

УДК 550.72+ 624.138.9+ 004.67

## ИЗМЕНЕНИЯ МИКРОБИОМА КАРБОНАТНОЙ ГЛИНЫ: ОПЫТ ПРИМЕНЕНИЯ

### АЛГОРИТМОВ АНАЛИЗА БОЛЬШИХ ДАННЫХ

О.А. Софinskaya<sup>1</sup>, Е.Ю. Полякова<sup>1</sup>, Н.Е. Гоголева<sup>2</sup>

<sup>1</sup>Казанский федеральный университет, Казань, Россия

<sup>2</sup>Институт клеточного и внутриклеточного симбиоза УрО РАН, Оренбург, Россия

ushik2001@mail.ru

## ANALYSIS OF CARBONATE CLAY MICROBIOME CHANGES

### USING BIG DATA ALGORITHMS

О.А. Sofinskaya<sup>1</sup>, Е.Y. Polyakova<sup>1</sup>, N.E. Gogoleva<sup>2</sup>

Kazan Federal University, Kazan, Russia

Institute for Cellular and Intracellular Symbiosis of Ural Branch of RAS, Orenburg, Russia

ushik2001@mail.ru

**Аннотация.** Исследование посвящено изучению связи состава микробного сообщества с некоторыми свойствами карбонатно-глинистого грунта. Разработан и опробован алгоритм, позволяющий выявить устойчивые микробные консорциумы и реконструировать их роль в укреплении глин. Изменение микробиома инициировано внесением донорского сообщества спелеотем в глину и вариацией условий замачивания: длительности и парциального давления CO<sub>2</sub>. В исследовании учтены данные таксономического профилирования по гену 16S рРНК, физико-химические и физико-механические свойства глины, свойства порового раствора. Выделено 8 условных микробных сообществ, отличающихся по трофическим стратегиям и адаптациям к различным условиям, и 5 бактериальных родов-доминантов с наибольшим потенциалом изменения глин. Полученные результаты пригодны для анализа потенциала и прогноза применения биогеотехнологий. Биоцементация как метод укрепления глинистых грунтов с использованием бактерий для осаждения минералов является перспективным направлением инженерной геологии. Биоцемент, по сравнению с традиционным цементом, обладает лучшими технологическими характеристиками и меньше загрязняет окружающую среду [1]. В настоящее время в качестве потенциальных бактерий-кандидатов для биоцементации изучают алкалофильных продуцентов уреазы, денитрификаторов, восстановителей железа, сульфатредукторов, а также продуцентов карбоангидразы. Однако в силу высокой вариативности состава микробного сообщества, анализ и прогноз его влияния на инженерные свойства грунта в течение длительного времени предполагает обработку больших данных. В связи с этим, целью проведенного исследования стала разработка масштабируемого алгоритма, который связывает

изменения состава микробного сообщества с изменениями наиболее часто определяемых в инженерной геологии признаков глинистого грунта и помогает выбору оптимальной стратегии его укрепления. Общая структура алгоритма задана спецификой получения и организации геомикробиологических и инженерно-геологических данных. На начальном этапе данные были организованы в реляционную базу, связывающую варианты условий эксперимента, измерения характеристик глины, геохимических параметров порового раствора и микробиомы получившихся сообществ. Путем присвоения категорий неколичественным свойствам и масштабирования количественных свойств глины сгенерированы признаки для дальнейшего анализа. Для полученных признаков рассчитаны первичные статистические зависимости и выбраны индикаторные функции. Неинформативные компоненты базы данных отбрасывали, согласно критерию вклада в дисперсии значений индикаторных функций. Анализ метагеномных данных таксономического профилирования начинали с выделения фоновых доминантов сообщества в контрольных средах и их исключения из рассмотрения. Далее проводили анализ сообществ экспериментальных вариантов. Были определены доминанты сообществ каждого экспериментального варианта; исключены незначимые группы организмов, исходя из критерия обилия менее 5% наиболее обильного доминанта. На следующем этапе было задано ветвление алгоритма на три направления: выделение доминантов, обилия которых скоррелированы между собой, выделение субдоминантов, скоррелированных с доминантами, и выделение доминантов, перспективных с точки зрения изменения свойств глины. Все корреляции между признаками были представлены в виде матриц значимых ранговых корреляций Спирмена. Первые две ветви имели общим итогом формирование условных сообществ и реконструкцию предпочтаемых ими условий. На основании априорных знаний о биохимических циклах, в которых участвуют члены условных сообществ были построены гипотезы о возможностях как изменений глины, так и сукцессии самих сообществ. При этом кластеризация вариантов эксперимента по составу условных сообществ позволила оценить правдоподобие реконструкции условий, поскольку последние были заданы экспериментальными вариантами. Третья ветвь алгоритма предполагала построение корреляционной матрицы обилий доминантов сообществ с абиотическими признаками глины и геохимической обстановки. Эту ветвь продолжал алгоритм градиентного бустинга, с помощью которого была определена значимость отобранных бактериальных родов для уровня каждой индикаторной функции глины, выраженной в тринарном виде. Итогом выполнения третьей ветви алгоритма было выделение биогеотехнологически перспективных бактериальных родов. Объединение результатов выполнения трех ветвей вело к сопоставлению устойчивым микробным консорциумам признаков грунта, реконструкции предпочтительного состава сообщества для укрепления глины. Для реализации описанного алгоритма применяли средства работы с базами данных, предобработки и статистической обработки массивов, классификации, агломеративной кластеризации и визуализации языка программирования Python: Pandas, Numpy, Scipy, Scikit-learn, Matplotlib, Seaborn. Приведенный алгоритм был опробован на образцах карбонатно-глинистого грунта и спелеотем в качестве источника донорского сообщества, свойства которых описаны в авторской работе [2]. Данные образцы экспонировали 4–180 дней в условиях полного влагонасыщения замкнуту циркулирующим потоком при температуре +12 °C в темноте. Вариации условий эксперимента с образцами задавали за счет длительности замачивания, уровней парциального давления CO<sub>2</sub> в атмосфере, внесения стимулирующей среды с донорским сообществом. На вход алгоритма подавали данные об условиях эксперимента, таксономическом разнообразии бактериального сообщества, содержании в глине общего органического углерода, карбонатов, смачиваемости, распределении по фракциям диаметра, агрегированности частиц, пределах пластичности, сжимаемости при одноосном сжатии, а также pH, ОВП и электропроводности порового раствора. В качестве индикаторных функций глины были выбраны признаки: нижний

предел пластичности, агрегированность (по И.М. Горьковой), карбонатность, сжимаемость при компрессионных дренированных испытаниях и смачиваемость (по контактному углу смачивания). Положительными изменениями глины считалось повышение первых трех функций и снижение последних двух.

## Результаты

На начальном этапе в базу данных вошло 21 свойство, характеризующее глину и геохимическую обстановку, и 3166 ASV (вариантов последовательностей ампликонов). Для глубокого анализа изменений глины и сообщества было отобрано 13 абиотических признаков, 39 доминантных и более 80 связанных с ними субдоминантных бактериальных родов. Участие донорского сообщества в формировании вариаций таксономического распределения бактерий в экспериментальных сообществах оценено по воспроизведимости присутствия доминантных представителей бактерий в экспериментальных грунтах после пересева. Из родов донорского сообщества спелеотем наиболее успешно воспроизвелись после пересева на глину *Pseudomonas* и *Paenibacillus*. На основании оценки корреляционных связей определено, что с пересевом пещерного сообщества в глину также связаны рода *Citrifermantans*, *Noviherbaspirillum*, *Stenotrophomonas*, *Sedimentibacter*. Выделено 8 групп доминантов условных сообществ, реконструированы их основные свойства и произведено соотнесение с экспериментальными вариантами. Большинство доминантов имели адаптации к анаэробным или микроаэрофильным условиям, а также колебаниям pH и температуры. В вариантах без донорского сообщества доминировали олиготрофы. В долгосрочных вариантах сокультивирования с донорским сообществом доминировали хемоорганотрофы с возможностью производства биопленки (внеклеточных полимеров); при этом, если в варианте создавалось повышенное парциальное давление CO<sub>2</sub>, то в доминанты выходили те из них, кто был способен использовать различные терминальные акцепторы электронов. Сообщества на основе хемоорганотрофов со способностью к образованию спор, ферментации, производству антибиотиков доминировали в краткосрочных вариантах и долгосрочных вариантах с повышенным содержанием CO<sub>2</sub>.

## Инженерные эффекты и микробиом

Выявлены доминирующие бактериальные рода с максимальным потенциалом полезных изменений свойств глин: *Sediminibacterium*, *Noviherbaspirillum*, *Caenimonas*, *Acinetobacter*, *un\_Mухососсaceae*. Фрагмент корреляционной матрицы инженерных признаков с обилием отдельных бактериальных родов приведен на рисунке 1. По результатам проведенного исследования, разработан алгоритм работы с геомикробиологическими данными, который пригоден для анализа потенциала и прогноза применения биогеотехнологий.

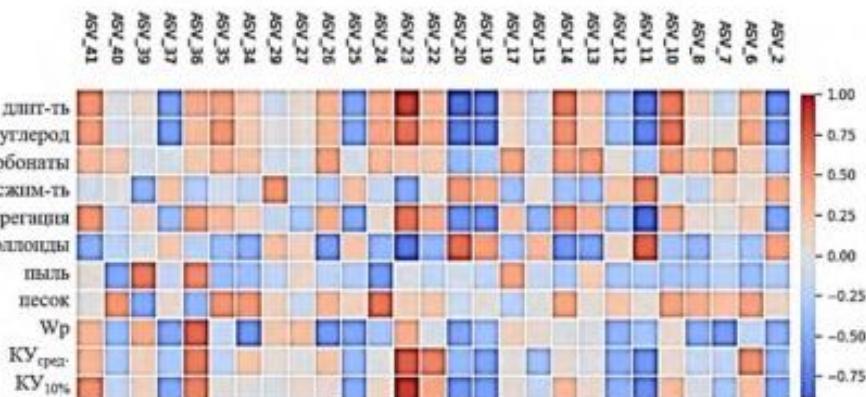


Рис. 1. Корреляционная матрица обилия некоторых бактериальных родов и свойств глины (фрагмент): Обозначения: «длит-ть» – длительность замачивания, «орг.углерод» – общий органический углерод, «сжим-ть» – сжимаемость глины при компрессионном дренированном испытании на одноосное сжатие, «коллоиды», «пыль», «песок» – содержание фракций частиц менее 2 мкм, 2–50 мкм и более 50 мкм, соответственно, «W<sub>p</sub>» – нижний предел пластичности, «КУсред.», «КУ10%» – среднее значение и 10 процентиль контактного угла смачивания частиц глины

#### Литература

1. Kim Y., Kang S., Jo K. et al. Carbonate mineral precipitation induced by microorganisms enriched from the cave water and biofilm in a lime-decorated lava tube // *Scientific Reports*. – 2025. – Vol. 15. – Art. No. 7182.
2. Sofinskaya O.A., Mannapova L.M., Usmanov R.M. et al. Biogeochemical interface development in a model carbonate-clayey soil // *Environmental Earth Sciences*. – 2024. – Vol. 83. – P. 6.