

*Materials and Methods.* We used algorithms LipoP 1.0, SignalP-4.0, TargetP, PSORTb, Cello, TMHMM and BOMP. Sequence of B-cell epitopes were predicted by the algorithm ABCpred or BCPREDS. Server modules BLASTX and CLASTALW were used for the searching of homologues proteins.

*Results.* As the first step of the work it was aimed to the analyzing the complete genome of reference strain MmmSc Gladysdale. It consisted in finding open reading frames and selecting the corresponding amino acid sequences. In total 1095 were used. The sequences were selected by the presence of lipobox and signal peptides that could indicate their transmembrane localization. Among these sequences it was conducted by the absence of  $\alpha$ - helices and  $\beta$ - barrels, which is typical for membrane lipoproteins. Additionally, prediction algorithms were used protein localization in the cell, making it possible in a complex with other proteins results select five of them. In sequences of these proteins B-cell epitopes localization were predicted.

*Conclusions.* A comprehensive theoretical analysis allowed us to find among the 1095 ORFs MmmSc strains Gladysdale five that with high probability encode secretory proteins. The results of these studies can be used for subsequent work towards a highly specific diagnostic test systems and effective vaccines against contagious bovine pleuropneumonia.

**Keywords:** contagious bovine pleuropneumonia, secreted proteins, bioinformatics

**УДК 619:578:616.98:578.831**

### **ГЕНОТИПУВАННЯ ПЕСТІВІРУСІВ ТВАРИН НА ОСНОВІ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ**

**Герілович А.П., Лиманська О.Ю., Горайчук І.В., Ареф'єв В.Л., Гема І.О.**

Національний науковий центр «Інститут експериментальної і клінічної ветеринарної медицини»,  
м. Харків, e-mail: antger@vet.kharkov.ua

*Вивчено філогенетичні зв'язки вірусів класичної чуми свиней різних генотипів, що циркулюють у різних географічних регіонах. Продемонстрована перспективність проведення філогенетичного аналізу для генотипування та молекулярного маркування мікроорганізмів.*

**Ключові слова:** вірус класичної чуми свиней, генотип, філогенетичний аналіз.

Найчастішою подією молекулярної еволюції біомакромолекул, ДНК та РНК, є нуклеотидні заміни, які накопичуються під час незалежної еволюції послідовностей, що дивергують від спільної батьківської форми. Середня кількість замін на один нуклеотидний сайт для двох гомологічних послідовностей біологічних молекул двох видів організмів визначає еволюційну відстань. Встановлення еволюційної відстані полягає у знаходженні різниці на рівні генетичного матеріалу та її взаємозв'язку між часом еволюції. Саме знання цієї залежності в подальшому робить можливим побудову філогенетичних дерев, визначення часу дивергенції таксонів на основі порівняння первинних структур генетичних біомакромолекул, реконструкцію історії біот, дослідження порівняно недавніх змін, вивчення еволюції генів у просторі та часі [1, 2]. У найпростішому випадку, коли заміни рідкісні (або час еволюції невеликий), можна припустити, що число замін у парі послідовностей прямо пропорційно часу їх еволюції. При цьому відстань між двома послідовностями відображає подвійний час, що минув з моменту дивергенції (за умови, що швидкості накопичення замін у двох послідовностях були однакові). Визначення еволюційних відстаней можна здійснити за допомогою низки методів:

- однопараметричного методу Джукса і Кантора, який використовує вірогідність заміни одного нуклеотиду на інший;
- методу Тадзими-Ней, який припускає рівність швидкостей заміщення серед сайтів і між заміщеннями за типом транзицій та трансверсій;
- двопараметричного методу Кімури, у якому вірогідності транзицій не дорівнюють вірогідностям трансверсій;
- трипараметричного методу Тамури, який враховує не тільки різну частоту транзицій та трансверсій, але й вміст гуаніну та цитозину.

Наведені методи основані на припущенні, що швидкості фіксації нуклеотидних замін у ході молекулярної еволюції даного сімейства гомологічних нуклеотидних послідовностей є рівномірними. Топологія побудованих з урахуванням еволюційних відстаней філогенетичних дерев дозволяє отримати достовірну інформацію, зокрема, про генотипову характеристику збудника, спектр ізолятів інфекційного агента, що циркулює на певній території [3, 4].

Зазначені процеси значно ускладнюють здійснення генотипування збудників інфекційних захворювань та, отже, потребують ретельного дослідження з боку фахівців ветеринарної медицини та молекулярної біології, що є важливим для розвитку галузі тваринництва України та розробки системи прогнозування та всебічного контролю спалахів інфекційних, зокрема вірусних, захворювань сільськогосподарських тварин.

**Метою** даної роботи є вивчення філогенетичних зв'язків основних пестівірусів тварин.

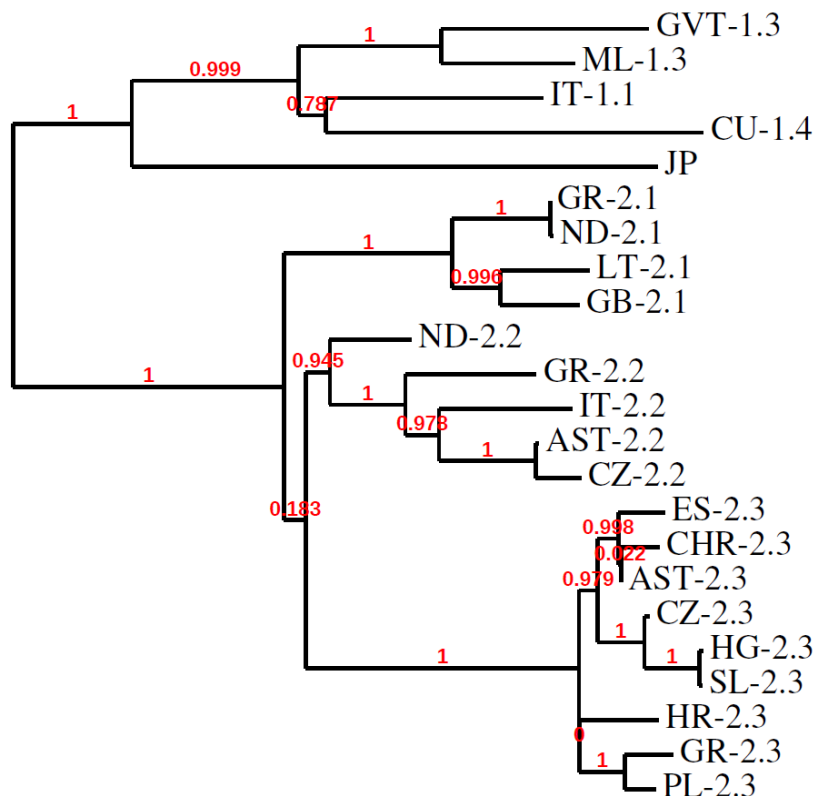
**Матеріали та методи.** Для проведення філогенетичного аналізу використовували програми Mega 4, ver. 4.0.2 [5], PhyML ver. 3.0 [6]. Для побудови традиційних дендрограм на основі послідовностей генів та геномних РНК пестівірусів використовували дистанційно-матричний метод – метод зв'язування найближчих сусідів (Neighbour joining) та максимальної економії (maximum parsimony). Як критерій достовірності топології отриманої дендрограми використовували індекси повторень (bootstrap).

Для проведення філогеографічних досліджень пестівірусів сільськогосподарських тварин нами з міжнародних баз даних були вибрані послідовності геномного матеріалу представників сімейства *Flaviviridae* роду *Pestivirus* – вірусу діареї (ВД) великої рогатої худоби (ВРХ) типу 1 (bovine viral diarrhoea virus 1, BVDV-1); вірусу діареї ВРХ типу 2 (bovine viral diarrhoea virus 2, BVDV-2); вірусу класичної чуми свиней (КЧС) (classical swine fever virus, CSFV). Усі послідовності отримано у форматах FASTA (\*.fasta) або GenBank (\*.gb).

**Результати роботи.** Геном пестівірусів представлений одонитковим ланцюгом молекули РНК позитивної полярності довжиною близько 12,3 тис. нуклеотидів (у низки ізолятів ВД ВРХ ця величина може значно варіювати) з однією відкритою рамкою зчитування, фланкованою нетрансльованими областями [7].

Відомо, що геномна РНК пестівірусів кодує поліпротеїн N<sup>pro</sup>-C-E<sup>ns</sup>-E1-E2-p7-NS2-3-NS4A- NS4B-NS5A-NS5B, який складається з близько 4000 амінокислотних залишків і розщеплюється клітинними та вірусними протеазами на 11–12 структурних і неструктурних білків [8]. N<sup>pro</sup> належить до N-кінцевої протеази; E<sup>ns</sup> – до оболонкових структурних глікопротеїнів з РНКазною активністю. Крім E<sup>ns</sup> до числа структурних протеїнів належать білок капсиду С і білки оболонки Е1 та Е2. Решта білків є неструктурними. При цьому NS2-3 характеризується множинною ферментативною активністю, зокрема хеліказною та нуклеозидтрифосфатази; для NS5B доведена активність, подібна до активності РНК-залежної РНК-полімерази.

Аналіз топології філогенетичних дерев, побудованих на основі послідовностей повногеномних РНК вірусу КЧС, свідчить про те, що кожний з відомих на цей час генотипів вірусу утворює чітко відокремлений кластер (рис. 1). Таким чином, результати філогенетичного аналізу мають важливе значення для генотипування мікроорганізмів, а отже, можуть бути основою їх молекулярного маркування.



**Рис. 1.** Філогенетичне дерево, побудоване на основі повністю секвенованих послідовностей геномних РНК вірусу КЧС, циркулюючого на території різних країн. **Позначки:** GVT – Гватемала; ML – Малайзія; GB – Великобританія; IT – Італія; JP – Японія; GR – Німеччина; ND – Нідерланди; CU – Куба; LT – Литва; AST – Австрія; CZ – Чехія; ES – Іспанія; CHR – Хорватія; HG – Угорщина; SL – Словаччина; PL – Польща

```

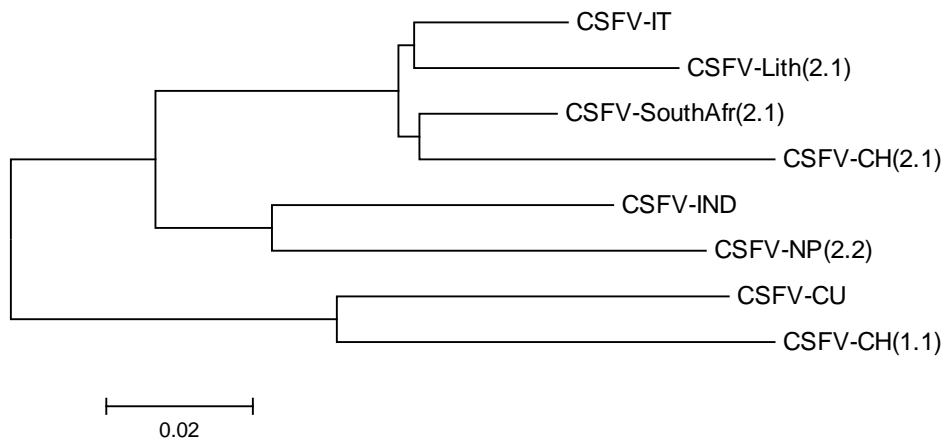
BVDV1-JP ACAACCACAG TACGGATATA CAGGAGGACT AAGCCTTTTC CTCATAGG
BVDV1-SWI . . . . .TGTTG. .G. . . .C. . . . .A. .AT. A G.A. .A. . . . .A
BVDV1-SWI . . . . .TGTTG. .G. . . .C. . . . .A. .AT. A G.A. .A. . . . .A
BVDV1-CH TTG. .GTT. . . .A. C. . . .C. . . .CAC. . . .G. .C. .CT. . . .
BVDV1-DM . .G. . .AG. .G. . .CC. .T. .A. . . .C. . . .G. .C. .T. . .A
BVDV1-GR TTG. . .AG. . .AA. C. . . . .G. C. . . .A. C. . .T. C. . .
BVDV1-Neth . .G. . .AG. .G. . .CC. .T. .A. . . .C. . . .G. .C. .T. . .A
BVDV1-CA . .C. .GTT. .G. .A. CG. .T. A. . . .GA. .G. . . . .AT. . . .
BVDV1-CL . . .G. .TT. .G. .A. C. . . .A. .ACAC. .A. . . . .CT. . . .
BVDV1-CL . . .GAGTG. .G. T. .CG. . .TA. .AT. C .G. .G. .C. . . . .
BVDV1-EG . . . .TGTT. . .A. C. . . . .CAC. . . . . .CT. . . . .
BVDV1-USA . . . .GTT. .G. .T. CT. .T. A. . . .AC .GC. . . . .A. . . . .
BVDV1-CH . . . .GTC. .G. .A. CG. .T. A. . . .GTC .G. .C. . . .AT. . . .
BVDV1-ARG GGG. .TGTT. . . .A. CT. . .C. . . .CTC .A. . . . .CT. . . .
BVDV1-ARG . . . .GTGA . . . .C. . . .T. A. . .T. C .A. .A. . . . .
BVDV1-JP . . . .GTT. .G. .A. CG. .T. A. . . .GTC .G. . . . .AT. . . .
BVDV2-JP . . . .TGTTA .TA. .C. . . .T. A. .A. .C .CC. A. . . .AA.GG. .A
BVDV2-GR . . . .TGTT. .TA. . . . .T. A. .A. T. .C. .A. . .AG.GG. .A
BVDV2-CH . . . .GTT. .TA. A. C. . .T. A. .A. . .CT. A. . .AG.GG. .A
BVDV2-USA . . . .TGTC. .TA. .C. . . .T. A. .A. . .CT. A. . .AG.GG. .A
BVDV2-USA . . . .TGTC. .TA. .C. . . .T. A. .A. . .CT. A. . .AG.GG. .A
BVDV2-ARG . .G. .TGTT. .TA. A. C. . .TGAA. .A. . .CT. A. . .G.GG. .
CSFV-CU . .GAAGTG. .AA. .C. .T . . . .AGAG . . .C. .C. .G. .C. .A
CSFV-IT . .GAAGTG. .GAA. .CC. T .A. .AGAG .A. . .C. .G. .C. .A
CSFV-IND . .GAGTG. .AA. .CT. T . . . .AGAG . . . .C. .A. . . .A
CSFV-Lith . .GAAGTG. .GAA. .CC. T .A. .AGAG . . . .C. .A. .C. .A
CSFV-NP . .GGAGTG. .AAA. CT. T . . . .AGAG . . . . .G. . . .A
CSFV-CH . .GAAGTG. .AAA. CC. T T. A. . .AGAG . . . .C. .C. .G. .C. .A
    
```

**Рис. 2.** Фрагмент множинного вирівнювання послідовностей гена, що кодує білок E2 пестівірусів тварин, циркулюючих на території різних країн. **Позначки:** USA – США; CH – Китай; IT – Італія; DM – Данія; IT – Італія; NZ – Нова Зеландія; SWI – Швейцарія; JP – Японія; GR – Німеччина; CA – Канада; CL – Чилі; Neth – Нідерланди; ARG – Аргентина; CU – Куба; NP – Непал; IND – Індонезія; Lith – Литва; EG – Єгипет

Дослідження варіабельних областей геномів вірусів дозволяють, зокрема, визначити напрямки їх еволюції, походження первісної інфекції, класифікувати ізоляти патогенів [9]. Множинне вирівнювання представлених у міжнародних базах даних послідовностей геномної РНК пестівірусів, що циркулюють у різних географічних регіонах, показало, що найменш консервативними є послідовності генів, які кодують протеїни N<sup>pro</sup> та E2 (рис. 2). Це свідчить про можливість генотипування та диференціювання вірусів з урахуванням результатів філогенетичного аналізу на основі послідовностей саме цих генів.

Результати філогенетичного аналізу, проведеного на основі повністю секвенованих послідовностей гена E2 вірусів КЧС, переконують у можливості проведення генотипування (встановлення підтипу) збудника на основі цього гена – адже кожний генотип вірусу утворює на дендрограмі чітко відокремлений кластер (рис. 3).

На цей час класична чума свиней, яку реєструють на території низки країн, завдає серйозного економічного збитку промисловому свиноводству. Проблема контролю розповсюдження вірусу КЧС є актуальною для таких країн, а також для тих, які є вільними від КЧС. Вивчення та розуміння механізмів еволюційного розвитку вірусу є необхідною складовою розробки надійних сучасних засобів діагностики та профілактики. Саме філогенетичний аналіз дозволяє оцінити ступінь еволюційної спорідненості різних штамів вірусу КЧС, з'ясувати походження спалахів, вивчати еволюцію вірусів, визначати генотипову належність збудника [10], що разом зі з'ясуванням шляхів його потрапляння в господарства є, безумовно, необхідним кроком при створенні системи довгострокового прогнозування спалахів та запобігання подальшого поширення КЧС.



**Рис. 3.** Філогенетичне дерево, побудоване на основі повністю секвенованих послідовностей гена E2 вірусу класичної чуми свиней, циркулюючого на території різних країн. **Позначки:** IT – Італія; Lith – Литва; SouthAfr – Південно-Африканська Республіка; IND – Індонезія; NP – Непал; CU – Куба; CH – Китай

З іншого боку, з метою мінімізації соціально-економічних наслідків спалахів КЧС у ряді країн свиней щеплюють живою вакциною. Крім того, значну увагу приділяють розробці вакцин нового покоління [11]. Ефективність вакцини значною мірою залежить від генотипу вірусу, виявленого на певній території. Наприклад, відомо, що протягом останніх десятиліть епідемії КЧС в популяціях диких кабанів та поголів'я свійських свиней у країнах Європи викликані штамами вірусу КЧС генотипу 2.3, які є вірулентними та викликають високу захворюваність та летальність [12]. Тобто необхідно враховувати поточну ситуацію на місцях та з'ясовувати, штами вірусу КЧС якого генотипу переважають у конкретному випадку. Отже, використання результатів філогенетичного аналізу, безумовно, сприятиме підвищенню ефективності вакцинації.

**Висновки.** Доведено варіабельність гену E2, що кодує білок оболонки вірусу КЧС. Проведено філогенетичний аналіз пестівірусів тварин на основі послідовностей гену E2, який продемонстрував можливість генотипування на основі саме цього гена. Продемонстрована перспективність проведення філогенетичного аналізу для молекулярного маркування вірусу класичної чуми свиней.

#### Список літератури

1. Riddle, B. R. The molecular phylogeographic bridge between deep and shallow history in continental biotas II [Text] / B. R. Riddle // Trends in Ecology and Evolution. - 1996. - Vol. 11. - P. 207-211.
2. Hewitt, G. M. Speciation, hybrid zones and phylogeography — or seeing genes in space and time [Text] / G. M. Hewitt // Molecular Ecology. - 2001. - Vol. 10. - P. 537-549.
3. Attwood, T.K. Progress in bioinformatics and the importance of being earnest [Text] / T.K. Attwood, C.J. Miller // Biotechnological Annu. Rev. - 2002. - № 8. - P. 1-54.
4. Леск, А. Введение в биоинформатику [Текст]: пер. с англ. – М.: Бином, 2009. – 318 с.
5. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0 [Text] / K. Tamura, J. Dudley, M. Nei, S. Kumar // Mol. Biol. Evolution. - 2007. - Vol. 24. - P. 1596-1599.
6. Phylogeny.fr: robust phylogenetic analysis for the non-specialist [Text] / A. Dereeper [et al.] // Nucleic Acids Res. - 2008. - Vol. 36. - P. 465-469.
7. Meyers, G. Molecular characterization of pestiviruses [Text] / G. Meyers, H.-J. Thiel // Adv. Virus Res. - 1996. - Vol. 47. - P. 53-118.
8. Genetic diversity of pestiviruses: identification of novel groups and implication for classification [Text] / P. Becher [et al.] // Virology. - 1999. - Vol. 262. - P. 64-71.
9. Белақ, С. Раннее предупреждение [Текст] / С. Белақ // Бюллетень МАГАТЭ. – 2006. - № 48. – С. 62-63.
10. Власова Анастасия Николаевна. Филогенетический анализ изолятов вируса классической чумы свиней и вируса репродуктивно-респираторного синдрома свиней, циркулирующих на территории России и Белоруссии : дисс. ... канд. Биол. наук: 03.00.03 / Власова Анастасия Николаевна. – Москва, 2003. – 121 с. – Библиогр. : с. 108-121.
11. Efficacy of marker vaccine candidate CP7\_E2alf against challenge with classical swine fever virus isolates of different genotypes [Text] / S. Blome [et al.] // Vet. Microbiol. – 2014. - Vol. 169, № 1-2. – P. 8-17.
12. Novel rope-based sampling of classical swine fever shedding in a group of wild boar showing low contagiousity upon experimental infection with a classical swine fever field strain of genotype 2.3 [Text] / S. Mouchantat [et al.] // Vet. Microbiol. – 2014. - Vol. 170, № 3-4. – P. 425-429.

## GENOTYPING OF THE ANIMAL PESTIVIRUSES BASED ON PHYLOGENETIC ANALYSIS

Gerilovich A.P., Limanskaya O.Yu., Gorajchuk I.V., Arefyev V.L., Gema I.A.

National Scientific Center "Institute experimental and Clinical  
Veterinary Medicine", Kharkov

*Goal.* The topology of constructed phylogenetic trees based on the evolutionary distance allows to obtain the reliable information, particularly about the genotypic characterization, the isolates spectrum of the infected agents circulating in a particular area. The aim of this work is to study the phylogenetic relationships of basic animal pestiviruses.

*Methods.* The softwares Mega 4, ver. 4.0.2, and PhyML, ver. 3.0, are used for the phylogenetic analysis. The methods of linking nearest neighbors (Neighbour joining) and maximum savings (maximum parsimony) are used for a dendrograms construction.

*Results.* The phylogenetic relationships of classical swine fever viruses (CSFV) circulating in different geographic regions are studied. It is shown that each of the currently known genotypes of CSFV forms a clearly separated cluster. It is shown that the phylogenetic analysis is a necessary component of effective vaccination, and the establishment of long-term forecasting of outbreaks and the prevention of further spread of CSF.

*Conclusion.* The variability of E2 gene encoding the coat protein of the CSFV is proved. A phylogenetic analysis of animal pestiviruses based on E2 gene sequences demonstrated the possibility of genotyping on the basis of this gene. The promising of phylogenetic analysis for the molecular labeling of the classical swine fever virus is demonstrated.

**Keywords:** classical swine fever virus, genotype, phylogenetic analysis.