

Реконструкция древних геномов и внутривидовая филогения *Aggregatibacter aphrophilus*

Иванова А.О.^{1*}, Павлова А.В.¹, Боцманов Е.И.¹, Цедилина Т.Р.¹,
Перевозчикова А.А.², Березина Н.Я.², Крицкий А.А.¹, Климук Е.И.¹,
Северинов К.В.¹

¹ ООО «Биотехнологический кампус», Москва, Россия

² Научно-исследовательский институт и Музей антропологии им. Д.Н. Анучина,
Московский государственный университет им. М.В. Ломоносова, Москва, Россия

* aivanova@biotc.ru

Ключевые слова: палеогеномика; филогенетика; биоинформатика; сборка *de novo*; *Aggregatibacter aphrophilus*

Мотивация и цель: *Aggregatibacter aphrophilus* – грамотрицательная, неподвижная, факультативно анаэробная палочка. В норме – комменсал ротовой полости, оппортунистический патоген, способный вызывать бактериемию, раневые инфекции, абсцессы мозга, инфекционный эндокардит у человека [1]. Входит в группу этиологических агентов инфекционного эндокардита НАСЕК [2]. Известно 23 полных генома штаммов данного вида, выделенных в течение 1932 – 2022 гг. В ходе палеогеномных исследований человеческих останков нами определено два древних генома *A. aphrophilus* возрастом 2170 ± 23 и 6012 ± 28 лет и проведен их функциональный и филогенетический анализ.

Методы и алгоритмы: Биоинформатический анализ проводили с использованием программ “FastP”, “SPAdes”, “metabat”, “PhyloPhlAn”, “checkm”, “roary”, “prokka”, “Blast”. Аутентификация древней ДНК осуществлялась при помощи “MapDamage2”. Филогенетический анализ был проведен с использованием “iqtree”. Для визуализации результатов использовали пакет “seaborn” (python) и “itol” [3].

Результаты: Реконструированы и проанализированы два полных генома *A. aphrophilus* возрастом 2170 ± 23 лет (AAA18C_KGZ_143BC) из зуба мужчины, погребенного в могильнике Кен-Су (Киргизия, Нарынская область), и 6012 ± 28 лет (AAA25C_RU_3899) из зуба мужчины, погребенного в могильнике Прогресс 2 (Россия, Ставропольский край).

Оба древних генома не принадлежат ни одной из филогенетических клад современных штаммов *A. aphrophilus*. В геноме AAA18C_KGZ_143BC отсутствуют регуляторные и структурные гены пилей адгезии *rcpABC/tadA-Z*. Эти гены присутствуют у всех современных клинических изолятов [4]. В геноме AAA25C_RU_3899 присутствуют некоторые гены кластера пилей адгезии (*rpcA*, *rpcB*, *rpcC*, *tadD*, *tadB*, *tadZ*). В обоих древних геномах отсутствуют гены адгезинов *aae* и *etaA*, характерные для современных клинических штаммов. Однако оба древних генома содержат кластер *lptA-G*, кодирующий основной фактор патогенности грамотрицательных бактерий, а также гены *lolA* и *oppA*, кодирующие белки, участвующие в колонизации макроорганизма. Гены, ассоциированные с устойчивостью к эльфамицину у современных изолятов, отсутствуют в древних геномах.

Выводы: Собраны и проанализированы два древних полных генома *A. aphrophilus* и продемонстрированы отличия от современных геномов этого вида, подтверждающие приобретение факторов патогенности и антибиотико-резистентности в ходе эволюции.

Финансирование: Исследование поддержано ООО «Биотехнологический кампус».

Reconstruction of ancient *Aggregatibacter aphrophilus* genomes and establishment of their intraspecific phylogeny

Ivanova A.O.^{1*}, Pavlova A.V.¹, Botsmanov E.I.¹, Tsedilina T.R.¹,
Perevozchikova A.A.², Berezina N.Y.², Kritsky A.A.¹, Klimuk E.I.¹, Severinov K.V.¹

¹ LLC “Biotechnological campus”, Moscow, Russia

² Research Institute and Museum of Anthropology named after D.N. Anuchin,
Lomonosov Moscow State University, Moscow, Russia

* aivanova@biotc.ru

Key words: paleogenomics; phylogenetics; bioinformatics; *de novo* assembly; *Aggregatibacter aphrophilus*

Motivation and Aim: *Aggregatibacter aphrophilus* is a Gram-negative, nonmotile, facultatively anaerobic rod bacterium, a member of the HACEK group of etiological agents of infective endocarditis [1]. A part of normal oral flora [2], it is an opportunistic pathogen causing bacteremia, wound infections, brain abscesses, and infective endocarditis in humans. The genomes of 23 *A. aphrophilus* strains collected between 1932–2022 are known. The purpose of this work was to establish evolutionary relationship between modern and ancient *A. aphrophilus* isolates.

Methods and Algorithms: Bioinformatics analysis was carried out using the software “FastP”, “SPAdes”, “metabat”, “PhyloPhlAn”, “checkm”, “roary”, “prokka”, “Blast”. Authentication of ancient DNA was conducted with “MapDamage2” package. Phylogenetic analysis was performed using “iqtree”. The “seaborn” python-based package and “itol” were used for the result visualization [3].

Results: Two complete genomes of *A. aphrophilus*, a 2170 ± 23 years old AAA18C_KGZ_143BC and a 6012 ± 28 years old (AAA25C_RU_3899), were reconstructed and analyzed. The ancient DNA used to reconstruct AAA18C_KGZ_143BC was isolated from a tooth of a man buried in the Ken-Su burial ground (Kyrgyzstan, Naryn region). The AAA25C_RU_3899 genome was reconstructed from DNA isolated from a tooth of a man buried in the Progress 2 burial ground (Russia, Stavropol Krai).

The ancient genomes do not belong to phylogenetic clades of modern *A. aphrophilus* strains. The AAA18C_KGZ_143BC genome lacks the *rcpABC/tadA-Z* regulatory and structural genes of adhesion pili. These genes are present in all extant clinical isolates [4]. In the AAA25C_RU_3899 genome, some adhesion pili genes (*rpcA*, *rpcB*, *rpcC*, *tadD*, *tadB*, *tadZ*) are present. In both ancient genomes, the *ae* and *emaA* adhesins genes characteristic of modern clinical strains are absent. However, both ancient genomes contain the *lptA-G* cluster encoding the main pathogenicity factor of Gram-negative bacteria, as well as the *lolA* and *oppA* genes encoding proteins involved in host

colonization. Genes associated with resistance to elfamycin in modern isolates are absent from ancient genomes

Conclusion: We have assembled and analyzed two ancient *A. aphrophilus* complete genomes. Comparisons with modern genomes of this species is consistent with acquisition of pathogenicity and antibiotic resistance factors during recent evolution of the species.

Funding: The study is supported by LLC “Biotechnological campus”.

Список литературы/References

1. Nørskov-Lauritsen N. Classification, identification, and clinical significance of Haemophilus and Aggregatibacter species with host specificity for humans. *Clin Microbiol Rev.* 2014;27(2):214-240. doi 10.1128/CMR.00103-13
2. Khaledi M., Sameni F., Afkhami H., Hemmati J. et al. Infective endocarditis by HACEK: a review. *J Cardiothorac Surg.* 2022;17(1):185. doi 10.1186/s13019-022-01932-5
3. Ancient DNA. Methods and protocols / ed. Beth Shapiro et al. Springer protocols. New York: Humana Press, 2019
4. Kittichotirat W., Bumgarner R.E., Asikainen S., Chen C. Identification of the pangenome and its components in 14 distinct Aggregatibacter actinomycetemcomitans strains by comparative genomic analysis. *PLoS One.* 2011;6(7):e22420. doi 10.1371/journal.pone.0022420