

# ВИРУС ИСПАНСКОГО ГРИППА: ШТРИХИ К ПОРТРЕТУ СПУСТЯ 100 ЛЕТ

Е.П. Харченко

*ФГБУН Институт эволюционной физиологии и биохимии им. И.М. Сеченова РАН, Санкт-Петербург, Россия*

**Резюме.** Цель исследования состояла в сравнительном анализе молекулярных характеристик генов и белков пандемических штаммов вируса гриппа А для выявления особенностей вируса пандемии испанского гриппа 1918 г. Выявлено, что, в отличие от других пандемических штаммов 1957, 1968, 1977 и 2009 гг., гены вируса испанского гриппа содержат оптимальное количество протяженных комплементарных последовательностей, которые позволяют обеспечить сборку наборов его 8 генов по модели специфического их связывания по единственно возможной схеме и вызывать инфекционный процесс минимальным числом вирионов. Другие пандемические штаммы содержали в своих генах либо избыточное, либо недостаточное количество комплементарных последовательностей, что предполагает сборку набора генов в вирионе по разным моделям, в том числе и по стохастической модели, реализация которой может приводить к образованию потомства в виде роя зависимых от комплементации полуинфекционных вирионов и, соответственно, к уменьшенной их инфекционности. Анализ гена НА вируса испанского гриппа показал исключение из его трансляционного кода квартета триплетов аргинина (CGG, CGA, CGC и CGU), которое строго соблюдается у всех штаммов подтипа H1N1 на протяжении минувшего столетия. Кодирование аргинина в гене НА осуществляется за счет двух других его кодонов: AGG и AGA. Наконец, третья особенность вируса испанского гриппа заключается в принадлежности его NP гена, в отличие от других пандемических штаммов, к птичьему варианту, характеризующемуся более высоким содержанием аргинина и уменьшением числа лизина и адаптированному к повышенной температуре организма птиц. Преобладание аргинина сообщает молекуле NP более высокий положительный заряд за счет его гуанидиновой группировки, обуславливающий более сильное электростатическое взаимодействие с молекулами РНК, и, соответственно, более высокую термостабильность РНП, чем у РНП вирусов гриппа человека. Потенциальным следствием присутствия в вирусе испанского гриппа NP птичьего варианта могла бы быть его высокая патогенность, поскольку лихорадка в организме хозяина создает температурный оптимум для его репликации. Полученные данные по особенностям генома вируса испанского гриппа и другим пандемическим штаммам в сочетании с информацией об особенностях их белков могут быть использованы для прогнозирования предпандемических штаммов и выявления траектории возможной пандемической опасности.

**Ключевые слова:** компьютерный анализ, испанский грипп, пандемические штаммы, сборка генома, трансляционный код, H1 гемагглютинин, нуклеопротеин.

---

**Адрес для переписки:**

Харченко Евгений Петрович  
194223, Россия, Санкт-Петербург, пр. Мориса Тореза, 44,  
ИЭФБ им. И.М. Сеченова РАН.  
Тел./факс: 8 (812) 552-70-31 (служебн.); 8 904 338-22-80 (моб.).  
E-mail: neuro.children@mail.ru

**Contacts:**

Eugene P. Kharchenko  
194223, Russian Federation, St. Petersburg, Toreza pr., 44,  
Sechenov Institute of Evolutionary Physiology and Biochemistry.  
Phone/Fax: +7 (812) 552-70-31 (office); +7 904 338-22-80 (mobile).  
E-mail: neuro.children@mail.ru

**Библиографическое описание:**

Харченко Е.П. Вирус испанского гриппа: штрихи к портрету спустя 100 лет // Инфекция и иммунитет. 2018. Т. 8, № 3. С. 325–334.  
doi: 10.15789/2220-7619-2018-3-325-334

**Citation:**

Kharchenko E.P. The Spanish influenza virus: treats to the portrait after 100 years // Russian Journal of Infection and Immunity = Infektsiya i immunitet, 2018, vol. 8, no. 3, pp. 325–334. doi: 10.15789/2220-7619-2018-3-325-334

## THE SPANISH INFLUENZA VIRUS: TREATS TO THE PORTRAIT AFTER 100 YEARS

Kharchenko E.P.

*I. Sechenov Institute of Evolutionary Physiology and Biochemistry, Russian Academy of Sciences, St. Petersburg, Russian Federation*

**Abstract.** The purpose of the study was to compare molecular characteristics of genes and proteins of pandemic influenza strains and find features of the 1918 Spanish influenza virus. Computer analysis has shown that the genes of the Spanish influenza virus in contrast to other pandemic strains contain optimal quantity of long complementary sequences that allow to obtain a supramolecular assembly of 8 virus RNA in according to a model ensuring selective packing of one copy of each virus RNA by the only possible scheme and high transmission to induce infection by single virions. Other pandemic strains contain redundant or insufficient quantity of complementary sequences that allow an assembly of its genome by means of some models including a stochastic one and occurrence of virions with incomplete genome, that is influenza viruses can exist primarily as a swarm of complementation-dependent semi-infectious virions. Analysis of an HA gene of the Spanish influenza virus found out exclusion from its translation code four triplets (CGG, CGA, CGC и CGU) coding arginine. This exclusion is observed in all the H1N1 strains isolated during 100 years. Coding arginine in an HA gene of H1N1 strains is provided by only triplets AGG and AGA. A NP gene of the Spanish influenza virus in contrast to other pandemic strains is avian-like and its NP protein is characterized by elevated quantity of arginine and decreased quantity of lysine that is considered as viral adaptation to avian body temperature. Prevalence of arginine provides more high positive charge for the Spanish influenza NP protein and its more powerful interaction with RNA and consequently more high thermal stability of the its RNP in comparison with the RNP of other pandemic strains. Potential consequence of existence of the avian-like NP in the Spanish influenza virus could be its high pathogenicity as infection fever creates optimal temperature for virus replication. These new data obtained by computer analysis of genomes in the Spanish influenza virus and other pandemic strains (altogether information about its proteins) can potentially be used to track pre-pandemic strains among circulating influenza A viruses and detect the formation of a possible trajectory of pandemic alert.

**Key words:** computer analysis, Spanish influenza virus, pandemic influenza strains, genome packing, translation code, H1 hemagglutinin, nucleoprotein.

Уже 100 лет отделяют нас от разразившейся в 1918 г. пандемии испанского гриппа, унесшей более 50 млн человеческих жизней. По данным ВОЗ, ежегодная смертность от инфекции вирусом гриппа (ВГ) составляет 200 000–500 000 человек. Если ориентироваться на верхнюю границу этой оценки, то за 100 лет из-за гриппа не стало около 100 млн человек.

В поиске подходов распознавания пандемических штаммов (ПШ) ВГ А по молекулярным характеристикам показано, что их внутренние белки [белки полимеразного комплекса (PB1, PB2 и PA), нуклеопротеин (NP), матриксные белки (M1 и M2) и неструктурный белок 2 (NS2)] характеризуются постоянством числа и позиций определенных аминокислот и наличием блоков протяженных инвариантных последовательностей. Выявленные паттерны инвариантности внутренних белков позволяют безошибочно идентифицировать ПШ. Разделенные по их возникновению десятками лет и отличающиеся составом подтипов гемагглютинаина (НА) и нейраминидазы (НА), ПШ имеют сильное сходство по внутренним белкам, образуя особое подмножество, от которого в разной степени «отдалены» неПШ. Это позволяет предположить, что в природе возникновение ПШ вируса гриппа А связано с конвергенцией их внутренних белков к обнаруженным инвариантам пандемичности [2].

Обязательные атрибуты ПШ — высокая трансмиссивность и тяжесть вызываемой ими инфекции. Оба признака определяются сложной совокупностью особенностей генома и кодируемых им продуктов. Перенесенные за 100 лет пандемии существенно отличаются по числу инфицированных людей и смертельных исходов, и вирус испанского гриппа (ВИГ) имеет по этим характеристикам наивысшие, а ПШ 2009–2010 гг. — наименьшие показатели [4]. Представляется естественным сформулировать вопрос: что же отличает ВИГ от виновников последующих пандемий? Последующими ПШ были возникшие вновь реассортантные ВГ А. Важнейшим этапом реассортации (обмена) генов (являющейся наиболее быстрым механизмом изменения генома), как и обычной репродукции ВГ, является процесс упаковки его 8 фрагментов генома в вирион. Однако по прошествии 100 лет механизмы комплектации генома вирионов ВГ по-прежнему остаются невыясненными. В первом приближении можно было бы предположить, что в числе факторов, определяющих трансмиссивность ВГ, не только скорость репликации фрагментов генома и синтеза структурных белков, но и стабильность генов и сборка вирионов, укомплектованных полным набором генов и способных дальше самостоятельно воспроизводиться при инфекции другой особи [8]. Но действительность пред-

ставляется сложнее [3]. Поскольку внутренние белки ПШ имеют тенденцию к конвергенции их первичных структур [2], то для ответа на поставленный вопрос об особенностях ВИГ нами был предпринят сравнительный анализ геномов ПШ, который позволил выявить некоторые существенные отличия ВИГ от других ПШ, потенциально способные влиять на трансмиссивность и патогенность. Их описание является предметом данного сообщения.

## Материалы и методы

Для сравнительного компьютерного анализа были использованы штаммы A/Brevig Mission/1/18 (ПШ1918), H2N2 A/Japan/305/1957 (ПШ1957), H3N2 A/ICHI/2/1968 (ПШ1968), H1N1 A/USSR/90/1977 (ПШ1977), H1N1 A/California/04/2009 (ПШ2009), H5N1 A/Indonesia/5/2005, H7N9A/Fujian/02151/2017, H9N2A/Guangdong/MZ058/2016, H9N2 A/SHANHAI/1/2013, а также Zaire Ebolavirus/H.sapiens-rec/LBR/2014/Makona-L2014 ZsG. Из-за отсутствия полных данных по первичной структуре гена НА для ПШ1918 анализ НА проводился по данным для H1N1 A/SouthCarolina/1/18. Помимо перечисленных штаммов при анализе трансляционного кода (ТК) НА были использованы более 100 других выделенных в разные годы штаммов подтипа H1N1. С помощью компьютерного анализа в РНК всех штаммов ВГ выявляли комплементарные последовательности (КП) и аутоКП. Для всех белков каждого ПШ определяли ТК, а для ПШ и птичьих штаммов по белку NP выполнен сравнительный анализ аминокислотного состава\*. Источником первичных структур белков и генов анализированных вирусных штаммов служили общедоступные в Интернете базы данных (<http://www.ncbi.nlm.nih.gov> и GenBank).

## Результаты

Как показано многими исследователями, важную роль в селективности упаковки фрагментов генома ВГ играют комплементарные взаимодействия между РНК в рибонуклеопротеинах (РНП), особенно те, что реализуются комплементарными последовательностями (КП), находящимися в пределах рамок считывания [6]. Для более полного выяснения роли изби-

**Таблица 1. Комплементарные последовательности между генами A/BrevigMission/1/18**

Table 1. Complementary sequences between the A/BrevigMission/1/18 gene segments

|                   |     |             |
|-------------------|-----|-------------|
| TGCATCCTGTGCAGCAA | PA  | (1413–1429) |
| ACGTAGGACAGGTCGTT | NA  | (239–224)   |
| AATTCAGTGGTCCACAG | PB2 | (1683–1698) |
| TGAAGTCAACCAGTGTC | M1  | (424–410)   |
| TTGATTTCAGCTGATA  | PB2 | (1141–1155) |
| AACTAAGTCGACTCT   | PA  | (1986–1973) |
| TCTTCAATCTTCTGA   | PB1 | (1320–1334) |
| AGAAGTTTGAAGACT   | NS  | (695–682)   |
| ACATTCATCTATTTC   | HA  | (1010–1024) |
| TGAAAGGTAGATAAG   | NP  | (309–296)   |
| CTTGGGACATTTGA    | PB2 | (1819–1832) |
| GAAACCTGTAAACT    | PA  | (753–741)   |
| CTGTTCCACCATTG    | PB1 | (2238–2251) |
| GACAAGGTAGTAAC    | NA  | (996–984)   |
| GATCTTATTTCTTC    | NP  | (1454–1467) |
| GTAGAATAAAGAAG    | NS  | (704–692)   |
| ATGCAACTTCAGTG    | NA  | (236–249)   |
| TACGGTGAAGTCAAC   | M1  | (429–417)   |

рательного связывания 8 генов в геном был выполнен поиск полного набора КП между разными генами для каждого штамма, включенного в исследование.

Анализ распространенности наибольшей длины КП среди генов одного и того же штамма ВГ показал, что у 5 ПШ и 3 птичьих исследованных штаммов полностью КП (пКП) длиной в 13 нуклеотидов (н.) почти не встречаются (только одна у ПШ1918), редкими являются пКП в 12 н. У многих генов, но не у всех, имеются пКП в 11 н., и очень частыми для всех подтипов ВГ являются пКП в 10 н. Но часто встречающиеся пКП создадут помехи для комплектования полноценного генома с 8 разными генами, благоприятствуя образованию преимущественно дефектных вирионов. Принять 11 н. в качестве минимальной длины пКП нельзя по той причине, что в разных белках часто встречается идентичные тетрапептиды, и если бы комплементарное узнавание РНК осуществлялось по 10–11 н., то в клетках происходило бы очень частое комплексование разных мРНК. Идентичные же пентапептиды в разных белках встречаются весьма редко, поэтому нижнее значение длины для КП можно было бы принять за 14–15 н., но такие для генов ВГ, как отмечено выше, практически не встречаются. Однако комплементарное узнавание может быть реализовано, если допустить, что связность генов обеспечи-

\* В статье используется международный код аминокислот: А — аланин, С — цистеин, D — аспарагиновая кислота, E — глутаминовая кислота, F — фенилаланин, G — глицин, H — гистидин, I — изолейцин, K — лизин, L — лейцин, M — метионин, N — аспарагин, P — пролин, Q — глутамин, R — аргинин, S — серин, T — треонин, V — валин, W — триптофан, Y — тирозин. Для обозначения нуклеиновых оснований используется следующая аббревиатура: А — аденин, Г — гуанин, Ц — цитозин, Т — тимин.

**Таблица 2. Распределение аутокомплементарных последовательностей среди генов разных штаммов вируса гриппа**

Figure 2. Distribution of the autocomplementary sequences amongst the genes of various influenza virus strains

| Гены<br>Genes | ПШ1918<br>PS1918 | ПШ1957<br>PS1957 | ПШ1968<br>PS1968 | ПШ1977<br>PS1977 | ПШ2009<br>PS2009 | H5N1* | H7N9* | H9N2* |
|---------------|------------------|------------------|------------------|------------------|------------------|-------|-------|-------|
| PB2           | 4                | 0                | 0                | 0                | 0                | 0     | 0     | 0     |
| PB1           | 8                | 0                | 0                | 0                | 2                | 10    | 8     | 8     |
| PA            | 6                | 2                | 2                | 2                | 0                | 6     | 6     | 4     |
| HA            | 4                | 0                | 0                | 0                | 6                | 0     | 6     | 12    |
| NP            | 12               | 6                | 6                | 12               | 4                | 0     | 4     | 6     |
| NA            | 2                | 0                | 0                | 0                | 6                | 0     | 0     | 0     |
| M1            | 0                | 0                | 0                | 0                | 0                | 4     | 0     | 0     |
| NS            | 0                | 0                | 0                | 0                | 0                | 0     | 2     | 2     |

**Примечания:** \* соответственно штаммы H5N1 A/Indonesia/5/2005, H7N9 A/Fujian/02151/2017 и H9N2 A/Guangdong/MZ058/2016.

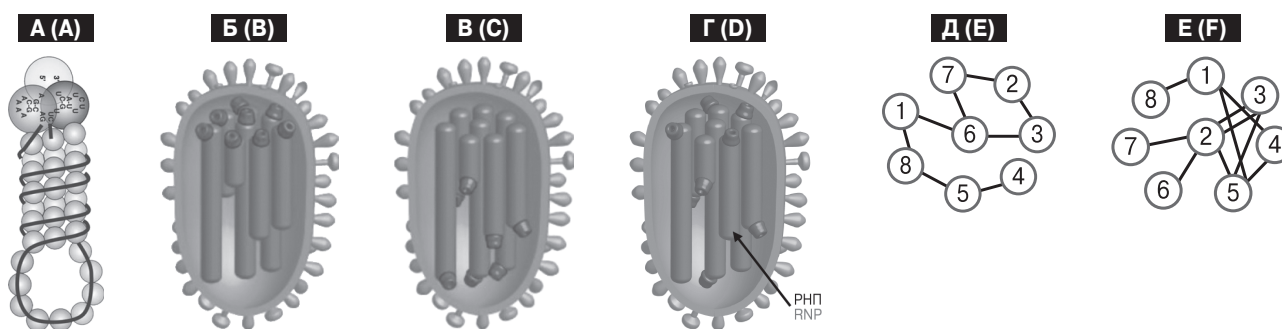
ПШ — пандемический штамм.

Notes: \* correspondingly H5N1 A/Indonesia/5/2005, H7N9 A/Fujian/02151/2017 and H9N2 A/Guangdong/MZ058/2016 strains. PS — pandemic strain.

ваются более длинными, но не полностью КП (нпКП), содержащими по одному некомплементарному нуклеотиду. Анализ показал, что лишь 5 штаммов из 10 исследованных содержали от 1 до 3 нпКП длиной в 16 н., нпКП длиной в 15 н. встречались (в количестве от 1 до 9) среди генов 6 из 8 штаммов, что также явно недостаточно для обеспечения полной связности всех генов у каждого штамма. Примечательно, что они отсутствовали у ПШ2009, а их максимальное число, по сравнению с другими, было у ПШ1918. С нпКП длиной в 14 н. связными оказываются лишь геномы ПШ1918 и ПШ1957. Полный перечень нпКП длиной от 17 до 14 н. в генах ПШ1918 представлен в табл. 1, а на рисунке Д изображен единственно возможный граф их связности. Среди других пандемических штаммов ПШ1918 первенствует по длине и оптимальному числу нпКП, Кроме того, в ге-

нах ПШ1918, по сравнению с другими штаммами, существенно выше содержание полностью аутоКП длиной в 11 н. (табл. 2), что может обеспечить им большую стабильность как при входе вириона в цитоплазму клетки при ее инфицировании, так и при завершении репликации вируса.

По сравнению с ПШ1918 у ПШ1957 не 8, а 10 нпКП между генами, и это количество уже не является оптимальным для формирования набора из 8 разных генов по двум причинам. В частности, PB2 и PB1 имеют нпКП более чем к 2 генам, что приводит к конкуренции между ними за связывание с другими генами и формированию комплексов с уменьшенным набором генов. Вторая причина связана с наличием у генов PB2 и PA двух нпКП друг к другу, что может приводить к формированию ансамблей генов, содержащих дубли одного и того же



**Рисунок. Модели упаковки РНП в вирионах вируса гриппа**

Figure. RNP packaging models in influenza virus virions

**Примечания:** А — модель РНП; В, С и D — варианты упаковки набора 8 РНП в вирионе; Д и Е модели — графы связности РНП для 8 генов соответственно для ПШ1918 и ПШ1957, в которых 1 — PB2, 2 — PB1, 3 — PA, 4 — HA, 5 — NP, 6 — NA, 7 — M1 и 8 — NS.

Notes: A — model of the RNP; B, C and D are variants of the packaging of the set of 8 RNP in the virion; E and F models are the RNP connectivity graphs for 8 genes for PS1918 and PS1957, respectively, in which 1 — PB2, 2 — PB1, 3 — PA, 4 — NA, 5 — NP, 6 — NA, 7 — M1 and 8 — NS.

гена либо состоящие только из цепи этих двух генов. На рисунке Е для ПШ1957 представлен лишь один из возможных графов укладки, обеспечивающий связность 8 разных генов. Обе отмеченные причины возникновения помех при избыточном количестве КП между одними и теми же парами генов особенно вероятны в случае ПШ1968, у которого 15 нпКП между генами. Примечательно, что при этой избыточности нпКП ген М1 оказывается вне связности с другими генами. По тем же упомянутым причинам, но при наличии 8 нпКП вне связности оказывается ген М1 у ПШ1977, а у ПШ2009 с 7 нпКП — вне связности ген NS.

Сравнение ТК генов у разных штаммов выявило у всех представителей подтипа Н1N1 общую особенность гена Н1. Она заключается в элиминации из ТК (табл. 3) квартета триплетов, кодирующих аргинин (CGA, CGG, CGC и CGT). С целью подтверждения этой уникальности ТК гена Н1 нами были проанализированы более сотни штаммов Н1N1, выделенных в течение столетия, и отклонений от этой консервативной характеристики Н1 не выявлено. Сопоставление ТК Н1 штаммов A/South Carolina/1/1918 и ПШ2009 свидетельствует не только о сохранении у них общей уникальности по кодированию аргинина, но и о близости по численным пропорциям использования триплетов остальных аминокислот. В частности, по подавляющему большинству аминокислот отмечается предпочтительное использование триплетов с максимальным содержанием аденина и/или тимина (урацила) с явной тенденцией наименьшего вовлечения триплетов с максимальным содержанием в них гуанина и/или цитозина.

Помимо гена Н1, в штаммах ПШ1918 и ПШ2009 исключение кодонов отмечается в генах NA и NS. Они характеризуются разным составом элиминированных кодонов, которые затрагивают не только аргинин, но и другие аминокислоты. В гене NP ПШ1918 и гене М1 ПШ2009 исключены по одному из триплетов, кодирующих соответственно лейцин и аланин. Для других подтипов NA свойственна иная картина исключения триплетов из ТК. У NA H5N1 A/Indonesia/5/2005 по аргинину исключены триплеты CGG, CGC и CGT, а также триплет GGC по глицину, а у H9N2 A/Shanghai/1/2013 по аргинину исключены триплеты CGC и CGT, а также триплет GCG по аланину. В NA ПШ1968 не используется триплет CGT, а у ПШ1957 3 триплета аргинина — CGT, CGC и CGA. О распространенности исключения триплетов из ТК свидетельствует анализ генов ВГ В и мононэгавирусов (например, у вируса Эбола). Таким образом, элиминация из ТК генов триплетов разных аминокислот, в том числе и триплетов аргинина, не является привилегией только ВГ подтипа Н1N1, но полная элиминация из ТК квартета триплетов одного и того же корня для аргинина, сохраняющаяся, по-видимому, более 100 лет, свойственна лишь для NA у ВГ подтипа Н1N1.

NP белок обеспечивает упаковку РНК в РНП, формируя остов, вокруг которого обвивается двойная спиральная структура РНК с антипараллельными ветвями (рис., А), при этом фосфатные группы РНК взаимодействуют с NP, а нуклеотидные основания экспонированы наружу. Априорно конформация фрагмента РНК в РНП будет определяться пространственной укладкой NP и составом его основных

**Таблица 3. Трансляционный код гемагглютининов Н1N1 пандемических штаммов**

Table 3. The translation code of the H1N1 pandemic strain hemagglutinins

| A/South Carolina/1/1918 |   |    |     |   | A/California/04/2009 |     |   |    |     |   |    |     |   |    |     |   |    |     |   |    |     |   |    |
|-------------------------|---|----|-----|---|----------------------|-----|---|----|-----|---|----|-----|---|----|-----|---|----|-----|---|----|-----|---|----|
| TCT                     | S | 7  | TGT | C | 10                   | TTT | F | 9  | TAT | Y | 17 | TCT | S | 8  | TGT | C | 9  | TTT | F | 8  | TAT | Y | 13 |
| TCC                     | S | 6  | TGC | C | 6                    | TTC | F | 10 | TAC | Y | 9  | TCC | S | 4  | TGC | C | 6  | TTC | F | 11 | TAC | Y | 14 |
| TCA                     | S | 16 | TGA | Z | 1                    | TTA | L | 11 | TAA | Z | 0  | TCA | S | 18 | TGA | Z | 0  | TTA | L | 5  | TAA | Z | 1  |
| TCG                     | S | 3  | TGG | W | 11                   | TTG | L | 9  | TAG | Z | 0  | TCG | S | 1  | TGG | W | 10 | TTG | L | 10 | TAG | Z | 0  |
| ACT                     | T | 7  | AGT | S | 6                    | ATT | I | 10 | AAT | N | 26 | ACT | T | 9  | AGT | S | 7  | ATT | I | 18 | AAT | N | 27 |
| ACC                     | T | 8  | AGC | S | 11                   | ATC | I | 8  | AAC | N | 16 | ACC | T | 2  | AGC | S | 9  | ATC | I | 7  | AAC | N | 14 |
| ACA                     | T | 20 | AGA | R | 12                   | ATA | I | 14 | AAA | K | 21 | ACA | T | 23 | AGA | R | 14 | ATA | I | 12 | AAA | K | 27 |
| ACG                     | T | 2  | AGG | R | 8                    | ATG | M | 8  | AAG | K | 12 | ACG | T | 3  | AGG | R | 4  | ATG | M | 7  | AAG | K | 15 |
| CCT                     | P | 2  | CGT | R | 0                    | CTT | L | 2  | CAT | H | 10 | CCT | P | 2  | CGT | R | 0  | CTT | L | 1  | CAT | H | 8  |
| CCC                     | P | 5  | CGC | R | 0                    | CTC | L | 6  | CAC | H | 3  | CCC | P | 4  | CGC | R | 0  | CTC | L | 5  | CAC | H | 7  |
| CCA                     | P | 9  | CGA | R | 0                    | CTA | L | 9  | CAA | Q | 10 | CCA | P | 9  | CGA | R | 0  | CTA | L | 15 | CAA | Q | 7  |
| CCG                     | P | 3  | CGG | R | 0                    | CTG | L | 14 | CAG | Q | 7  | CCG | P | 5  | CGG | R | 0  | ATG | L | 10 | CAG | Q | 8  |
| GCT                     | A | 6  | GGT | G | 10                   | GTT | V | 5  | GAT | D | 15 | GCT | A | 5  | GGT | G | 9  | GTT | V | 7  | GAT | D | 13 |
| GCC                     | A | 7  | GGC | G | 2                    | GTC | V | 9  | GAC | D | 10 | GCC | A | 9  | GGC | G | 4  | GTC | V | 5  | GAC | D | 13 |
| GCA                     | A | 18 | GGA | G | 22                   | GTA | V | 9  | GAA | E | 26 | GCA | A | 17 | GGA | G | 13 | GTA | V | 17 | GAA | E | 24 |
| GCG                     | A | 6  | GGG | G | 10                   | GTG | V | 9  | GAG | E | 9  | GCG | A | 2  | GGG | G | 14 | GTG | V | 7  | GAG | E | 11 |

аминокислот (лизина, аргинина и гистидина). Хотя все структурные белки ВГ А, по сравнению с поверхностными белками НА и NA, характеризуются выраженной эволюционной консервативностью, у NP птичьих штаммов следует ожидать некоторые особенности. Более высокая температура тела птиц, по сравнению с человеком, должна бы влиять на прочность комплекса РНК-NP, и одним из молекулярных механизмов адаптации ВГ в организме птиц могло бы быть изменение состава основных аминокислот, в частности аргинина и лизина, имеющих существенно отличные потенциалы ионизации и, соответственно, разную силу взаимодействия с фосфатными группировками РНК в составе РНП.

Сравнительный анализ аминокислотного состава NP у ПШ ВГ и у птичьих штаммов выявил (табл. 4), что ПШ1918, в отличие от других ПШ, содержит NP птичьего типа, а особенностью последнего является измененное соотношение содержания аргинина и лизина: более низкое содержание лизина и повышенное содержание аргинина при практически неизменном количественном соотношении дикарбоновых аминокислот — аспарагиновой и глутаминовой.

В других белках (даже в PB1, PB2 и PA, отличающихся противоположной по суммарному заряду полярностью) птичьих и человеческих штаммов ВГ таких сдвигов в содержании основных аминокислот не выявлено. Аргинин по своим размерам значительно больше лизина и имеет гуанидиновый радикал с более мощным ионизационным потенциалом, чем NH<sub>2</sub>-группа у лизина. Поэтому повышение содержания аргинина может быть существенным фактором в повышенной термостабильности РНП у ВГ птичьего происхождения.

## Обсуждение

Выполненный компьютерный анализ геномов и белков ПШ позволил дополнить «портрет» ВИГ тремя новыми «штрихами»: 1) его гены содержат более протяженные (и оптимальные по количеству) нпКП, потенциально способные участвовать в комплектации полного на-

бора 8 генов; 2) в нуклеотидной последовательности НА гена исключен из ТК квартет триплетов аргинина, что оказалось консервативным признаком для всех ВГ подтипа H1N1; 3) NP ВИГ, в отличие от других ПШ, принадлежит к птичьему варианту, для которого характерно иное соотношения аргинина и лизина. Каждая из этих характеристик потенциально способна влиять на различные уровни репликации ВИГ в организме хозяина, преломляясь и в механизмах патогенности.

*Комплементарные последовательности и упаковка генома.* Исторически объяснение механизма упаковки 8 фрагментов генома в состав вириона ВГ основывалось на двух гипотезах. Согласно первой комплектование генов является случайным процессом [7]. Но по элементарным математическим расчетам по стохастической модели при инфекционном процессе возможно образование лишь 0,4% вирионов с полным комплектом генов и способных соответственно быть самостоятельно инфекционными. По второй гипотезе, процесс упаковки генов характеризуется селективностью, обеспечиваемой совокупностью сигналов взаимного узнавания генов — содержащимися в них КП. В качестве таких сигналов, как показано экспериментально, служат не только концевые последовательности, фланкирующие открытую рамку считывания генов, но и последовательности в пределах рамки считывания [6], что заметно усилило аргументацию в пользу гипотезы селективного комплектования генов в вирионы ВГ. Однако существование множества подтипов ВГ, огромная численность штаммов и распространенность реассортации между разными подтипами ВГ практически исключают возможность существования универсальной сети сигналов распознавания РНП в процессе упаковки генома.

Трехмерный анализ комплексов РНП выявил несколько вариантов их взаимного расположения в вирионе по модели «7+1» (рис., Б, В, Г) — группирование по окружности в разных сочетаниях 7 генов вокруг центрально расположенного гена [10], которую условно можно обозначить как «моновирионную» модель,

**Таблица 4. Аминокислотный состав нуклеопротеинов разных подтипов вируса гриппа**

Figure 4. Amino acid compositions in the nucleoproteins of various influenza virus subtypes

|      |                        | K  | R  | H | D  | E  | P  | C | L  | I  | V  | A  | Y  | W | F  | G  | M  | N  | Q  | S  | T  |
|------|------------------------|----|----|---|----|----|----|---|----|----|----|----|----|---|----|----|----|----|----|----|----|
| H1N1 | A/Brevig Mission/1/18  | 16 | 54 | 6 | 22 | 37 | 17 | 6 | 30 | 31 | 22 | 39 | 15 | 6 | 17 | 41 | 25 | 25 | 21 | 40 | 28 |
| H2N2 | A/Japan/305/1957       | 24 | 46 | 5 | 24 | 35 | 18 | 6 | 31 | 28 | 23 | 39 | 15 | 6 | 17 | 42 | 25 | 28 | 21 | 38 | 27 |
| H3N2 | A/AICHI/2/1968         | 23 | 48 | 5 | 24 | 36 | 18 | 6 | 32 | 27 | 22 | 41 | 15 | 6 | 17 | 40 | 26 | 27 | 21 | 39 | 25 |
| H1N1 | A/USSR/90/1977         | 23 | 47 | 5 | 23 | 36 | 17 | 6 | 32 | 30 | 22 | 40 | 15 | 6 | 17 | 41 | 24 | 26 | 21 | 39 | 28 |
| H1N1 | A/California/04/2009   | 21 | 50 | 7 | 23 | 36 | 16 | 6 | 32 | 28 | 26 | 40 | 13 | 6 | 17 | 42 | 23 | 26 | 21 | 40 | 25 |
| H5N1 | A/Indonesia/5/2005     | 15 | 55 | 6 | 21 | 37 | 16 | 6 | 32 | 29 | 25 | 39 | 14 | 6 | 18 | 41 | 23 | 28 | 22 | 38 | 27 |
| H7N9 | A/Fujian/02151/2017    | 15 | 55 | 6 | 21 | 37 | 16 | 6 | 32 | 29 | 23 | 40 | 13 | 6 | 18 | 41 | 24 | 29 | 22 | 38 | 27 |
| H9N2 | A/Guangdong/MZ058/2016 | 14 | 56 | 6 | 21 | 37 | 16 | 6 | 32 | 29 | 24 | 40 | 13 | 6 | 18 | 41 | 23 | 28 | 22 | 39 | 27 |

предполагающую упаковку каждого вириона ансамблем всех 8 генов ВГ. Следует заметить, что множественность вариантов комплектования генов по модели «7+1» уже не служит свидетельством строгого детерминирования комплексования генов в вирионах ВГ. Более того, последующее исследование авторов модели «7+1» отдельных вирионов ВГ А и В выявило, что до 20% вирионов (в зависимости от штамма) содержат 5, 6 или 7, а не 8 генов, то есть являются не полностью укомплектованными по генному составу. Авторы, сохраняя приверженность модели селективного комплектования генов, объясняют неполноту состава генома в отдельных вирионах «большой гибкостью процесса упаковки генома, чем предполагалось ранее» [9], не допуская включение в него случайности и игнорируя результаты более раннего исследования [3], в котором уже было показано, что до 90% вирионов ВГ А (в зависимости от штамма) не экспрессируют, по крайней мере, один из главных белков. По мнению авторов [3], следствием легкости реассортации генов между разными штаммами являются трудности в комплектовании полноценных вирионов, и большинство их из-за дефектности способны лишь на один раунд инфекции, не влияя заметно на проявление самого инфекционного процесса. Само же существование ВГ А рассматривается, главным образом, как рой зависимых от комплементации полуинфекционных вирионов и, следовательно, традиционные анализы инфективности ВГ, возможно, сильно искажают истинный инфекционный потенциал вирусной популяции. По существу это означает использование ВГ также популяционной модели генома, которая совершенно не исключает ее сосуществование в природе с моновирионной моделью упаковки генома. Весомой аргументацией в пользу популяционной модели служат данные по морфологии ПШ2009, в частности существование у них палочковидных структур и образования вирионами кластеров [1]. Передача ВГ кластерами позволяет сохранить его контагиозность благодаря возможности реализации комплементации генов, если вирионы содержат неполные наборы генов. Палочковидная форма вирионов хорошо адаптирована к модели сборки генома по варианту рисунка Г, когда полюса РНП, связывающихся через КП в РНК, смещены относительно друг друга, что увеличивает общую длину ансамбля генов, а также устраняет ограничения на позицию КП в РНК и допускает упаковку повторяющихся ансамблей генов. Вариативность морфологии ВГ должна определяться согласованными изменениями в поверхностных белках НА и NA и белками матрикса М1 и М2.

Оба упомянутых исследования [3, 9], как и данные по морфологии ВГ, сильно подрывают аргументацию в пользу существования одной модели упаковки полноценного генома ВГ и скорее свидетельствуют о сосуществовании по крайней мере 2 моделей. Нами впервые специально исследовано распределение КП среди генов разных подтипов ВГ А, которое на примере ПШ и птичьих штаммов показало варьирование их распределения в генах. Число КП в генах штамма может быть оптимальным, недостаточным или избыточным. В случае оптимального количества КП, как это выявлено при моделировании упаковки генома ПШ1918, в вирионе формируется полный набор 8 генов по модели «7+1» либо последовательного связывания генов через КП с образованием завитка, имитирующего модель «7+1». Сама модель «7+1» может допускать несколько сценариев, проиллюстрированных на рисунке (Б, В и Г) [12] и обусловленные тем, в какой ветви «шпильчатой» структуры РНК в РНП (рис., А) и на каком уровне расположены КП в РНП соответствующих пар генов. К примеру, если КП расположены в антипараллельных ветвях РНК в РНП, то их РНП располагаются однонаправленно относительно друг друга, то есть круговые петли их РНП располагаются по одну сторону. Варианты Б и В упаковки РНП на рисунке [12] представляются возможными, но маловероятными (поскольку требуют согласованного позиционирования КП во всех 8 генах ВГ), что аргументируется разными позициями нПКП в генах ПШ1918 (табл. 1) и других ПШ. Реалистичным представляется вариант Г на рисунке, допускающий разнонаправленное расположение РНП и смещение их полюсов. Преимущество существования между генами оптимального количества КП связано с образованием полноценных по набору генов вирионов, способных самостоятельно вызвать инфекционный процесс при заражении хозяина минимальным числом вирионов, то есть одним вирионом.

При избыточном числе КП, как в случае ПШ1957, потенциально возможна реализация нескольких конкурирующих вариантов образования наборов генов. Один из них основан на модели «7+1». Другие реализуются в наборы с разным уменьшенным числом РНП, и, следовательно, часть образующегося потомства — это популяция зависимых от комплементации полуинфекционных вирионов, репликация которых в клетках хозяина будет невозможной при заражении единичным вирионом. Следовательно, инфекционность таких штаммов ВГ будет сильно варьировать и уступать таковой у вирусов с оптимальным числом КП. При недостаточности количества КП между генами (как это характерно для ПШ1968

и ПШ2009), практически все потомство вируса из-за хаотичности в образовании наборов генов представляет собой в разной степени зависимые от комплементации полуинфекционные вирионы. Если обратиться к статистике минувших за столетие пандемий, то в ряду ПШ1918—ПШ1957—ПШ1968—ПШ2009 прослеживается последовательное уменьшение по числу инфицированных людей и смертельных исходов от ПШ1918 к ПШ2009, которое согласуется с особенностями распределения КП в генах этих ПШ и с возможностью существования в природе разных моделей упаковки генома ВГ.

С биологической точки зрения, саму возможность существования двух моделей интеграции генома ВГ можно рассматривать как проявление вырожденности механизма интеграции генов, и они не исключают друг друга. Вырожденность функции (или механизма) предполагает существование нескольких путей ее (его) реализации, при этом возможно вовлечение разных участников. Классическим примером вырожденности функции служит генетический код, в котором кодирование одной и той же аминокислоты представлено разными кодонами — от одного до шести. Если оценивать обе модели интеграции генома ВГ, то популяционная модель интеграции генома в аспекте воспроизведения/выживания потомства обладает преимуществом, поскольку предоставляет широкий выбор в комплекции вирионов генами при инфекционном процессе и меньшей чувствительностью к мутационным дефектам генов. При моновирионной модели интеграции генома преимущество самостоятельного воспроизведения каждым вирионом может быть утрачено в результате быстрого накопления в генах мутаций, не совместимых с воспроизведением потомства, например, из-за изменений в самих КП, приводящих к утрате «узнавания» генов друг другом, необходимого для формирования ансамбля из 8 разных генов.

*Особенности трансляционного кода.* Нет оснований объяснять особенности ТК гена Н1 различиями в организации генома и трансляционного механизма между хозяином и ВГ, поскольку исключение каких-либо кодонов отсутствует в генах полимеразного комплекса ВГ, трансляция которых вместе с мРНК остальных белков происходит в одном и том же клеточном компартменте, что отвергает, соответственно, предположение о детерминации элиминации триплетов из РНК Н1 недостаточной концентрацией загруженных аминокислотами тРНК в клеточном пуле. Возможно, что феномен исключения из ТК гена Н1 квартета триплетов аргинина связан с обеспечением поверхностными белками ВГ адаптивной изменчивости, но он не распространяется на ген NA и на другие

подтипы NA. Из-за постоянства проявления на протяжении 100 лет в громадном множестве штаммов подтипа Н1N1 мало оснований полагать, что этот феномен является случайным по своей природе. Он не связан с функционированием самого белка NA (так как аргинин в генетическом коде представлен 6 триплетами и не исключен из самого NA за счет кодирования его двумя другими триплетами — AGG и AGA) и остается пока не объяснимым.

*Особенности NP.* Среди ограничений, налагаемых на трансмиссивность вирусных реассортантов, можно особенно выделить температурный фактор, поскольку тройка хозяев (птицы, свинья и человек) штаммов ВГ существенно контрастирует по температуре их тела. В жизненном цикле ВГ температуро- и pH-зависимыми являются конформационные изменения NA и процесс слияния оболочки вируса с эндосомальной мембраной [5, 11]. Другое ограничение должно быть связано с NP, так как прочность взаимодействия NP с вирусной РНК в РНП будет определяться, помимо особенностей структуры белка NP, температурой тела хозяина ВГ.

Априорно можно полагать, что процессы репликации и транскрипции, сопряженные с диссоцированием NP от РНК, имеют температурный оптимум, детерминированный температурой тела конкретного вида животного, под которую адаптирован аминокислотный состав NP циркулирующего в нем штамма ВГ. У человека температура тела равна 36,6°C, у свиней — 38–40°C, а у птиц — 42,5°C. Столь существенная разница температур тела птиц и человека должна обязательно отразиться на аминокислотном составе NP их ВГ, в частности на содержании в них положительно заряженных аргинина и лизина, так как через эти аминокислоты осуществляется электростатическое (наиболее сильное) взаимодействие NP с фосфатными радикалами геномной РНК в вирионе, и от силы этого взаимодействия будут зависеть скорость репликации и транскрипции и прежде всего возможность их самой реализации. Для NP у ВГ птиц должна быть характерна по крайней мере более высокая, по сравнению с ВГ человека, положительная заряженность, что и проявляется в более повышенном содержании у них аргинина (табл. 4). Преобладание аргинина сообщает за счет его гуанидиновой группировки молекуле NP более высокий положительный заряд. Это обуславливает более сильное электростатическое взаимодействие с молекулами РНК, и, соответственно, более высокую термостабильность, чем в NP у ВГ человека. Особенности эволюции NP, являющегося высококонсервативным белком, проявляются преимущественно в синонимичных заменах заряженных аминокислот. При не-



изменности сумм соответственно положительно (лизина, аргинина, гистидина) и отрицательно (глутаминовой и аспарагиновой кислот) заряженных аминокислот и постоянстве их положения в первичной структуре при адаптации к тому или иному виду менялись количественные соотношения в парах лизин/аргинин и глутаминовая кислота/аспарагиновая кислота, независимо от подтипа их HA и NA. У всех ВГ в NP содержание аргинина значительно выше, чем лизина.

По-видимому, реассортанты ВГ птиц с воспринятым NP геном от ВГ человека будут резко ограничены в распространении среди птиц, но потенциально способны циркулировать в человеческой популяции при условии функциональной совместимости NP гена (белка) с другими генами (белками) самого ВГ и хозяина. Для реассортанта, сочетающего птичий NP и остальные гены от ВГ человека, неоптимальными будут условия репликации до подъема температуры у зараженного им субъекта, но повышение температуры у него при развитии инфекции окажется благом для репликации такого реассортанта и еще большим отягощением состоянием зараженного субъекта. Возможно, что присутствие в ПШ1918 NP гена птичьего происхождения (предполагается, что ПШ1918 полностью птичьего происхождения [14]) внесло также свой вклад в особенно высокую его патогенность. В этой связи примечательным представляется сопоставление ПШ1918, ПШ2009 и современных птичьих штаммов H5N1, H7N9 и H9N2 по трансмиссивности и патогенности для человека. Птичьим штаммам свойственна слабая трансмиссивность среди людей, но они

вызывают высокую смертность в случае инфицирования. Для ПШ1918 и ПШ2009 свойственна высокая трансмиссивность, объясняемая их принадлежность обоим к подтипу H1N1, но они сильно различаются по вызываемому ими уровню смертности. Сходны ПШ1918 и птичий ВГ по вызываемой ими высокой смертности и содержанию птичьего NP. При анализе факторов, обуславливающих патогенность ВГ, и конструировании живых вакцинных штаммов нельзя, по-видимому, игнорировать природу NP-гена.

В завершение хотелось бы подчеркнуть, что благодаря целеустремленности исследователей и развитию методов науки в период 1995–2005 гг. были воссозданы гены ВИГ [13], что явилось блестящим вкладом в вирусологию и открыло возможность заглянуть в прошлое — в истоки одной из наиболее опустошительных пандемий гриппа в современной истории человечества. Помимо ранее известных данных о различиях между ВИГ и ПШ2009 (одиночные мутации в разных генах, наличие стоп-кодонов в белке PB1-F2 или COOH-концевая делеция в NS1 белке [1]), выявленные нами данные по КП в генах ВИГ, как и особенности его NP и HA, дополняют описание его «портрета» и углубляют молекулярную интерпретацию его непревзойденных другими ПШ трансмиссивность и патогенность для человека. В сочетании с информацией по особенностям белков ПШ [2] новые данные по анализу геномов ПШ расширяют возможности выявления и отслеживания предПШ и прогнозирования формирования пандемической опасности и могут быть полезными для конструирования новых вакцин.

## Список литературы/References

1. Киселев О.И. Геном пандемического вируса гриппа A/H1N1V-2009. СПб.—М.: Компания «Димитрейд График Групп», 2011. 163 с. [Kiselev O.I. Genom pandemicheskogo virusa grippa A/H1N1V-2009. [The genome of pandemic influenza virus A/H1N1V-2009]. *St. Petersburg—Moscow: Dimitrade Grafic Group, 2011. 163 p.*]
2. Харченко Е.П. Инвариантные паттерны внутренних белков пандемических вирусов гриппа // Инфекция и иммунитет. 2015. Т. 5, № 4. С. 323–330. [Kharchenko E.P. The invariant patterns of the internal proteins of pandemic influenza viruses. *Infektsiya i immunitet = Russian Journal of Infection and Immunity, 2015, vol. 5, no. 4, pp. 323–330. doi: 10.15789/2220-7619-2015-4-323-330 (In Russ.)*]
3. Brooke C.B., Ince W.L., Wrammert J., Ahmed R., Wilson P.C., Bennink J.R., Yewdel J.W. Most influenza A virions fail to express at least one essential viral protein. *J. Virol., 2013, vol. 87, no. 6, pp. 3155–3162. doi: 10.1128/JVI.02284-12*
4. Chan M. Statement to press by director — general of the World Health Organization 11 June 2009. World now at the start of 2009 influenza pandemic.
5. Daniels R.S., Downie J.C., Hay A.J., Knossow M., Skehel J.J., Wang M.L., Wiley D.C. Fusion mutants of the influenza virus hemagglutinin glycoprotein. *Cell, 1985, vol. 40, no. 2, pp. 431–439.*
6. Gerber M., Isel C., Moules V., Marquet R. Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol., 2014, vol. 22, no. 8, pp. 446–455. doi: 10.1016/j.tim.2014.04.001*
7. Hutchinson E.C., von Kirchbach J.C., Gog J.R., Digard P. Genome packaging in influenza A virus. *J. Gen. Virol., 2010, vol. 91, pt. 2, pp. 313–328. doi: 10.1099/vir.0.017608-0*
8. Lakdawala S.S., Fodor E., Subbarao K. Moving on out: transport and packaging of influenza viral RNA into virions. *Annu. Rev. Virol., 2016, vol. 3, pp. 411–427. doi: 10.1146/annurev-virology-110615-042345*
9. Nakatsu S., Sagara H., Sakai-Tagawa Y., Sugaya N., Noda T., Kawaoka Y. Complete and incomplete genome packaging of influenza A and B viruses. *MBio, 2016, vol. 7(5), pp. e01248-16. doi: 10.1128/mBio.01248-16*

10. Noda T., Sugita Y., Aoyama K., Hirase A., Kawakami E., Miyazawa A., Sagara H., Kawaoka Y. Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat. Commun.*, 2012, vol. 3: 639. doi: 10.1038/ncomms1647
11. Skehel J.J., Bayley P.M., Brown E.B., Martin S.R., Waterfield M.D., White J.M., Wilson I.A., Wiley D.C. Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. *Proc. Natl. Acad. Sci. USA*, 1982, vol. 79, no. 4, pp. 968–972.
12. Sugita Y., Sagara H., Noda T., Kawaoka Y. The configuration of viral ribonucleoprotein complexes within the influenza A virion. *J. Virol.*, 2013, vol. 87, no. 23, pp. 12879–12884. doi: 10.1128/JVI.02096-13
13. Taubenberger J.K., Hultin J.V., Morens D.M. Discovery and characterization of the 1918 pandemic influenza virus in historical context. *Antivir. Ther.*, 2007, vol. 12, no. 4, pt. B, pp. 581–591.
14. Taubenberger J.K., Reid A.H., Lourens R.M., Wang R., Jin G., Fannin T.G. Characterization of the 1918 influenza virus polymerase genes. *Nature*, 2005, vol. 437, no. 7060, pp. 889–893.

---

**Автор:**

**Харченко Е.П.**, д.б.н., ведущий научный сотрудник Института эволюционной физиологии и биохимии им. И.М. Сеченова РАН, Санкт-Петербург, Россия.

**Author:**

**Kharchenko E.P.**, PhD, MD (Biology), Senior Researcher, Sechenov Institute of Evolutionary Physiology and Biochemistry of the Russian Academy of Sciences, St. Petersburg, Russian Federation.

---

Поступила в редакцию 18.10.2017  
Принята к печати 13.06.2018

Received 18.10.2017  
Accepted 13.06.2018