
7. Brenner S., Williams S.R., Vermaas E.H., Storck T., Moon K., McCollum C., Mao J.I., Luo S., Kirchner J.J., Eletr S., DuBridge R.B., Burcham T., Albrecht G. In vitro cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. *Proc. Natl. Acad. Sci. USA.* 2000. V. 97(4). P. 1665-1670. doi: 10.1073/pnas.97.4.1665.
8. Chemeris A.V., Bikbulatova S.M., Chemeris D.A., Baimiev Al.Kh., Knyazev A.V., Kuluev B.R., Maksimov I.V. Should we be afraid of GMOs? The view of outside observers on the hysteria around. *Biomics.* 2014. V. 6. S. 77-138.
9. Chemeris A.V., Chemeris D.A., Baimiev Al.Kh., Knyazev A.V., Kuluev B.R., Maksimov I.V. The fight against GMOs as neolysenkoism. *Biomics.* 2015. V.7. C. 1-39.
10. Clelland C.T., Risca V., Bancroft C. Hiding messages in DNA microdots. *Nature.* 1999. V. 399(6736). P. 533-534. doi: 10.1038/21092.
11. Cook L.J., Cox J.P. Methylated DNA labels for marking objects. *Biotechnol. Lett.* 2003. V. 25(1). P. 89-94. doi: 10.1023/a:1021713625176.
12. Cox J.P. Bar coding objects with DNA. *Analyst.* 2001. V. 126(5). P. 545-547. doi: 10.1039/b101850g.
13. Cox J.P. Long-term data storage in DNA. *Trends Biotechnol.* 2001a. V. 19(7). P. 247-250. doi: 10.1016/s0167-7799(01)01671-7.
14. Feynman R.P. There's Plenty of Room at the Bottom: An Invitation to Enter a New Field of Physics. *Engin. Sci. (CalTech).* 1960. V. 23. P. 22-36.
15. Fister K., Fister I. Jr., Murovec J., Bohanec B. DNA labelling of varieties covered by patent protection: a new solution for managing intellectual property rights in the seed industry. *Transgenic Res.* 2017. V. 26. P. 87-95.
16. Fister K., Iztok Fister J.R. How to Store Wikipedia into a Forest Tree: Initial Idea. *MSIVISM.* 2014. P. 45-52.
17. Fister K., Iztok Fister J.R., Murovec J. The Potential of Plants and Seeds in DNA-Based Information Storage. *Understanding Information.* 2017. P. 69-81. doi: 10.1007/978-3-319-59090-5_4.
18. Garafutdinov R.R., Sakhabutdinova A.R., Chemeris A.V. Long-term room temperature storage of DNA molecules. *Biomics.* 2020. V. 12(4). P. 552-563. DOI: 10.31301/2221-6197.bmcs.2020-49 (In Russian)
19. Gibson D.G., Benders G.A., Andrews-Pfannkoch C., Denisova E.A., Baden-Tillson H., Zaveri J., Stockwell T.B., Brownley A., Thomas D.W., Algire M.A., Merryman C., Young L., Noskov V.N., Glass J.I., Venter J.C., Hutchison C.A. 3rd, Smith H.O. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science.* 2008.

References

- V. 319(5867). P. 1215-20. doi: 10.1126/science.1151721.
20. Gibson D.G., Glass J.I., Lartigue C., Noskov V.N., Chuang R.Y., Algire M.A., Benders G.A., Montague M.G., Ma L., Moodie M.M., Merryman C., Vashee S., Krishnakumar R., Assad-Garcia N., Andrews-Pfannkoch C., Denisova E.A., Young L., Qi Z.Q., Segall-Shapiro T.H., Calvey C.H., Parmar P.P., Hutchison C.A. 3rd, Smith H.O., Venter J.C. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*. 2010. V. 329(5987). P. 52-6. doi: 10.1126/science.1190719.
21. Gimalov F.R., Garafutdinov R.R., Kuluev B.R., Privalov L.Yu., Kiryanova O.Yu., Mikhailenko K.I., Urmanceev S.F., Chemeris A.V. Tracing in hydrogeology of underground watercourses using DNA tracers. *Biomics*. 2021. V.13(3). P.244-253. DOI: 10.31301/2221-6197.bmcs.2021-16 (In Russian)
22. Gustafsson C. For anyone who ever said there's no such thing as poetic gene. *Nature*. 2009. V. 458. P. 703. doi: 10.1038/458703a.
23. Hao M., Qiao H., Gao Y., Wang Z., Qiao X., Chen X., Qi H. A mixed culture of bacterial cells enables an economic DNA storage on a large scale. *Commun. Biol.* 2020. V. 3. P. 416. doi: 10.1038/s42003-020-01141-7.
24. Hashiyada M. Development of biometric DNA ink for authentication security. *Tohoku J. Exp. Med.* 2004. V. 204. P. 109-117.
25. Haughton D., Balado F. BioCode: two biologically compatible Algorithms for embedding data in non-coding and coding regions of DNA. *BMC Bioinformatics*. 2013. V. 14. P. 121. doi: 10.1186/1471-2105-14-121.
26. Heider D., Barnekow A. DNA Watermarking: Challenging Perspectives for Biotechnological Applications. *Curr. Bionf.* 2011. V. 6. P. 375-382. doi: 10.2174/157489311796904646.
27. Heider D., Barnekow A. DNA watermarks: a proof of concept. *BMC Mol. Biol.* 2008. V. 9. P. 40. doi: 10.1186/1471-2199-9-40.
28. Heider D., Barnekow A. DNA-based watermarks using the DNA-Crypt algorithm. *BMC Bioinformatics*. 2007. V. 8. P. 176. doi: 10.1186/1471-2105-8-176.
29. Heider D., Kessler D., Barnekow A. Watermarking sexually reproducing diploid organisms. *Bioinformatics*. 2008. V. 24(17). P. 1961-1962. doi: 10.1093/bioinformatics/btn342.
30. Heider D., Pyka M., Barnekow A. DNA watermarks in non-coding regulatory sequences. *BMC Res. Notes*. 2009. V. 2. P. 125. doi: 10.1186/1756-0500-2-125.
31. Holst-Jensen A., Spilsberg B., Arulandhu A.J., Kok E., Shi J., Zel J. Application of whole genome shotgun sequencing for detection and characterization of genetically modified organisms and derived products. *Anal. Bioanal. Chem.* 2016. V. 408(17). P. 595-614. doi: 10.1007/s00216-016-9549-1.
32. Jain S., Farnoud F., Schwartz M., Bruck J. Duplication-Correcting Codes for Data Storage in the DNA of Living Organisms. *IEEE Trans. Inform. Theory*. 2017. V. 63. P. 4996-5010. doi: 10.1109/TIT.2017.2688361.
33. Jiao S., Goutte R. Code for encryption hiding data into genomic DNA of living organisms. *9th Intern. Conf. Signal Processing*. 2008. doi: 10.1109/ICOSP.2008.4697576.
34. Jupiter D.C., Ficht T.A., Samuel J., Qin Q.M., de Figueiredo P. DNA watermarking of infectious agents: progress and prospects. *PLoS Pathog.* 2010. V. 6(6). e1000950. doi: 10.1371/journal.ppat.1000950.
35. Kac E. Genesis: a transgenic artwork. Art, Technology, Consciousness mind@large. R.Ascott (ed.). 2000. P.17-19.
36. Kong X.Z., Deuber C.A., Kittilä A., Somogyvári M., Mikutis G., Bayer P., Stark W.J., Saar M.O. Tomographic Reservoir Imaging with DNA-Labeled Silica Nanotracers: The First Field Validation. *Environ. Sci. Technol.* 2018. V. 52(23). P. 13681-13689. doi: 10.1021/acs.est.8b04367.
37. Lee S-H. DNA sequence watermarking based on random circular angle. *Digital Signal Processing*. 2014. V. 25. P. 173-189. doi: 10.1016/j.dsp.2013.11.010.
38. Lee S-H. DWT based coding DNA watermarking for DNA copyright protection. *Inform. Sci.* 2014. V. 273. P. 263-286. doi: 10.1016/j.ins.2014.03.039.
39. Lee S-H., Lee E-J., Hwang W-J., Kwon K-R. Reversible DNA data hiding using multiple difference expansions for DNA authentication and storage. *Multimedia Tools & Applications*. 2018. V. 77. P. 19499-19526. 28p. doi: 10.1007/s11042-017-5379-1.
40. Lindsey J Cook, Jonathan P L Cox Methylated DNA labels for marking objects. *Biotechnol. Lett.* 2003. V. 25(1). P. 89-94. doi: 10.1023/a:1021713625176.
41. Liss M., Daubert D., Brunner K., Kliche K., Hammes U., Leiherer A., Wagner R. Embedding permanent watermarks in synthetic genes. *PLoS One*. 2012. V. 7(8). e42465. doi: 10.1371/journal.pone.0042465.
42. Luescher A.M., Koch J., Stark W.J., Grass R.N. Silica-encapsulated DNA tracers for measuring aerosol distribution dynamics in real-world settings. *Indoor Air*. 2021. Oct 21. doi: 10.1111/ina.12945.
43. Marillonnet S., Klimyuk V., Gleba Y. Encoding technical information in GM organisms. *Nat. Biotechnol.* 2003. V. 21(3). P. 224-6. doi: 10.1038/nbt0303-224.

44. Marillonnet S., Klimyuk V., Gleba Y. Identity tags revisited. *Nat. Biotechnol.* 2003. V. 21(9). P. 974-975. doi: 10.1038/nbt0903-974.
45. Mora C.A., Paunescu D., Grass R.N., Stark W.J. Silica particles with encapsulated DNA as trophic tracers. *Mol. Ecol. Resour.* 2015. V. 15(2). P. 231-41. doi: 10.1111/1755-0998.12299.
46. Morono Y., Kitagawa W., Kimura N., Noda N., Nakamura K., Kamagata Y. "Mark the gene": a method for nondestructive introduction of marker sequences inside the gene frame of transgenes. *Appl. Environ. Microbiol.* 2007. V. 73(15). P. 4915-4921. doi: 10.1128/AEM.00068-07.
47. Mueller S. Are Market GM Plants an Unrecognized Platform for Bioterrorism and Biocrime? *Front. Bioeng. Biotechnol.* 2019. V. 7. P. 121. doi: 10.3389/fbioe.2019.00121.
48. Mueller S. On DNA Signatures, Their Dual-Use Potential for GMO Counterfeiting, and a Cyber-Based Security Solution. *Front. Bioeng. Biotechnol.* 2019a. V. 7. P. 189. doi: 10.3389/fbioe.2019.00189. eCollection 2019.
49. Mueller S., Jafari F., Roth D. A covert authentication and security solution for GMOs. *BMC Bioinformatics.* 2016. V. 17(1). P. 389. doi: 10.1186/s12859-016-1256-6.
50. Mueller S., Jafari F., Roth D. A covert authentication and security solution for GMOs. *BMC Bioinformatics.* 2016. V. 17(1). P. 389. doi: 10.1186/s12859-016-1256-6.
51. Pauli U. Identity tags. *Nat. Biotechnol.* 2003. V. 21(6). P. 598. doi: 10.1038/nbt0603-598.
52. Sakhabutdinova A.R., Mikhailenko K.I., Garafutdinov R.R., Kiryanova O.Yu., Sagitova M.A., Sagitov A.M., Chemeris A.V. Non-biological application of DNA molecules. *Biomics.* 2019. V.11(3). P. 344-377. DOI: 10.31301/2221-6197.bmcs.2019-28 (In Russian)
53. Scotoni M., Koch J., Julian T.R., Clack L., Pitol A.K., Wolfensberger A., Grass R.N., Sax H. Silica nanoparticles with encapsulated DNA (SPED) - a novel surrogate tracer for microbial transmission in healthcare. *Antimicrob. Resist. Infect. Control.* 2020. V. 9(1). P. 152. doi: 10.1186/s13756-020-00813-7.
54. Shimanovsky B., Feng J., Potkonjak M. Hiding data in DNA. *LNCS.* 2003. P. 373-386. doi:10.1007/3-540-36415-3_24.
55. Shipman S.L., Nivala J., Macklis J.D., Church G.M. CRISPR-Cas encoding of a digital movie into the genomes of a population of living bacteria. *Nature.* 2017. V. 547(7663). P. 345-349. doi: 10.1038/nature23017.
56. Smith G.C., Fiddes C.C., Hawkins J.P., Cox J.P. Some possible codes for encrypting data in DNA. *Biotechnol. Lett.* 2003. V. 25(14). P. 1125-1130. doi: 10.1023/A:1024539608706.
57. Tai W-L., Wang C.C.N., P.C.Y. Sheu, Tsai J.J.P. Data Hiding in DNA for Authentication of Plant Variety Rights. *J. Electr. Sci. Technol.* 2013. V. 11(1). P. 38-43. doi: 10.3969/j.issn.1674-862X.2013.01.008
58. Ullrich C., Luescher A.M., Koch J., Grass R.N., Sax H. Silica nanoparticles with encapsulated DNA (SPED) to trace the spread of pathogens in healthcare. *Antimicrob. Resist. Infect. Control.* 2022. V. 11(1). P. 4. doi: 10.1186/s13756-021-01041-3.
59. Venter J.C. Life at the speed of light. From the double helix to the dawn of digital life. Viking. N-Y. 2013. 240 P.
60. Vershinina Z.R., Kuluev B.R., Maksimov I.V., Mikhailova E.V., Gumerova G.R., Maleev G.V., Knyazev A.V., Baimiev An.Kh., Baimiev Al.Kh., Chemeris A.V. GMOs cannot be banned! *Biomics.* 2020. V.12(1). pp. 80-120. doi: 10.31301/2221-6197.bmcs.2020-6 (In Russian)
61. Wong P.C., Wong K-K., Foote H. Organic data memory using the DNA approach. *Commun. ACM.* 2003. V. 46(1). P.95-98. doi: 10.1145/602421.602426
62. Yachie N., Ohashi Y., Tomita M. Stabilizing synthetic data in the DNA of living organisms. *Syst. Synth. Biol.* 2008. V. 2(1-2). P. 19-25. doi: 10.1007/s11693-008-9020-5.
63. Yamamoto N., Kajiura H., Takeno S., Suzuki N., Nakazawa Y. A watermarking system for labeling genomic DNA. *Plant Biotechnol.* 2014. V. 31. P. 241-248. doi: 10.5511/plantbiotechnology.14.0609b.