WYDZIAŁ BIOLOGII
i OCHRONY ŚRODOWISKA
Uniwersytet Łódzki

InterDOC-STARt
Interdyscyplinarne Studia Doktoranckie

Jakub Lach

Charakterystyka bioróżnorodności i potencjału biotechnologicznego mikroorganizmów halofilnych

Characteristics of biodiversity and biotechnological potential of halophilic microorganisms

Praca doktorska

wykonana w Katedrze Mikrobiologii Molekularnej
Instytutu Mikrobiologii, Biotechnologii i Immunologii

Promotor:

- **Dr hab. Paweł Stączek, prof. UŁ**
  z Katedry Mikrobiologii Molekularnej
  Uniwersytetu Łódzkiego

Promotor pomocniczy:

- **Dr Agnieszka Matera-Witkiewicz**
  z Pracowni Przesiewowych Testów Aktywności Biologicznej i Gromadzenia Materiału Biologicznego
  Uniwersytetu Medycznego im. Piastów Śląskich we Wrocławiu

Łódź, 2023

Składam serdeczne podziękowania dla ***Pana dr. hab. Pawła Stączka, prof. UŁ****, oraz dla* ***Pani dr Agnieszki Matery-Witkiewicz*** *za nieocenioną pomoc, opiekę merytoryczną oraz cierpliwość okazaną w okresie przygotowania niniejszej pracy doktorskiej.*

Dziękuję również ***Panu dr. Dominikowi Strapagielowi, prof. UŁ*** *za wieloletnią współpracę, możliwość wzięcia udziału w licznych projektach naukowych i wsparcie w rozwoju warsztatu badawczego.*

Szczególne podziękowania należą się również moim współpracownikom z Pracowni Biobank oraz Katedry Mikrobiologii Molekularnej Uniwersytetu Łódzkiego za wspaniałą atmosferę w pracy i ogromne pokłady dobrej energii.

Spis treści

Źródła finansowania i współpraca

Praca doktorska została przygotowana w ramach projektu "InterDOC-STARt – Interdyscyplinarne Studia Doktoranckie na Wydziale BiOŚ UŁ" - Program Operacyjny Wiedza Edukacja Rozwój 2014-2020, Oś priorytetowa III. Szkolnictwo wyższe dla gospodarki i rozwoju, Działanie 3.2 Studia doktoranckie. Nr projektu: POWR.03.02.00- IP.08-00-DOK/16. Realizowany w latach 2018-2022. Kierownik – prof. dr hab. Agnieszka Marczak.

Praca doktorska powstała we współpracy z Pracownią Biobank Uniwersytetu Łódzkiego oraz Pracownią Przesiewowych Testów Aktywności Biologicznej i Gromadzenia Materiału Biologicznego/Biobankiem UMW Uniwersytetu Medycznego im. Piastów Śląskich we Wrocławiu.

Dorobek naukowy

Publikacje wchodzące w skład rozprawy doktorskiej

Artykuły opublikowane:

Lach J., Jęcz P., Strapagiel D., Matera-Witkiewicz A., Stączek P. (2021) The Methods of Digging for “Gold” within the Salt: Characterization of Halophilic Prokaryotes and Identification of Their Valuable Biological Products Using Sequencing and Genome Mining Tools. *Genes* 2021, DOI: 10.3390/genes12111756

IF: 4.414, MEiN: 100 pkt.

Lach J., Królikowska K., Baranowska M., Krupińska M., Strapagiel D., Matera-Witkiewicz A., Stączek P. A First Insight into the Polish Bochnia Salt Mine Metagenome. *Environ Sci Pollut Res* 2023, DOI: 10.1007/s11356-023-25770-7

IF: 5.190, MEiN: 100 pkt.

Artykuły zaakceptowane, oczekujące na publikację:

Lach J., Strapagiel D., Matera-Witkiewicz A., Stączek P. Draft Genomes of Halophilic Chromohalobacter and Halomonas Strains Isolated from Brines of The Carpathian Foreland, Poland. *Journal of Genomics* 2023, DOI: 10.7150/jgen.80829

MEiN: 100 pkt.

Lach J., Strapagiel D., Matera-Witkiewicz A., Stączek P. Draft Genomes of Halophilic Archaea Strains Isolated from Brines of the Carpathian Foreland, Poland. *Journal of Genomics* 2023, DOI: 10.7150/jgen.82493

MEiN: 100 pkt.

Manuskrypt w recenzji:

Lach J., Krupińska M., Mikołajczyk A., Strapagiel D., Stączek P., Matera-Witkiewicz A. Novel AMP from saline environments - promising glimmer for inhibition of multidrug resistant E. faecalis and S. aureus infections?. *Eur J Pharm Biopharm* 2023

IF: 5.589, MEiN: 100 pkt.

Suma: IF = 15,193; MEiN = 500 pkt.

Pozostały dorobek naukowy

Borówka, P., **Lach, J.**, Bakuła, Z., van Ingen, J., Safianowska, A., Brzostek, A., Dziadek, J., Strapagiel, D., Jagielski, T. Draft Genome Sequences of Mycobacterium kansasii Clinical Strains. *Genome Announc.* 2017, DOI: 10.1128/genomeA.00406-17

MEiN: 70 pkt.

Jarczak J., Grochowalski Ł., Marciniak B., **Lach J.**, Słomka M., Sobalska-Kwapis M., Lorkiewicz W., Pułaski Ł., Strapagiel D. Mitochondrial DNA variability of the Polish population. *European Journal of Human Genetics* 2019, DOI: 10.1038/s41431-019-0381-x

IF: 5.351, MEiN: 100 pkt.

Brodzik K., Krysztopa-Grzybowska K., Polak M., **Lach J.**, Strapagiel D., And Zasada AA.; Analysis of the Amino Acid Sequence Variation of the 67–72p Protein and the Structural Pili Proteins of Corynebacterium diphtheria for their Suitability as Potential Vaccine Antigens. *Polish Journal of Microbiology* 2019, DOI: 10.33073/pjm-2019-025

IF: 2.019, MEiN: 40 pkt.

Zawilak-Pawlik A., Zarzecka U., Żyła-Uklejewicz D., **Lach J.**, Strapagiel D., Tegtmeyer N., Böhm M., Backert S., Skorko-Glonek J. Establishment of serine protease htrA mutants in Helicobacter pylori is associated with secA mutations. *Scientific Reports* 2019, DOI: 10.1038/s41598-019-48030-6

IF: 4.996, MEiN: 140 pkt.

Jarczak, J., **Lach, J.**, Borówka, P., Gałka, M., Bućko, M., Marciniak, B., Strapagiel, D. BioSCOOP–Biobank Sample Communication Protocol. New approach for the transfer of information between biobanks. *Database - The Journal of Biological Databases and Curation* 2019, DOI: 10.1093/database/baz105

IF: 4.462, MEiN: 100 pkt.

Jagielski T., Borówka P., Bakuła Z., **Lach J.**, Marciniak B., Brzostek A., Dziadek J., Dziurzyński M., Pennings L., van Ingen J., Žolnir-Dovč M., Strapagiel D. Genomic Insights into the Mycobacterium kansasii Complex: An Update. *Frontiers in Microbiology* 2020, DOI: 10.3389/fmicb.2019.02918

IF: 6.064, MEiN: 100 pkt.

Minias A., Żukowska L., **Lach J.**, Jagielski T., Strapagiel D., Kim S.Y., Koh W.J., Adam H., Bittner R., Truden S., Žolnir M., Dziadek, J. Subspecies-specific sequence detection for differentiation of Mycobacterium abscessus complex. *Scientific Reports* 2020, DOI: 10.1038/s41598-020-73607-x

IF: 4.996, MEiN: 140 pkt.

Majak P., Molińska K., Latek M., Rychlik B., Wachulec M., Błauż A., Budniok A., Gruchała M., **Lach J.**, Sobalska-Kwapis M., Baranowska M., Królikowska K., Strapagiel D., Majak J., Czech D., Pałczy C., Kuna P. Upper-airway dysbiosis related to frequent sweets consumption increases the risk of asthma in children with chronic rhinosinusitis. *Pediatric Allergy and Immunology* 2021, DOI: 10.1111/pai.13417

IF: 5.644, MEiN: 100 pkt.

Brzostek A., Gąsior F., **Lach J.**, Żukowska L., Lechowicz E., Korycka-Machała M., Strapagiel D., Dziadek J. ATP-Dependent Ligases and AEP Primases Affect the Profile and Frequency of Mutations in Mycobacteria under Oxidative Stress. *Genes* 2021, DOI: 10.3390/genes12040547

IF: 4.141, MEiN: 100 pkt.

Wajda A., Walczuk E., Stypińska B., **Lach J.**, Yermakovich D., Sivitskaya L., Romanowska-Próchnicka K., Wysocki T., Jarończyk M., Paradowska-Gorycka A. AHR-dependent genes and response to MTX therapy in rheumatoid arthritis patients. *Pharmacogenomics Journal* 2021, DOI: 10.1038/s41397-021-00238-4

IF: 3.245, MEiN: 140 pkt.

Bekier A., Kawka M., **Lach J.**, Dziadek J., Paneth A., Gatkowska J., Dzitko K., Dziadek B. Imidazole-thiosemicarbazide derivatives as potent anti-Mycobacterium tuberculosis compounds with antibiofilm activity. *Cells* 2021, DOI: 10.3390/cells10123476

IF: 7.666, MEiN: 140 pkt.

Molińska K., Latek M., Rychlik B., **Lach J.**, Strapagiel D., Majak J., Błażowski Ł., Jerzyńska J., Kuna P., Majak P. House dust mite sensitization and frequent antibiotic courses may suppress remission of rhinosinusitis and asthma symptoms in young children. *Allergy* 2022, DOI: 10.1111/all.15085

IF: 14.710, MEiN: 140 pkt.

Korycka-Machała M., Kawka M., **Lach J.**, Płocińska R., Bekier A., Dziadek B., Brzostek A., Płociński P., Strapagiel D., Szczesio M., Gobis K., Dziadek J. 2,4-Disubstituted pyridine derivatives are effective against intracellular and biofilm-forming tubercle bacilli. *Frontiers in Pharmacology* 2022, DOI: 10.3389/fphar.2022.1004632

IF: 5.988, MEiN: 100 pkt.

Latek M., Łacwik P., Molińska K., Błauż A., **Lach J.**, Rychlik B., Strapagiel D., Majak J., Molińska J., Czech D., Seweryn M., Kuna P., Pałczyński C., Majak P. Effect of an intranasal corticosteroid on quality of life and local microbiome in young children with chronic rhinosinusitis: a randomized clinical trial. *JAMA Pediatrics* 2023, DOI: 10.1001/jamapediatrics.2022.6172

IF: 26.796, MEiN: 200 pkt.

Suma: IF = 96.078, MEiN = 1610

Komunikaty zjazdowe

Lach J., Marciniak B., Borówka P., Strapagiel D. Species classification of Mycobacterium avium and Mycobacterium intracellulare based on whole genome analysis. 14th Horizons in Molecular Biology Symposium, 11-15.09.2017, Göttingen, Niemcy

Lach J., Marciniak B., Strapagiel D. Zarządzanie bezpieczeństwem danych w polskich biobankach. Krajowa Naukowo-Szkoleniowa Konferencja Biobanków Polskich, 13-14.10.2017, Wrocław

Sobalska-Kwapis M., **Lach J.**, Królikowska K., Grochowalski Ł., Marciniak B., Strapagiel D.; Metagenomics as a complex analysis in Biobank Lab. Polish Scientific Networks, 21.06.2018, Łódź

Jarczak J., **Lach J.**, Borówka P., Gałka M., Marciniak B., Dominik Strapagiel D.; BioSCOOP – protokół komunikacyjny informacji opisującej próbki w biobankach. II Krajowa Naukowo-Szkoleniowa Konferencja Biobanków Polskich, 11-12.10.2018, Wrocław

Strapagiel D., Marciniak B., **Lach J.**, Topolski P.; Bezpieczeństwo informacji – podsumowanie audytów wewnątrz Polskiej Sieci Biobanków. II Krajowa Naukowo-Szkoleniowa Konferencja Biobanków Polskich, 11-12.10.2018, Wrocław

Lach J., Strapagiel D. Sekwencjonowanie de novo genomów bakteryjnych z zastosowaniem długich odczytów – różne podejścia, II Wrocławska Konferencja Sekwencjonowania Następnej Generacji, 26.10.2018, Wrocław

Janik-Superson K., Hejduk J., **Lach J.**, Strapagiel D.; Sekwencjonowanie NGS mikrobiomu guano Myotis myotis - badanie pilotażowe. XXVII Ogólnopolska Konferencja Chiropterologiczna, 16-18.11.2018, Białowieża

Knot A., Chrabąszcz M., **Lach J.**, Sikora M.; Sexual functioning in female patients with systemic sclerosis – single centre study. 15th Warsaw International Medical Congress, 9-12.05.2019, Warszawa

Lach J., Marciniak B., Walczak W., Strapagiel D.; BioFace - narzędzie do wymiany informacji na temat próbek w biobankach. III Krajowa Naukowo-Szkoleniowa Konferencja Biobanków Polskich, 06-08.11.2019, Łódź

Słomka M., Jarczak J., Grochowalski Ł., Marciniak B., **Lach J.**, Sobalska-Kwapis M., Lorkiewicz W., Pułaski Ł., Strapagiel D. Genetic diversity of Polish population based on the analysis of mitochondrial DNA. III Krajowa Naukowo-Szkoleniowa Konferencja Biobanków Polskich, 06-08.11.2019, Łódź

Strapagiel D., Marciniak B., Jarczak J., **Lach J.**; Distributed IT systems within Polish Biobanking Network, III Krajowa Naukowo-Szkoleniowa Konferencja Biobanków Polskich, 06-08.11.2019, Łódź

Lach J., Królikowska K., Strapagiel D., Matera-Witkiewicz A., Stączek P.; Graduation towers microbiome – biobanking and analyzing data from specific microbial environment. Europe Biobank Week 2020, 17-20.11.2020, online

Janik-Superson K., Krawczyk D., Baranowska M., Tończyk G., **Lach J.**, Seweryn M., Królikowska K., Strapagiel D., Bącela-Spychalska K. Optimalization of the ponds eDNA procedures for research prokaryotes and eukaryotes biodiversity. International Conference on DNA Barcoding and Biodiversity 2022, 25-27.05.2022, Sofia, Bulgaria

Baranowska M, Krawczyk D., Janik-Superson K., Królikowska K., Seweryn M., **Lach J.**, Tończyk G., Taugbøl A., Strand D., Brodecki J., Strapagiel D., Bącela-Spychalska K. How does urbanisation impact pond biodiversity in a geographical gradient? International Conference on DNA Barcoding and Biodiversity 2022, 25-27.05.2022, Sofia, Bulgaria

Janik-Superson K., Burlińska K., Gottfried I., Gottfried T., Wojtowicz B., **Lach J.**, Strapagiel D. A Comparison of the Diet Specialization of Two Forest-dwelling Bat Species: *Barbastella barbastellus* and *Plecotus auratus*. 19th International Bat Research Conference, 07-12.08.2022, Austin, USA

Lach J., Strapagiel D., Matera-Witkiewicz A., Stączek P. Nowe genomy mikroorganizmów halofilnych wyizolowanych ze źródeł solankowych w południowej Polsce. XXIX Ogólnopolski Zjazd Polskiego Towarzystwa Mikrobiologów, 15-17.09.2022, Warszawa

Lach J., Królikowska K., Baranowska M., Krupińska M., Strapagiel D., Agnieszka Matera-Witkiewicz A., Stączek P. Salt mine metagenomics – biodiversity characterization and novel natural products identification. European Biotechnology Congress 2022, 5-7.10.2022, Praga, Czechy

Staże naukowe

W trakcie przygotowywania pracy doktorskiej w ramach projektu "InterDOC-STARt – Interdyscyplinarne Studia Doktoranckie na Wydziale BiOŚ UŁ zostały zrealizowane dwa staże naukowe.

W terminie 02.04.2022 - 30.06.2022 zrealizowano staż naukowy pod opieką Profesora Francisco Eduardo Rodriguez Valera w Grupie Badawczej Genomiki Ewolucyjnej, Katedrze Nauk o Roślinach i Mikrobiologii Uniwersytetu Miguela Hernandeza w Elche, Hiszpania. Staż został sfinansowany w ramach projektu InterDOC-STARt – Interdyscyplinarne Studia Doktoranckie na Wydziale BiOŚ UŁ.

W ramach stażu zapoznano się ze szlakami analiz bioinformatycznych stosowanymi w odwiedzanym zespole badawczym. Zapoznano się z metodami tworzenia złożeń genomów z metagenomów (MAGs), oceny jakości otrzymanych MAGs i ich adnotacji. Dodatkowo zaznajomiono się także z narzędziami bioinformatycznymi wytworzonymi w odwiedzanym zespole takimi jak Flex2 czy RaFAH i rozwinięto umiejętności w zakresie ich zastosowania.

W terminie 17.10.2022 – 28.10.2022 zrealizowano staż naukowy na Wydziale Farmaceutycznym Uniwersytetu Medycznego im. Piastów Śląskich we Wrocławiu w Pracowni Przesiewowych Testów Aktywności Biologicznej i Gromadzenia Materiału Biologicznego/Biobank UMW.

Staż posłużył do wykonania eksperymentów związanych z oceną aktywności przeciwdrobnoustrojowej peptydów zidentyfikowanych w ramach wcześniejszych analiz bioinformatycznych. W trakcie stażu poznano metody stosowane w tym zakresie oraz opracowano otrzymane wyniki. Ponadto przygotowano zarys artykułu dotyczącego badanych peptydów.

Wstęp

Środowiska zasolone to jedne z najbardziej rozpowszechnionych ekosystemów na świecie. Prym w tym względzie wiodą środowiska morskie zajmujące ponad 70% powierzchni ziemi, których średnie zasolenie wynosi około 3,5% NaCl. Habitaty te charakteryzuje także ogromna bioróżnorodność (Stat et al., 2017; Baker et al., 2021; Ward et al., 2022). Istnieją jednak siedliska cechujące się wyższym stopniem zasolenia niż środowiska morskie. Można do nich zaliczyć na przykład słone jeziora, takie jak Morze Martwe na pograniczu Izraela i Jordanii, czy też Wielkie Jezioro Słone w USA, gdzie woda może osiągać ponad 20% zasolenie. Jednym z najbardziej zasolonych tego typu habitatów jest jezioro Assale w Etiopii, gdzie zasolenie potrafi przekroczyć 40% NaCl (Cavalazzi et al., 2019). Innymi przykładami środowisk o wysokim zasoleniu mogą być miejsca stworzone przez człowieka w celu pozyskiwania soli, takie jak saliny czy kopalnie soli kamiennej. W przypadku takich siedlisk często dochodzi do niemal całkowitego wysycenia obecnych tam solanek, co tworzy bardzo wymagające warunki do przetrwania dla organizmów żywych (Fernández et al., 2014; Cycil et al., 2020). Poza wysokim zasoleniem, w wielu przypadkach środowiska te charakteryzują się również innymi niesprzyjającymi dla życia parametrami fizykochemicznymi, co sprawia, że bytujące tam organizmy można zaliczyć do poliekstremofilów. Przykładem takiego siedliska może być obszar Dallol w Etiopii. Woda w kolorowych stawach znajdująca się w tym miejscu charakteryzuje się bardzo niską aktywnością (rozumianą jako stosunek ciśnienia pary wodnej nad powierzchnią solanki do ciśnienia pary wodnej nad powierzchnią czystej wody w tej samej temperaturze i przy tym samym ciśnieniu całkowitym), pH zbliżonym do zera oraz niemal całkowitym wysyceniem solą. Warunki w tych zbiornikach skrajnie nie sprzyjają zasiedleniu przez organizmy żywe, jednak pomimo tego udaje się w nich wykryć obecność mikroorganizmów, które z racji na swoją tolerancję wobec wysokiego zasolenia nazywamy halofilami (Belilla et al., 2019; Gómez et al., 2019).

Halofile są to organizmy, dla których obecność soli w środowisku jest niezbędna do wzrostu i przetrwania. Do tej grupy zaliczyć można przedstawicieli wszystkich trzech domen życia (Oren, 2011). W zależności od przynależności filogenetycznej i sposobu pozyskiwania energii halofile można podzielić na heterotroficzne, fototroficzne i metanogenne Archaea, heterotroficzne, litotroficzne i fototroficzne Bakterie oraz heterotroficzne i fototroficzne Eukarionty (Oren, 2011). Ze względu na duże zróżnicowanie oraz dominującą rolę mikroorganizmów prokariotycznych w środowiskach o wysokim zasoleniu, w przedstawionej pracy skupiono się właśnie na tej grupie.

Biorąc pod uwagę wymagania wobec poziomu zasolenia koniecznego dla zapewnienia optymalnego wzrostu, halofile można podzielić na halofile słabe, które optimum wzrostu osiągają przy zasoleniu w przedziale 1-6% NaCl, halofile umiarkowane, dla których optymalne warunki to zasolenie na poziomie 6-15% NaCl oraz halofile skrajne, które do optymalnego wzrostu wymagają ponad 15% NaCl w środowisku (Ventosa and Arahal, 2009). Poza halofilami istnieje także

inna grupa mikroorganizmów zdolna do życia w środowisku, w którym obecne są znaczne ilości soli, nazywana halotolerantami. Jak sama nazwa wskazuje, mikroorganizmy te charakteryzują się tolerancją na zasolenie, jednak obecność soli w środowisku nie jest wymagana do ich optymalnego wzrostu. W przypadku mikroorganizmów, których nie jesteśmy w stanie hodować w kontrolowanych warunkach, często bardzo ciężko jest jednoznacznie stwierdzić czy dany mikroorganizm jest halofilem czy halotolerantem (Ventosa and Arahal, 2009).

Ważnym czynnikiem niesprzyjającym przetrwaniu organizmów żywych w środowiskach zasolonych jest wysokie ciśnienie osmotyczne, związane z dużą ilością substancji rozpuszczonych. W związku z tym, jednym z głównych przystosowań mikroorganizmów bytujących w takich środowiskach do przetrwania, jest wytworzenie mechanizmów pozwalających na przeciwdziałanie nadmiernej osmozie. Można wyróżnić dwie główne strategie mikroorganizmów w tym obszarze: „salt-out" oraz „salt-in" (Gunde-Cimerman et al., 2018). Pierwsza z nich wykorzystywana jest głównie przez bakterie i polega na biosyntezie lub absorbcji ze środowiska substancji o charakterze osmolitów lub osmoprotektantów. Substancje te pozwalają na wyrównanie ciśnienia osmotycznego tworzonego przez wysoką zawartość soli w środowisku i pozwalają ochronić komórkę przed jego wpływem. Do substancji wykorzystywanych w tej strategii zaliczyć można poliole, cukry, betainę czy też ektoinę. W strategii „salt-in" natomiast, jako czynnik wyrównujący ciśnienie osmotyczne, komórki akumulują sól. W większości przypadków jest to chlorek potasu. Strategia ta jest wykorzystywana głównie przez halofile skrajne należące do Archaea oraz niektórych przedstawicieli Eubakterii (Oren, 2011; Gunde-Cimerman et al., 2018). Jak wspomniano wcześniej, wysokie zasolenie często nie jest jedynym niesprzyjającym czynnikiem w środowisku bytowania mikroorganizmów halofilnych. W wielu przypadkach muszą one sobie radzić z wysokim poziomem promieniowania UV, niską zawartością tlenu, wysokim ciśnieniem, skrajnymi temperaturami i pH, czy też obecnością metali ciężkich lub substancji toksycznych. W związku z tym halofile wytworzyły wiele szlaków metabolicznych pozwalających im przetrwać w radykalnie nieprzyjaznym środowisku. Ta właściwość determinuje ich znaczącą wartość dla przemysłu biotechnologicznego (Yin et al., 2015; Amoozegar et al., 2017; Dutta and Bandopadhyay, 2022). Z powodu ograniczonej dostępności substancji odżywczych w środowisku, u części z halofili pojawiła się zdolność do wytwarzania substancji o działaniu przeciwdrobnoustrojowym, które mogą hamować wzrost pokrewnych gatunków mikroorganizmów. Dzięki temu są one w stanie skuteczniej konkurować o zasoby odżywcze. Do istotnych i charakterystycznych substancji należących do tej grupy zaliczyć można halocyny (Kumar et al., 2021).

Interesującym przykładem środowiska o wysokim zasoleniu jest Kopalnia Soli Bochnia, jedna z najstarszych kopalni soli w Europie mająca bezdyskusyjną wartość historyczną. Została ona umieszczona na Liście Światowego Dziedzictwa UNESCO (UNESCO, 2023). Obecnie w kopalni znajduje się dziewięć poziomów sięgających nawet na 350 m poniżej poziomu wejścia, które niegdyś służyły za chodniki górnicze. Znaczna część dawnych wyrobisk jest

otwarta dla turystów i kuracjuszy korzystających z leczniczych właściwości komór solnych (Puławska et al., 2021). Mikrobiom związany z kopalnią nie był dotychczas przedmiotem wyczerpujących badań, co determinuje jego potencjał naukowo-badawczy. Przesłanką ku temu może być fakt zidentyfikowania przed kilku laty nowego gatunku halofilnego Archaea, nazwanego *Halorhabdus rudnickae,* w próbce pochodzącej z technicznego ujęcia solanki, położonej niedaleko, nieczynnej kopalni soli w Baryczy (Albuquerque et al., 2016a).

Jednymi z najczęściej identyfikowanych bakterii w środowiskach zasolonych są mikroorganizmy należący do typów Actinobacteria, Bacteroidetes, Cyanobacteria, Proteobacteria i Firmicutes. Bakterie często dominują w środowiskach o przeciętnym i niskim poziomie zasolenia, jednak wraz ze wzrostem ilości soli w środowisku często też rośnie udział Archaea. Wśród Archaea najczęściej można wykryć te należące do typu Euryarchaeota, w szczególności do klasy Halobacteria oraz Methanococci (Oren, 2008, 2020).

Dzięki rozwojowi metod związanych z sekwencjonowaniem DNA środowiskowego, coraz dokładniej poznawane są także taksony halofili, dla których nie ma obecnie opracowanych metod hodowli w warunkach laboratoryjnych. Do tej grupy zaliczyć można niektóre Archaea należące do supertypu DPANN. Nazwa ta jest akronimem stworzonym od liter nazw pięciu pierwszych typów mikroorganizmów odkrytych w tym taksonie, czyli Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota i Nanohaloarchaeota (Rinke et al., 2013). Mikroorganizmy z tej grupy charakteryzują się mniejszym rozmiarem komórek niż inne Archaea i nadal są niezbyt dobrze poznane (Pal et al., 2020). Innym niedawno odkrytym, dzięki wykorzystaniu metod molekularnych, supertypem Archaea, w którym można znaleźć mikroorganizmy halofilne jest Asgardarchaeota zawierający typy Lokiarchaeota, Thorarchaeota, Odinarchaeota oraz Heimdallarchaeota. Ze względu na charakterystykę genomiczną grupa ta jest traktowana ,jako jeden z kluczowych etapów przejściowych między mikroorganizmami Prokariotycznymi a Eukariotycznymi. Wiąże się to głównie z obecnością w ich genomie homologów kluczowych genów eukariotycznych takich jak Sec23/24 czy domeny TRAPP (Zaremba-Niedzwiedzka et al., 2017).

Biorąc pod uwagę charakterystykę tych mikroorganizmów, halofile zyskują coraz bardziej na znaczeniu w przemyśle biotechnologicznym. Jednym z podstawowych czynników wpływających na zainteresowanie sektora biotechnologicznego jest fakt, że produkcję z użyciem halofili można prowadzić w warunkach wysokiego zasolenia co sprawia, że trudniej jest doprowadzić do kontaminacji hodowli, finalnie ograniczając ryzyko i koszt produkcji. Ponadto, dzięki rozwojowi narzędzi biologii molekularnej i inżynierii genetycznej, możliwa jest skuteczna manipulacja szczepami produkcyjnymi w celu ich dostosowania do potrzeb procesu (Liu et al., 2019). Poza samym walorem procesowym, halofile są także producentami wielu cennych bioproduktów. Zaliczyć do nich można ektoinę i hydroksyektoinę, które są osmoprotektantami. Ponadto substancje te wykazują działanie ochronne wobec makrocząsteczek, komórek i tkanek. Ektoina znalazła zastosowanie w przemyśle kosmetycznym, ale prowadzone są również

badania mające na celu określenie ich skuteczności i potencjalnego zastosowania w chorobach oczu czy układu oddechowego (Pastor et al., 2010; Bilstein et al., 2021b, 2021a). Innym istotnym bioproduktem wytwarzanym przez halofile są polihydroksyalkaniany (PHA), czyli biodegradowalne i biokompatybilne poliestry, akumulowane przez wiele z mikroorganizmów halofilnych. W przypadku tych substancji rozważa się ich wykorzystanie do produkcji bioplastiku czy też biopaliw. Jednym z najlepiej przebadanych polimerów z tej grupy jest polihydroksymaślan (PHB) oraz poli(kwas hydroksymasłowy-co-hydroksywalerianowy) (PHBV) (Yin et al., 2015). Ponadto halofile są producentami wielu ekstremozymów, czyli enzymów aktywnych w warunkach przekraczających tolerancję klasycznych enzymów. Enzymy takie, jak chociażby proteazy stabilne w warunkach wysokiego zasolenia, uzyskiwane z *Halococcus agarilyticus* czy też *Haloferax mediteranei,* są istotne dla przemysłu spożywczego i mleczarskiego. Lipazy pozyskiwane ze szczepów halofilnych mogą natomiast znaleźć zastosowanie w przemyśle tekstylnym i farmaceutycznym, a także w innych gałęziach przemysłu jako detergenty. Ksylenazy znalazły natomiast zastosowanie w przemyśle papierniczym (Dutta and Bandopadhyay, 2022). Inną ważną grupą bioproduktów związanych z halofilami są barwniki. Produkowane są one w celu zapewnienia dodatkowej ochrony przed czynnikami środowiskowymi. Produkcja karotenoidów przez algi z rodzaju *Deinococcus* sp. jest obecnie jednym z najistotniejszych sukcesów związanych z wykorzystaniem mikroorganizmów halofilnych w przemyśle. Substancje te w wielu przypadkach działają jako skuteczne antyoksydanty (Liu et al., 2019). Ponadto rozważa się wykorzystanie halofili do produkcji biosurfaktantów, które są niezwykle wartościowym choć nadal bardzo drogim w produkcji bioproduktem (Schultz and Rosado, 2020).

Poza potencjałem przemysłowym halofile mają też potencjał w zakresie wytwarzania nowych bioproduktów o działaniu przeciwdrobnoustrojowym i przeciwnowotworowym. Przykładem takich bioproduktów mogą być peptydy przeciwdrobnoustrojowe produkowane przez halofilne Archaea, nazywane halocynami. Do najlepiej scharakteryzowanych na poziomie molekularnym można zaliczyć halocyny H4, C8 i S8 (Kumar et al., 2021). W większości przypadków badacze rozpatrują halocyny jako istotny czynnik w konkurencji z innymi mikroorganizmami produkującymi związki o właściwościach AMP (*antimicrobial peptides*). To co należy podkreślić, to fakt, iż niezwykle rzadko aktywność halocyn jest testowana względem szczepów istotnych z klinicznego punktu widzenia. Ponadto część bioproduktów halofili testowana jest w roli czynników przeciwnowotworowych i w tym obszarze osiągane są pewne postępy. Za przykład mogą posłużyć w tym wypadku karotenoidy, które są najlepiej przebadanym bioproduktem halofilii pod kątem aktywności przeciwnowotworowej. Zarówno ekstrakty karotenoidów pozyskiwane z halofilnych Euakterii jak i Archaea wykazują obiecujące działanie przeciwnowotworowe. Najprawdopodobniej jest to związane z ich aktywnością przeciwutleniającą. Przykładem bakteryjnego ekstraktu karotenoidów może być ekstrakt z *Kocuria* sp. QWT-12, który redukuje żywotność linii komórkowych MCF-7, MDA-MB-468 i MDA-MB-231. Natomiast za przykład dla halofilnych

Archaea może posłużyć ekstrakt karotenoidów z *Halobacterium halobium*, który wykazuje aktywność przeciw linii ludzkiego wątrobiaka HepG2 (Corral et al., 2020).

Cel pracy

Celem prezentowanej pracy doktorskiej było scharakteryzowanie bioróżnorodności mikrobiomów wybranych środowisk o wysokim zasoleniu oraz ocena potencjału biotechnologicznego bytujących tam mikroorganizmów w oparciu o dane genomiczne i metagenomiczne, ze szczególnym uwzględnieniem właściwości przeciwdrobnoustrojowych.

W ramach niniejszej pracy postawiono następujące cele szczegółowe:

1. Scharakteryzowanie bioróżnorodności mikrobiomu Kopalni Soli Bochnia w oparciu o dane z sekwencjonowania wygenerowanych amplikonów fragmentów genów kodujących 16S rRNA oraz sekwencjonowania metagenomów w zebranych próbach środowiskowych;
2. Identyfikacja klastrów genów związanych z biosyntezą wtórnych metabolitów oraz genów kodujących peptydy przeciwdrobnoustrojowe w metagenomach z Kopalni Soli Bochnia;
3. Zsekwencjonowanie oraz scharakteryzowanie genomów szczepów mikroorganizmów wyizolowanych z solanek pobranych na obszarze Pogórza Karpackiego;
4. Wytypowanie w oparciu o analizę bioinformatyczną metagenomów z Kopalni Soli Bochnia oraz miejskich tężni solankowych w Łodzi sekwencji genów kodujących peptydy przeciwdrobnoustrojowe o największym prawdopodobieństwie aktywności przeciwdrobnoustrojowej dla celów syntezy i badań *in vitro*;
5. Określenie aktywności przeciwdrobnoustrojowej syntetycznych peptydów na wybranej grupie wielolekoopornych szczepów patogennych, określenie ich toksyczności względem komórek ludzkich oraz zamodelowanie potencjalnego mechanizmu działania za pomocą metod *in silico*.

Materiały i metody

Pozyskanie prób środowiskowych z Kopalni Soli Bochnia i ich przetwarzanie

Część próbek środowiskowych wykorzystanych w prezentowanej pracy została pobrana z Kopalni Soli Bochnia (49°58′09″N 20°25′03″E) znajdującej się na obszarze Pogórza Karpackiego nieopodal Krakowa. Próbki zostały pobrane z siedmiu miejsc gromadzenia się solanki (tzw. rząpi), znajdujących się na terenie kopalni. Solanka w tych zbiornikach charakteryzowała się niemal całkowitym wysyceniem NaCl, przy zasoleniu w zakresie od 31.3 do 33.3% NaCl. Odczyn solanek był zbliżony do neutralnego i mieścił się w granicach od 6.6 do 7.4. Większe różnice pomiędzy stanowiskami można było zauważyć w zakresie stężeń jonów. Najwyższym poziomem mineralizacji charakteryzowały się solanki ze stanowisk POP27 i RA6. Zauważono tendencje, że solanki ze stanowisk znajdujących się na pierwszym, najwyższym poziomie, kopalni charakteryzowały się mniejszym poziomem mineralizacji, niż solanki z pozostałych dwóch poziomów, z których pobrano próby. Natomiast najwyższe stężenie jonów SO_4^{2-} wynoszące 111 g/l zidentyfikowano w solance ze stanowiska POP27 znajdującego się na trzecim poziomie kopalni.

W Kopalni Soli Bochnia pobrano 24 próbki wody ze zbiorników solanki znajdujących się na trzech poziomach kopalni. Próbki te były pobierane dwukrotnie, pierwszy raz pod koniec 2017 roku, kiedy to pozyskano łącznie 17 próbek: 10 z poziomu pierwszego kopalni (Danielowiec), 4 z poziomu trzeciego (Wernier) i 3 z poziomu czwartego (August); a drugi raz na początku 2019 r., gdy zebrano siedem próbek: 4 z poziomu pierwszego i 3 z poziomu czwartego. Próbki te zostały wykorzystywane do późniejszego sekwencjonowania wygenerowanych amplikonów fragmentów genów kodujących 16S rRNA. W 2017 roku, natomiast w 2019 roku Ponadto do sekwencjonowania metagenomów zostało wybrane po jednej próbce z pierwszego, drugiego, trzeciego i szóstego poziomu kopalni. Miejsca, z których zostały pobrane próby wskazywane były przez przedstawicieli zarządcy kopalni i ze względów bezpieczeństwa wynikających z aktualnych warunków w kopalni, część stanowisk nie zawsze była dostępna.

Próbki solanki były zbierane do sterylnych 50 ml probówek typu Falcon i były przechowywane w 4°C. W celu przygotowania do izolacji DNA, 2 ml porcje solanki były przenoszone do sterylnych probówek typu Eppendorf i wirowane przez 10 min przy 14 000 × g. Po wirowaniu supernatant był usuwany z probówki, dodawano do niej kolejną porcję solanki i powtarzano wirowanie. Dla każdej z próbek procedurę powtarzano pięciokrotnie w celu zagęszczenia biomasy. Następnie przeprowadzono izolację DNA za pomocą zestawu PowerSoil DNA Isolation Kit (Mo Bio Laboratories Inc., Carlsbad, CA, USA).

Pozyskanie prób środowiskowych z tężni solankowych i ich przetwarzanie

W celu identyfikacji genów kodujących peptydy przeciwdrobnoustrojowe poddano analizie cztery próbki środowiskowe pochodzące z miejskich tężni solankowych zlokalizowanych w Łodzi. Po jednej próbce w postaci fragmentów gałązek tarniny pochodziło z tężni solankowych zlokalizowanych w Parku Podolskim (51°74′14″N 19°49′17″E), na osiedlu Botanik (51°75′02″N 19°40′30″E) oraz osiedlu Mikołaja Reja (51°78′83″N 19°41′73″E). Ponadto pobrano próbkę solanki pochodzącą ze źródła Tadeusz w Zabłociu (49°90′77″N 18°77′06″E), która jest wykorzystywana do wypełniana wcześniej wspomnianych tężni. Próbki zostały pobrane w 2019 roku.

Próbki solanki z tężni solankowych były zbierane i przygotowywane do izolacji DNA w sposób analogiczny jak w przypadku próbek solanki z Kopalni Soli Bochnia. W przypadku tarniny 1 cm fragmenty gałązek były płukane i wytrząsane w sterylnej, wolnej od nukleaz wodzie, a następnie woda ta była wykorzystywana do izolacji DNA. Po wstępnym przygotowaniu próbek zgodnie z wcześniej opisaną procedurą, przeprowadzono izolację DNA za pomocą zestawu PowerSoil DNA Isolation Kit (Mo Bio Laboratories Inc., Carlsbad, CA, USA).

Izolacja szczepów mikroorganizmów halofilnych

Szczepy środowiskowe mikroorganizmów, których genomy zostały zsekwencjonowane w prezentowanej pracy, wyizolowano z zebranych próbek środowiskowych w Pracowni Biologii Bakterii Wydziału BiOŚ UŁ zgodnie z metodyką przedstawioną w pracy Albuquerque et al. (Albuquerque et al., 2016b). Pokrótce, próbki solanki zostały pobrane do sterylnych probówek typu Falcon, a następnie od 1 do 5 ml solanki zostało przesączonych przez filtry membranowe (Gelman typ GN-6; wielkość porów 0,45 µm; średnica 47 mm). Następnie filtry zostały umieszczone na powierzchni podłoża stałego dla halofili DSMZ 372. Poziom zasolenia podłoża wahał się w zależności od szczepu w przedziale od 15 do 25% NaCl. W zależności od szczepu hodowla była prowadzona w temperaturze 28°C lub 37°C. Otrzymane czyste hodowle izolatów były przechowywane w głębokim zamrożeniu (-80°C) w obecności 10% DMSO. Izolację genomowego DNA przeprowadzono z nowo przygotowanych hodowli drobnoustrojów, wykorzystując zestaw QIAamp DNA Mini Kit (Qiagen, Hilden, Niemcy), zgodnie z wytycznymi producenta.

Sekwencjonowanie DNA i udostępnienie danych

W oparciu o protokół opracowanym przez firmę Illumina o nazwie „16S Metagenomic Sequencing Library Preparation Illumina” przygotowano biblioteki

DNA składające się z amplikonów fragmentów genów kodujących 16S rRNA. Natomiast izolaty DNA przeznaczone do sekwencjonowania metagenomów zostały przygotowane do sekwencjonowania za pomocą zestawu Vazyme TruePrep DNA Library Prep Kit V2 for Illumina (Nanjing, China). W trakcie przygotowywania bibliotek do sekwencjonowania równolegle przygotowano negatywną próbę kontrolną w celu wykrycia ewentualnych kontaminacji w trakcie preparatyki. Biblioteki DNA składające się z amplikonów fragmentów genów kodujących 16S rRNA zostały zsekwencjonowane w trybie sparowanych odczytów na sekwenatorze Illumina MiSeq (Illumina, San Diego, USA), natomiast biblioteki metagenomów zostały zsekwencjonowane z użyciem sekwenatora Illumina NextSeq 500 (Illumina, San Diego, USA) również w trybie sparowanych odczytów.

W przypadku sekwencjonowania genomów szczepów mikroorganizmów wyizolowanych ze środowisk zasolonych wykorzystano po 1 ng wcześniej wyizolowanego DNA, którego ilość i jakość została potwierdzona za pomocą metody fluorymetrycznej z wykorzystaniem platformy Qubit (ThermoFisher, Waltham, USA). Do przygotowania bilbiotek DNA o sparowanych końcach wykorzystano zestaw Nextera XT DNA sample preparation kit (Illumina Inc., San Diego, USA). Przygotowane w ten sposób biblioteki DNA zostały poddane sekwencjonowaniu na sekwenatorze Illumina NextSeq 500 (Illumina, San Diego, USA).

Dane uzyskane z sekwencjonowania prób środowiskowych pobranych w Kopalni Soli Bochnia zostały udostępnione na serwerach National Center for Biotechnology Information jako BioProject o numerze dostępowym PRJNA848445. Genomy zsekwencjonowanych szczepów bakterii zostały udostępnione w BioProjectach o numerach dostępowych PRJNA899688, PRJNA899690, PRJNA899692, PRJNA899693, PRJNA899694 a genomy szczepów Archaea w BioProjectach o numerach dostępowych PRJNA900050, PRJNA900052 i PRJNA900053.

Analiza danych

W pierwszym etapie analizy, dla odczytów ze wszystkich typów sekwencjonowań przeprowadzono kontrolę jakości z zastosowaniem narzędzia FastQC (Andrews, 2022). W następnym kroku sekwencje adaptorowe i sekwencje o niskiej jakości zostały usunięte za pomocą narzędzia Trim Galore! v. 0.6.4 (Krueger, 2022).

W dalszych etapach, szlaki analiz dla wyników z sekwencjonowania amplikonów fragmentów genów kodujących 16S rRNA, metagenomów oraz genomów wyizolowanych szczepów przebiegały w odmienny sposób. W przypadku analizy wyników z sekwencjonowania genów 16S rRNA analizy były prowadzone z wykorzystaniem platformy QIIME 2 2021.8 (Bolyen et al., 2019). Narzędzie DADA2 zostało wykorzystane do odszumienia danych,

połącznia odczytów *forward* (R1) i *reverse* (R2) oraz wygenerowania tabeli ASV (*amplicon sequence variants*) (Callahan et al., 2016). Zarówno odczyty R1, jak i R2, zostały przycięte na 20 nukleotydzie od końca 5' i odpowiednio na nukleotydzie 245 i 230 końca 3'. W dalszej kolejności, na podstawie otrzymanej tabeli ASV obliczono miary α- i β- różnorodności. W celu ograniczenia różnic w wynikach związanych z głębokością sekwencjonowania poszczególnych próbek wykorzystano metodę próbkowania i z każdej z próbek wylosowano 11542 kontigi (co odpowiadało liczbie kontigów w próbce o najniższej głębokości sekwencjonowania, którą zdecydowano się pozostawić w analizie). Porównanie α-różnorodności między grupami zostało wykonane za pomocą testu Kruskala-Wallis (Kruskal and Wallis, 1952). Dla porównań β-różnorodności wykorzystywany był test PERMANOVA (Anderson, 2008). Klasyfikacja taksonomiczna został wykonana w oparciu o bazę danych SILVA. Kontigi zostały poddane adnotacji funkcjonalnej za pomocą PICRUSt2 v. 2.4.1 (Douglas et al., 2020). Analiza ANCOM została przeprowadzona w celu identyfikacji różnic w częstości występowania poszczególnych taksonów (Mandal et al., 2015).

Dla danych z sekwencjonowania metagenomów, w pierwszej kolejności przeprowadzono składanie *de novo* metagenomów za pomocą MEGAHIT v.1.2.9, przy minimalnej długości otrzymywanych odczytów określonej na poziomie 1000 pz (Li et al., 2015). Jakość złożeń została oceniona z pomocą metaQUAST (Mikheenko et al., 2016). Następnie kontigi zostały zgrupowane do MAGs (*metagenome-assembled genomes*) za pomocą metaBAT2 v. 2:2.15 (Kang et al., 2019). W celu dereplikacji uzyskanych MAGs wykorzystano dRep v3.2.2 (Olm et al., 2017). MAGs o długości większej niż 50000pz, kontaminacji mniejszej niż 25% i kompletności powyżej 75%, zostały poddane dalszej analizie. Adnotacje genów wykonano za pomocą DFAST v. 1.2.14 (Tanizawa et al., 2018). Za pomocą antiSMASH v.6.0.1 (Blin et al., 2021) i deepBGC v. 0.1.27 (Hannigan et al., 2019) zidentyfikowano klastry genów związanych z biosyntezą wtórnych metabolitów. Klasyfikacja taksonomiczna MAGs została przeprowadzona za pomocą gtdbtk v. 1.5.1 (Chaumeil et al., 2020) a na poziomie odczytów za pomocą Kaiju v. 1.8.2 (Menzel et al., 2016). Identyfikacja sekwencji peptydów przeciwdrobnoustrojowych (AMP) została wykonana za pomocą marcel v. 0.3.1 (Santos-Júnior et al., 2020). Właściwości peptydów i ich klasyfikacja funkcjonalna została przeprowadzona za pomocą platformy iAMP-2L (Xiao et al., 2013).

Genomy wyizolowanych mikroorganizmów, po wstępnym oczyszczeniu danych opisanym wcześniej, zostały złożone *de novo* za pomocą SPAdes v3.15.0 (Bankevich et al., 2012). Kontigi krótsze niż 500 pz lub o głębokości sekwencjonowania niższej niż 2 zostały odfiltrowane ze złożeń. W celu sprawdzenia jakości złożeń wykorzystano CheckM (Parks et al., 2015) i Quast (Mikheenko et al., 2018). Adnotacje genomów wykonano za pomocą narzędzi Prokka v.1.14.0 (Seemann, 2014)i eggNOG-mapper v.5.0.0 (Huerta-Cepas et al., 2019). Za pomocą narzędzia antiSMASH v.6.0.1 (Blin et al., 2021) zidentyfikowano klastry genów związanych z biosyntezą wtórnych metabolitów. Adnotację taksonomiczną wykonano za pomocą gtdbtk v. 1.5.1 (Chaumeil et al., 2020). Serwer TYGS został wykorzystany do przeprowadzenia analizy filogenomicznej (Meier-Kolthoff and Göker, 2019). Do wykonania drzew

filogenetycznych dla genów 16S rRNA i *gyrB* wykorzystano narzędzie MEGA X (Kumar et al., 2018).

Analiza peptydów przeciwdrobnoustrojowych

W oparciu o dane metagenomiczne wytypowano sekwencje peptydów o potencjalnym działaniu przeciwdrobnoustrojowym zgodnie z wcześniej opisaną metodyką. Sekwencje, dla których prawdopodobieństwo posiadania aktywności przeciwdrobnoustrojowej przekraczało 70% i które zostały oznaczone jako potencjalnie niehemolityczne zostały wytypowane do dalszych analiz. W oparciu o zidentyfikowane sekwencje, u komercyjnego dostawcy Pepmic Co. Ltd (Suzhou, Chiny), zamówiono syntezę trzech peptydów, najlepiej rokujących pod względem ocenionej *in silico* aktywności przeciwdrobnoustrojowej. Zsyntezowane peptydy posiadały certyfikat kontroli jakości.

Aktywność zsyntetyzowanych peptydów została zbadana *in vitro* względem siedmiu szczepów referencyjnych z kolekcji ATCC (*Acitenobacter baumannii* 19606, *Kliebsiella pneumoni*ae 700603, *Staphylococcus aureus* 43300, *Escherichia coli* 25922, *Enterococcus faecalis* 29212, *Pseudomonas aeruginosa* 27853 i *Candida albicans* 10231). Aktywność przeciwdrobnoustrojowa została zbadana metodą mikrorozcieńczeń z pomiarem spektrofotometrycznym (λ = 580 nm) w momencie podania peptydu i po 24 godzinach. Aktywność przeciwdrobnoustrojowa peptydów była badana w zakresie stężeń od 0.5 µg/mL do 256 µg/mL. Dodatkowym testem, który wykonano po pomiarze spektrofotometrycznym był test z użyciem 1% roztworu chlorku 2,3,5-trójfenylotetrazolu (TTC), który pozwala na wizualną ocenę żywotności drobnoustrojów. Zastosowanie obu opisanych metod pozwoliło na jednoczesne określenie minimalnego stężenia hamującego (MIC) oraz minimalnego stężenia bakteriobójczego (MBC) i/lub minimalnego stężenia grzybobójczego (MFC) badanego peptydu. W przypadku gdy nie było możliwości wyznaczenia MIC wyznaczano minimalną wartość stężenia przy której doszło do zahamowania wzrostu mikroorganizmów o 50% (MIC50). Analizie poddano również interakcję peptydów z antybiotykami referencyjnymi względem badanych szczepów mikroorganizmów. W tym celu przeprowadzono dwa eksperymenty. Pierwszym z nich był test mający na celu ocenę MBC, polegający na zbadaniu efektu skojarzonego działania peptydu w stężeniu odpowiadającym MIC50 z obniżonymi dawkami antybiotyku referencyjnego. Kolejnym krokiem było natomiast przeprowadzenie testu synergicznego w podejściu krzyżowym, gdzie peptyd oraz antybiotyk referencyjny były dodawane łącznie, w seriach rozcieńczeń, w celu zidentyfikowania ewentualnej interakcji, a w szczególności synergizmu ich działania.

W dalszej kolejności oceniono cytotoksyczność badanych peptydów względem linii pierwotnej komórek nabłonkowych kanalika proksymalnego nerki (RPTEC) z zastosowaniem czerwieni neutralnej. Badanie zostało przeprowadzone zgodnie z wytycznymi ISO:10993.

Oceniono również aktywność hemolityczną badanych peptydów względem ludzkich erytrocytów. Peptydy w stężeniu odpowiadającym ich MIC50 były inkubowane przez 3 godziny z 5% (v/v) zawiesiną erytrocytów .Po tym czasie próbki zostały odwirowane przez 15 min z przyśpieszeniem 1500×g a dla supernatantu zmierzono absorbancję przy 540 nm celem oznaczenia poziomu wolnej hemoglobiny.

Omówienie wyników

Cykl tematycznie spójnych publikacji wchodzących w skład pracy doktorskiej otwiera artykuł **„The Methods of Digging for "Gold" within the Salt: Characterization of Halophilic Prokaryotes and Identification of Their Valuable Biological Products Using Sequencing and Genome Mining Tools"**. Artykuł jest przeglądem publikacji opisujących strategie eksperymentalne nakierowane na identyfikację nowych bioproduktów mikroorganizmów stosowanych w badaniach nad mikroorganizmami halofilnymi. Skupiono się w nim w szczególności na metodach wykorzystujących sekwencjonowanie kwasów nukleinowych oraz analizie danych, która pozwala na wysokowydajne i ukierunkowane poszukiwania genów i klastrów genów związanych z wytwarzaniem bioproduktów. W publikacji omówiono również zalety i wady poszczególnych technologii sekwencjonowania, ze szczególnym zwróceniem uwagi na technologie trzeciej generacji obejmujące sekwencjonowanie nanoporowe i sekwencjonowanie SMRT (Single Molecule Real-Time). Wskazano, że zwiększenie długości odczytów, na które pozwala wykorzystanie metod sekwencjonowania trzeciej generacji korzystnie wpływa na jakość otrzymywanych wyników analiz, a w szczególności na ciągłość otrzymywanych złożeń genomów i metagenomów. Zauważono jednak również, że technologie te nadal mają wiele problemów, w tym związanych z wysokimi wymaganiami co do jakości izolatów DNA, które muszą być wykorzystywane w sekwencjonowaniu oraz, w przypadku sekwencjonowania nanoporowego, wysoki odsetek błędów w generowanych odczytach. Ponadto w artykule podsumowano informacje na temat narzędzi bioinformatycznych wykorzystywanych do identyfikacji w genomach i metagenomach sekwencji związanych z syntezą bioproduktów. Przedstawiono narzędzia wykorzystywane do identyfikacji klastrów genów związanych z biosyntezą wtórnych metabolitów, enzymów i peptydów przeciwdrobnoustrojowych. W artykule zwrócono uwagę, że rozwój baz danych zawierających dobrze scharakteryzowane genomy mikroorganizmów halofilnych i metagenomy środowisk zasolonych jest kluczowy dla dalszego rozwoju metod analitycznych oraz identyfikacji nowych bioproduktów wytwarzanych przez organizmy halofilne, które będą mogły być wykorzystane w przemyśle. Tego typu dane mogą być wykorzystywane zarówno bezpośrednio, w procesach identyfikacji interesujących genów i klastrów genów w procesach analitycznych, jak i do tworzenia nowych klasyfikatorów w tym klasyfikatorów wykorzystujących metody uczenia maszynowego, dla których kluczowym jest, aby na etapie ich przygotowania wykorzystywać jak najlepsze zbiory uczące, które mogą zostać zaaplikowane do zoptymalizowania algorytmu klasyfikacji.

Pierwsza praca eksperymentalna wchodząca w skład niniejszej pracy doktorskiej pt.: **„A First Insight into the Polish Bochnia Salt Mine Metagenome"** dotyczy charakterystyki mikrobiomu Kopalni Soli Bochnia. Z uwagi na swój unikatowy charakter związany z odizolowaniem od innych środowisk zasolonych oraz znaczną wartość historyczną lepsze poznanie

mikrobiomu Kopalni Soli Bochnia ma ogromną wartość poznawczą. Ponadto w dotychczasowych badaniach mikrobiom związany z kopalniami soli nie został zbyt dobrze poznany. Jedynie nieliczne prace opisują skład taksonomiczny mikrobiomu środowisk o podobnej charakterystyce z innych stron świata, a ocena potencjału biotechnologicznego mikroorganizmów zasiedlających tego typu środowiska w oparciu o dane metagenomiczne, nie była przeprowadzana. Biorąc pod uwagę wyniki sekwencjonowania amplikonów fragmentów genów kodujących 16S rRNA oraz wyniki sekwencjonowania metagenomów po raz pierwszy zobrazowano skład taksonomiczny mikrobiomu Kopalni Soli Bochnia oraz oceniono potencjał biotechnologiczny bytujących tam mikroorganizmów.

We wszystkich próbkach poddanych analizie w oparciu o dane z sekwencjonowania amplikonów fragmentów genów kodujących 16S rRNA zidentyfikowano 3655 unikatowych ASVs, które należały do 46 typów, 93 klas, 209 rzędów, 352 rodzin i 635 rodzajów należących do Eubakterii i Archaea. W pierwszej kolejności oceniono, czy występują istotne różnice w bioróżnorodności pomiędzy próbkami pobranymi z tych samych stanowisk w różnych latach. Wykazano, że próbki z tych samych stanowisk tworzą wspólne klastry niezależnie od sezonu, w którym zostały pobrane. Ponadto porównanie dystansu między próbkami w oparciu o odległość Bray'a-Curtis'a również nie wykazało istotnej statystycznie różnicy pomiędzy poszczególnymi sezonami (pseudo-F = 1,020224, p = 0,381). W dalszej kolejności oceniono różnicę w bioróżnorodności oraz składzie mikrobiomu poszczególnych stanowisk i poziomów kopalni. Wykazano, że próbki tworzą trzy główne klastry. Dwa odseparowane klastry składały się z próbek pochodzących ze stanowisk RDG (pierwszy poziom kopalni) oraz RD10 (pierwszy poziom kopalni)/Wernier (trzeci poziom kopalni), trzeci klaster zawierał próbki ze wszystkich pozostałych stanowisk. Jako taksony najmocniej wpływające na klastrowanie próbek wskazano klasy Halobacteria, Gammaproteobacteria i Clostridia oraz rodziny *Halomicrobiaceae, Thiohalorhabdaceae* i *Halomonadaceae*. Nie wykazano istotnej statystycznie różnic indeksu Shannona pomiędzy stanowiskami lub poziomami kopalni. Wykazano natomiast istotne statystycznie różnice pomiędzy poziomami kopalni w oparciu o odległość Bray'a-Curtis'a. Różnice zidentyfikowano pomiędzy pierwszym i trzecim (pseudo-F = 3.633673, p = 0.002) oraz pomiędzy pierwszym i czwartym poziomem kopalni (pseudo-F = 2.268532, p = 0.015). Można przypuszczać, że różnice te związane mogą być z różnym poziomem zmineralizowania solanek oraz z różnicami w grubości warstwy skalnej pokonywanej przez zasilające je wody gruntowe. Analiza ANCOM wykazała natomiast, że mikroorganizmy należące do klasy Chlamydiae stanowią istotnie większą część mikrobiomu stanowisk znajdujących się na pierwszym poziomie kopalni niż tych z pozostałych poziomów. Średni udział tej klasy w mikrobiomie na pierwszym poziomie kopalni wynosił 3,5%, na czwartym 0,04% a na trzecim poziomie takson ten był nieobecny. Zaobserwowanie tak wysokiego odsetka sekwencji przypisanych do klasy Chlamydiae jest interesujące ze względu na fakt,, że nie jest to takson typowy dla środowisk o wysokim zasoleniu. Należy jednak zauważyć, że już wcześniej raportowano obecność tego taksonu w próbkach solanek, jednak nie w tak znacznym procencie. Co

ciekawe, większość sekwencji należących do Chlamydiae zostało przypisanych do rodziny *Simkaniaceae*, a we wcześniejszych badaniach wykazano obecność tych mikroorganizmów w wodach powierzchniowych. Zatem można przypuszczać, że razem z tego typu wodami trafiły do solanek na najwyższym poziomie kopalni. Innym ważnym odkryciem jest zjawisko, dla którego wykazano bardzo wysoki udział sekwencji przypisanych do typu Patescibacteria. Udział ten wynosił 21,6% w odróżnieniu od średniej dla pozostałych próbek wynoszącej 3,5%. Ta grupa taksonomiczna charakteryzuje się silnie zredukowanym genomem, którego wielkość zazwyczaj nie przekracza miliona par zasad. Przedstawiciele tego typu są powszechnie identyfikowani w próbkach z warstw wodonośnych, gdzie warunki oscylują między tlenowymi, a beztlenowymi oraz mogą być związane z obiegiem siarki w przyrodzie.

Analiza wykonana w oparciu o dane metagenomiczne wykazała nieco inną charakterystykę taksonomiczną mikrobiomu Kopalni Soli Bochnia. W odróżnieniu od wyników z sekwencjonowania amplikonów fragmentów genów kodujących 16S rRNA, w danych metagenomicznych dominowały sekwencje należące do Archaea. Z danych metagenomicznych udało się odzyskać 16 unikatowych MAGs, z których 12 należało do Archaea, co również wskazuje, że były to organizmy o największym udziale w badanych próbach. Przypisanie gatunku było możliwe tylko dla jednego MAG oznaczonego jako WSO4.4, który został zidentyfikowany, jako należący do gatunku *Halobacterium bonnevillei*. W przypadku pozostałych MAGs podobieństwo ich sekwencji do znanych gatunków było niewystarczające, aby je do nich zaklasyfikować, co sugeruje, że w Kopalni Soli Bochni bytują liczne, dotąd nieznane gatunki mikroorganizmów. Co ciekawe, dwa z otrzymanych MAGs należą do mikroorganizmów z rodziny *Nanosalinaceae* przypisanych do nadtypu DPANN, który zawiera obecnie niemożliwe do wyhodowania w warunkach laboratoryjnych mikroorganizmy żyjące w symbiozie z mikroorganizmami z klasy Halobacteria. Jest to nadal dość słabo znana grupa taksonomiczna, w związku z czym pozyskanie dodatkowych danych na temat ich występowania i sekwencji genomicznych jest niezwykle cenne. Różnice w składzie taksonomicznym występujące między wynikami sekwencjonowania metagenomów i sekwencjonowania amplikonów fragmentów genów kodujących 16S rRNA mogą wynikać z pewnej niedoskonałości metod opartych o sekwencjonowanie fragmentów genów markerowych i nadreprezentacji w ich wypadku sekwencji należących chociażby do Proteobacteria, a także niedoszacowaniu sekwencji należących do Archaea. Problem ten może być związany ze specyficznością starterów reakcji PCR zastosowanych w badaniu, które najprawdopodobniej charakteryzują się większym powinowactwem do sekwencji bakteryjnych co powoduje późniejsze rozbieżności w wynikach sekwencjonowania metagenomów i amplikonów.

W celu określenia potencjału biotechnologicznego badanych środowisk podjęto próbę identyfikacji klastrów genów związanych z biosyntezą wtórnych metabolitów oraz sekwencji kodujących peptydy przeciwdrobnoustrojowych w złożonych metagenomach. Klastry genów związane z biosyntezą terpenów okazały się najbardziej rozpowszechnione i występowały w każdej z próbek. Największą liczbę klastrów genów związanych z biosyntezą wtórnych

metabolitów zidentyfikowano w próbce POP27 pochodzącej z trzeciego poziomu kopalni. Jeden z klastrów kodował pełen operon *ectABC* związany z produkcją ektoiny. W próbce tej zidentyfikowano również kilka klastrów zorganizowanych wokół genu kodującego syntazę fitoetenu, co może sugerować, że są związane z produkcją karotenoidów. Ponadto była to jedyna próbka, w której zidentyfikowano klastry związane z produkcją sideroforów i tiopeptydów. W pozostałych próbkach zidentyfikowano natomiast dwa klastry związane z produkcją lantipeptydów klasy drugiej, operon *ectABC* podobnie jak w przypadku próbki POP27 oraz klaster związany z produkcja białek typu *RiPP*, który najprawdopodobniej związany jest produkcją bakteriocyny podobnej do Linocin M18. Ponadto we wszystkich próbkach zidentyfikowano sekwencje kodujące peptydy przeciwdrobnoustrojowe. Najwięcej sekwencji tego typu zidentyfikowano w metegenomie próbki WSO4 a najmniej w próbce W81.

Podsumowując, w przedstawionym powyżej artykule scharakteryzowano mikrobiom Kopalni Soli Bochnia zarówno pod kątem jego bioróżnorodności jak i potencjału biotechnologicznego wchodzących w jego skład mikroorganizmów. Wykazano, że skład mikrobiomów poszczególnych stanowisk jest stabilny w czasie, oraz że występują różnice w składzie mikrobiomów poszczególnych poziomów kopalni. Odzyskano 16 unikatowych MAGs, z których 15 należy najprawdopodobniej do nieznanych dotychczas gatunków. Ponadto zidentyfikowano w metagenomach liczne klastry genów związanych z biosyntezą wtórnych metabolitów oraz geny kodujące sekwencje peptydów przeciwdrobnoustrojowych.

W kolejnym etapie badań zsekwencjonowano i scharakteryzowano genomy pięciu szczepów halofilnych bakterii oraz trzech szczepy halofilnych Archaea wyizolowanych ze środowisk o wysokim zasoleniu z obszaru Przedgórza Karpackiego. Szczepy te zostały wytypowane do sekwencjonowania całogenomowego na podstawie wstępnej analizy sekwencji 16S rRNA, która wykazała, że szczepy te najprawdopodobniej nie należą do żadnego ze znanych dotychczas gatunków halofili. Na podstawie uzyskanych wyników opracowano dwa artykuły wchodzące w skład niniejszej pracy doktorskiej: **„Draft Genomes of Halophilic Chromohalobacter and Halomonas Strains Isolated from Brines of The Carpathian Foreland, Poland”** oraz **“Draft Genomes of Halophilic Archaea Strains Isolated from Brines of the Carpathian Foreland, Poland”**.

W przypadku genomów halofilnych bakterii szczepy oznaczone jako 11-W, 296-RDG, oraz 48-RD10 zostały przypisane do rodzaju *Chromohalobacter,* a szczepy 11-S5 i 25-S5 do rodzaju *Halomonas*. W toku dalszych porównań zidentyfikowano duże rozbieżności pomiędzy analizowanymi sekwencjami a genomami dostępnymi w bazach danych, co sugeruje, że mamy do czynienia z nieopisanymi dotąd gatunkami. Analizowane genomy składały się z kontigów w liczbie od 38 do 321, przy wielkości złożeń w przedziale od 3,6 do 3,9 Mpz. Kompletność genomów była oszacowana na 98,71-99,86% przy kontaminacji na poziomie 0,54-8,42%. Zawartość par GC wynosiła od 60,11 do 66,46%. Na drzewie filogenetycznym stworzonym w oparciu o sekwencje genów 16S rRNA

szczepy 296-RDG i 48-RD10 klastrowały wspólnie i lokowały się najbliżej *Chromohalobacter canadensis*. Natomiast szczep 11-W był najbliższy gatunkowi *Chromohalobacter sarecensis*. Szczepy 11-S5 i 25-S5 utworzyły wspólny klaster i uplasowały się najbliżej *Halomonas sediminicola*. Przynależność szczepów 296-RDG i 48-RD10 oraz szczepów 11-S5 and 25-S5 do tych samych gatunków została potwierdzona za pomocą wyniku analizy dDDH, której wartość wyniosła dla tych par odpowiednio 87.5% i 89.7%. W wyniku adnotacji funkcjonalnej, w analizowanych genomach oznaczono klastry genów związane z biosyntezą wtórnych metabolitów. W szczepach należących do *Chromohalobacter* zidentyfikowano większą liczbę klastrów niż w szczepach należących do rodzaju *Halomonas.* We wszystkich genomach zidentyfikowano chociaż fragmentaryczne klastry genów związanych z biosyntezą ektoiny. W szczepach należących do *Chromohalobacter* zidentyfikowano ponadto klastry związane z biosyntezą sideroforów, kofaktorów reakcji redoks oraz arylopolienów.

Genomy szczepów halofilnych Archaea obejmowały od 33 do 177 kontigów. Wielkość genomów mieściła się w przedziale 2,7-3,0 Mpz. Kompletność genomów wynosiła 95,97-97,93%, a kontaminacja od 0,0% do 0,95%. Zawartość par GC w genomach została oceniona na 63,77-68,77%. W oparciu o analizę filogenetyczną i filogenomiczną szczep Boch-4 został przypisany do gatunku *Haloarcula hispanica*. Szczep Boch-26 natomiast został przypisany jedynie do rodzaju *Halorubrum*. W przypadku szczepu POP-27 przypisanie taksonomiczne było niejednoznaczne. Przypisanie przeprowadzone w oparciu o parametr ANI oraz drzewo filogenetyczne dla genu *gyrB* sugerowało, że szczep ten należy do gatunku *Halopenitus malekzadehii*. Natomiast parametr dDDH pomiędzy szczepem POP-27 a *Halopenitus malekzadehii* wyniósł 59,2%, co wskazywało na to, że genomy te nie należą do jednego gatunku mikroorganizmu. W związku z rozbieżnymi rezultatami uzyskanymi za pomocą różnych metod, zdecydowano się potraktować szczep POP-27, jako należący do rodzaju *Halopenitus,* ale prawdopodobnie nie należący do gatunku *Halopenitus malekzadehii*. W toku analizy funkcjonalnej, w każdym ze szczepów zidentyfikowano po dwa klastry genów związanych z biosyntezą terpenów. Klastry różniły się między sobą, jednak każdy z nich był zorganizowany wokół genu kodującego syntazę fitoetenu, co może sugerować, że są związane z produkcją karotenoidów. Ponadto szczep Boch-4 kodował klaster genów związany najprawdopodobniej z syntezą peptydów typu lasso.

W powyższych pracach scharakteryzowano genomy pięciu szczepów halofilnych bakterii oraz trzech szczepów halofilnych Archaea. W oparciu o analizę filogenomiczną wykazano, że genomy szczepów bakteryjnych należą najprawdopodobniej do dotychczas nieznanych gatunków. W przypadku Archaea dwa z analizowanych genomów należą do mikroorganizmów z rodzaju *Halopenitus*. Jeden z tych genomów jest trudny do jednoznacznego przypisania taksonomicznego i może stanowić odrębny gatunek w obrębie rodzaju *Halopenitus* lub też nieopisany szczep w obrębie gatunku *H. malekzadehii*. Drugi genom najprawdopodobniej należy do nieznanego dotąd gatunku z rodzaju *Halorubrum*, a ostatni z genomów należy do nowego szczepu *Haloarcula hispanica.* Ponadto w toku analiz scharakteryzowano profil funkcjonalny genów

kodowanych w poszczególnych genomach oraz zidentyfikowano kodowane w nich klastry genów związanych z biosyntezą wtórnych metabolitów.

W związku z rosnącym problem z antybiotykoodpornością koniecznym jest, aby koncentrować się i wspierać prace nad opracowaniem nowych substancji przeciwdrobnoustrojowych. Szczególna uwaga powinna być poświęcona substancjom o nowych mechanizmach działania, które będą odmienne od tych charakteryzujących obecnie stosowane antybiotyki. Znane są peptydy przeciwdrobnoustrojowe produkowane przez mikroorganizmy halofilne, jednak niewiele z nich zostało przebadanych pod kątem aktywności przeciw mikroorganizmom patogennym. W większości prac ich aktywność była badana głównie względem szczepów pokrewnych dla szczepu, z którego wyizolowano dany peptyd.

Ostatnim etapem badań było przeprowadzenie analiz *in silico* bazujących na sekwencjach metagenomowych otrzymanych w wyniku sekwencjonowania DNA z próbek środowiskowych pochodzących z Kopalni Soli Bochnia oraz z tężni solankowych, mających na celu wybór sekwencji peptydów o najwyższym prawdopodobieństwie działania przeciwdrobnoustrojowego (powyżej 70%). Finalnie, bazując na zsyntezowanych komercyjnie peptydach dokonano walidacji i ewentualnego potwierdzenia działania przeciwdrobnoustrojowego badanych peptydów . Wyniki uzyskane w ramach tego etapu przedstawiono w artykule **„Novel AMP from saline environments - promising glimmer for inhibition of multidrug resistant E. faecalis and S. aureus infections?”**.

W toku analiz bioinformatycznych, w danych metagenomicznych zidentyfikowano 271 sekwencji, które potencjalnie kodowały peptydy przeciwdrobnoustrojowe. 191 z nich zostało określonych jako potencjalnie hemolityczne i w związku z tym usunięto je z dalszych analiz, jako stwarzające zagrożenie w ewentualnych zastosowaniach terapeutycznych. Prawdopodobieństwo tego, że dana sekwencja koduje AMP dla pozostałych 80 sekwencji znajdowało się w przedziale od 50,05 do 76,2%. Do syntezy wybrano jedynie trzy peptydy, dla których prawdopodobieństwo było większe od 70% i oznaczono je jako P1, P2 i P3. Zgodnie z wstępną klasyfikacją peptydy P1 i P3 powinny charakteryzować się aktywnością przeciwbakteryjną, a peptyd P2 aktywnością przeciwgrzybową. Każdy z wytypowanych peptydów charakteryzował się dodatnim ładunkiem netto oraz zawierał od 32 do 41% aminokwasów hydrofobowych co może sugerować, że mogą mieć one właściwości amfipatyczne. Dla analizowanych peptydów wyznaczono wartości MIC50 dla szczepów *Acitenobacter baumanii* 19606, *Kliebsiella pneumoniae* 700603, *Staphylococcus aureus* 43300, *Escherichia coli* 25922, *Enterococcus faecalis* 29212, *Pseudomonas aeruginosa* 27853 i *Candida albicans* 10231 z kolekcji ATCC. W przypadku peptydu P1 wyznaczono MIC50 na poziomie 32 µg/ml względem *E. faecalis* 29212 oraz dla peptydu P3 na poziomie 32 µg/ml względem *E. faecalis* 29212 i *S. aureus* 43300, a także 256 µg/ml względem *E. coli* 25922. Ponadto przeanalizowano interakcję peptydów w połączeniu z antybiotykami referencyjnymi jednak nie wykazano żadnych korzystnych interakcji. W dalszej kolejności zbadano cytotoksyczność oraz aktywność

hemolityczną peptydów przy stężeniach MIC50. Żywotność komórek w teście cytotoksyczności wynosiła 100-107% po 24, 100-115% po 48 i 99-107% po 72 godzinach. Aktywność hemolityczna wynosiła natomiast odpowiednio 6,92% i 7,39% dla peptydów P1 i P2. Wartość ta nie odbiegała od próby kontrolnej, gdzie procent hemolizy wynosił 7,63%, co świadczyło o braku cytotoksyczności względem erytrocytów i potwierdzało przewidywania analiz *in silico*. W kolejnym etapie, za pomocą dokowania molekularnego zweryfikowano możliwość działania analizowanych peptydów poprzez analogiczne punkty uchwytów jak dla antybiotyków referencyjnych. Wykazano, że istnieje możliwość takiego działania i przedstawiono potencjalny sposób wiązania peptydów z receptorami. Uzyskane wyniki udowodniły, że peptydy P1 oraz P3 stanowią wykazują znaczny potencjał aplikacyjny z racji na udowodnioną aktywność przeciwdrobnoustrojową oraz niską toksyczność. Peptydy o podobnej aktywności przeciwdrobnoustrojowej były i są obiektami pogłębionych badań nad ich wykorzystaniem klinicznym, w tym prowadzone są z ich zastosowaniem badania kliniczne (Koo and Seo, 2019). W związku z tym można zakładać, że analizowane peptydy również mogą zostać poddane dalszym badaniom o podobnym charakterze.

Wnioski

Przeprowadzone badania pozwoliły na wyciągnięcie następujących wniosków:

- Scharakteryzowano mikrobiom Kopalni Soli Bochnia. Wykazano, że jest on stabilny w czasie oraz opisano różnice w składzie mikrobiomu pomiędzy poziomami kopalni. Poprzez identyfikację 16 nowych MAGs, z których 15 należy do nieznanych dotąd mikroorganizmów halofilnych udowodniono, że Kopalnia Soli Bochnia jest unikatowym rezerwuarem bioróżnorodności, w którym można odnaleźć nieznane dotąd gatunki mikroorganizmów.
- Na podstawie danych metagenomicznych wykazano, że mikroorganizmy bytujące w Kopalni Soli Bochnia charakteryzują się znacznym potencjałem biotechnologicznym objawiającym się występowaniem w ich genomach klastrów genów związanych z biosyntezą wtórnych metabolitów takich jak ektoina, siderofory, czy też karotenoidy.
- Scharakteryzowano genomy ośmiu szczepów mikroorganizmów halofilnych. Wykazano, że siedem z opisanych szczepów najprawdopodobniej należy do nieznanych dotąd gatunków mikroorganizmów halofilnych. Ponadto wykazano, że w części genomów analizowanych szczepów zidentyfikowano klastry genów związanych z biosyntezą wartościowych bioproduktów.
- W analizowanych metagenomach zidentyfikowano szereg sekwencji kodujących peptydy przeciwdrobnoustrojowe, z których trzy zostały zsyntetyzowane i scharakteryzowane. Wykazano, że dwa z nich, opisane jako P1 i P3, posiadały aktywność przeciwdrobnoustrojową oraz nie wykazywały działania toksycznego względem komórek ludzkich. W związku z tym mogą być kandydatami do dalszych badań jako nowe substancje przeciwdrobnoustrojowe.

Bibliografia

Albuquerque, L., Kowalewicz-Kulbat, M., Drzewiecka, D., Staczek, P., d'Auria, G., Rosselló-Móra, R., et al. (2016a). Halorhabdus rudnickae sp. nov., a halophilic archaeon isolated from a salt mine borehole in Poland. *Syst. Appl. Microbiol.* 39, 100–105. doi:10.1016/j.syapm.2015.12.004.

Albuquerque, L., Kowalewicz-Kulbat, M., Drzewiecka, D., Staczek, P., d'Auria, G., Rosselló-Móra, R., et al. (2016b). Halorhabdus rudnickae sp. nov., a halophilic archaeon isolated from a salt mine borehole in Poland. *Syst. Appl. Microbiol.* 39, 100–105. doi:10.1016/j.syapm.2015.12.004.

Amoozegar, M. A., Siroosi, M., Atashgahi, S., Smidt, H., and Ventosa, A. (2017). Systematics of haloarchaea and biotechnological potential of their hydrolytic enzymes. *Microbiol. (United Kingdom)* 163, 623–645. doi:10.1099/mic.0.000463.

Anderson, M. J. (2008). A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* 26, 32–46. doi:10.1111/j.1442-9993.2001.tb00081.x.

Andrews, S. (2022). FastQC: a quality control tool for high throughput sequence data. Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Baker, B. J., Appler, K. E., and Gong, X. (2021). New Microbial Biodiversity in Marine Sediments. *Ann. Rev. Mar. Sci.* 13, 161–175. doi:10.1146/annurev-marine-032020-014552.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi:10.1089/cmb.2012.0021.

Belilla, J., Moreira, D., Jardillier, L., Reboul, G., Benzerara, K., López-García, J. M., et al. (2019). Hyperdiverse archaea near life limits at the polyextreme geothermal Dallol area. *Nat. Ecol. Evol.* 3, 1552–1561. doi:10.1038/s41559-019-1005-0.

Bilstein, A., Heinrich, A., Rybachuk, A., and Mösges, R. (2021a). Ectoine in the Treatment of Irritations and Inflammations of the Eye Surface. *Biomed Res. Int.* 2021, 8–10. doi:10.1155/2021/8885032.

Bilstein, A., Werkhäuser, N., Rybachuk, A., and Mösges, R. (2021b). The Effectiveness of the Bacteria Derived Extremolyte Ectoine for the Treatment of Allergic Rhinitis. *Biomed Res. Int.* 2021, 8–10. doi:10.1155/2021/5562623.

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema, M. H., et al. (2021). AntiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 49, W29–W35. doi:10.1093/nar/gkab335.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-

Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi:10.1038/s41587-019-0209-9.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi:10.1038/nmeth.3869.

Cavalazzi, B., Barbieri, R., Gómez, F., Capaccioni, B., Olsson-Francis, K., Pondrelli, M., et al. (2019). The Dallol Geothermal Area, Northern Afar (Ethiopia) - An Exceptional Planetary Field Analog on Earth. *Astrobiology* 19, 553–578. doi:10.1089/ast.2018.1926.

Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2020). GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36, 1925–1927. doi:10.1093/bioinformatics/btz848.

Corral, P., Amoozegar, M. A., and Ventosa, A. (2020). Halophiles and their biomolecules: Recent advances and future applications in biomedicine. *Mar. Drugs* 18. doi:10.3390/md18010033.

Cycil, L. M., DasSarma, S., Pecher, W., McDonald, R., AbdulSalam, M., and Hasan, F. (2020). Metagenomic Insights Into the Diversity of Halophilic Microorganisms Indigenous to the Karak Salt Mine, Pakistan. *Front. Microbiol.* 11, 1–14. doi:10.3389/fmicb.2020.01567.

Douglas, G. M., Maffei, V. J., Zaneveld, J. R., Yurgel, S. N., Brown, J. R., Taylor, C. M., et al. (2020). PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* 38, 685–688. doi:10.1038/s41587-020-0548-6.

Dutta, B., and Bandopadhyay, R. (2022). Biotechnological potentials of halophilic microorganisms and their impact on mankind. *Beni-Suef Univ. J. Basic Appl. Sci.* 11. doi:10.1186/s43088-022-00252-w.

Fernández, A. B., Vera-Gargallo, B., Sánchez-Porro, C., Ghai, R., Papke, R. T., Rodriguez-Valera, F., et al. (2014). Comparison of prokaryotic community structure from Mediterranean and Atlantic saltern concentrator ponds by a metagenomic approach. *Front. Microbiol.* 5, 1–12. doi:10.3389/fmicb.2014.00196.

Gómez, F., Cavalazzi, B., Rodríguez, N., Amils, R., Ori, G. G., Olsson-Francis, K., et al. (2019). Ultra-small microorganisms in the polyextreme conditions of the Dallol volcano, Northern Afar, Ethiopia. *Sci. Rep.* 9, 1–9. doi:10.1038/s41598-019-44440-8.

Gunde-Cimerman, N., Plemenitaš, A., and Oren, A. (2018). Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. *FEMS Microbiol. Rev.* 42, 353–375. doi:10.1093/femsre/fuy009.

Hannigan, G. D., Prihoda, D., Palicka, A., Soukup, J., Klempir, O., Rampula, L., et al. (2019). A deep learning genome-mining strategy for biosynthetic gene cluster prediction. *Nucleic Acids Res.* 47, e110. doi:10.1093/nar/gkz654.

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. doi:10.1093/nar/gky1085.

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 2019, 1–13. doi:10.7717/peerj.7359.

Koo, H. B., and Seo, J. (2019). Antimicrobial peptides under clinical investigation. *Pept. Sci.* 111. doi:10.1002/pep2.24122.

Krueger, F. (2022). Trim Galore. Available at: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

Kruskal, W., and Wallis, A. (1952). Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* 47, 583–621.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi:10.1093/molbev/msy096.

Kumar, V., Singh, B., van Belkum, M. J., Diep, D. B., Chikindas, M. L., Ermakov, A. M., et al. (2021). Halocins, natural antimicrobials of Archaea: Exotic or special or both? *Biotechnol. Adv.* 53, 107834. doi:10.1016/j.biotechadv.2021.107834.

Li, D., Liu, C. M., Luo, R., Sadakane, K., and Lam, T. W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi:10.1093/bioinformatics/btv033.

Liu, C., Baffoe, D. K., Zhan, Y., Zhang, M., Li, Y., and Zhang, G. (2019). Halophile, an essential platform for bioproduction. *J. Microbiol. Methods* 166, 1–8. doi:10.1016/j.mimet.2019.105704.

Mandal, S., Van Treuren, W., White, R. A., Eggesbø, M., Knight, R., and Peddada, S. D. (2015). Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb. Ecol. Heal. Dis.* 26. doi:10.3402/mehd.v26.27663.

Meier-Kolthoff, J. P., and Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* 10. doi:10.1038/s41467-019-10210-3.

Menzel, P., Ng, K. L., and Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7. doi:10.1038/ncomms11257.

Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., and Gurevich, A. (2018). Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* 34, i142–i150. doi:10.1093/bioinformatics/bty266.

Mikheenko, A., Saveliev, V., and Gurevich, A. (2016). MetaQUAST: Evaluation

of metagenome assemblies. *Bioinformatics* 32, 1088–1090. doi:10.1093/bioinformatics/btv697.

Olm, M. R., Brown, C. T., Brooks, B., and Banfield, J. F. (2017). DRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* 11, 2864–2868. doi:10.1038/ismej.2017.126.

Oren, A. (2008). Microbial life at high salt concentrations: Phylogenetic and metabolic diversity. *Saline Systems* 4, 1–13. doi:10.1186/1746-1448-4-2.

Oren, A. (2011). Extremophiles Handbook. *Extrem. Handb.* doi:10.1007/978-4-431-53898-1.

Oren, A. (2020). "The microbiology of red brines," in *Advances in Applied Microbiology* (Elsevier Inc.), 57–110. doi:10.1016/bs.aambs.2020.07.003.

Pal, S., Biswas, R., Misra, A., Sar, A., Banerjee, S., Mukherjee, P., et al. (2020). Poorly known microbial taxa dominate the microbiome of hypersaline Sambhar Lake salterns in India. *Extremophiles* 24, 875–885. doi:10.1007/s00792-020-01201-0.

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi:10.1101/gr.186072.114.

Pastor, J. M., Salvador, M., Argandoña, M., Bernal, V., Reina-Bueno, M., Csonka, L. N., et al. (2010). Ectoines in cell stress protection: Uses and biotechnological production. *Biotechnol. Adv.* 28, 782–801. doi:10.1016/j.biotechadv.2010.06.005.

Puławska, A., Manecki, M., Flasza, M., and Styszko, K. (2021). Origin, distribution, and perspective health benefits of particulate matter in the air of underground salt mine: a case study from Bochnia, Poland. *Environ. Geochem. Health* 43, 3533–3556. doi:10.1007/s10653-021-00832-2.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J. F., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499, 431–437. doi:10.1038/nature12352.

Santos-Júnior, C. D., Pan, S., Zhao, X. M., and Coelho, L. P. (2020). Macrel: Antimicrobial peptide screening in genomes and metagenomes. *PeerJ* 8, 1–20. doi:10.7717/peerj.10555.

Schultz, J., and Rosado, A. S. (2020). Extreme environments: a source of biosurfactants for biotechnological applications. *Extremophiles* 24, 189–206. doi:10.1007/s00792-019-01151-2.

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi:10.1093/bioinformatics/btu153.

Stat, M., Huggett, M. J., Bernasconi, R., Dibattista, J. D., Berry, T. E., Newman, S. J., et al. (2017). Ecosystem biomonitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. *Sci. Rep.* 7, 1–11. doi:10.1038/s41598-017-12501-5.

Tanizawa, Y., Fujisawa, T., and Nakamura, Y. (2018). DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34, 1037–1039. doi:10.1093/bioinformatics/btx713.

UNESCO (2023). Wieliczka and Bochnia Royal Salt Mines. Available at: https://whc.unesco.org/en/list/32/.

Ventosa, A., and Arahal, D. R. (2009). “Physico-chemical characteristics of hypersaline environments and their biodiversity,” in *Extremophiles*.

Ward, D., Melbourne-Thomas, J., Pecl, G. T., Evans, K., Green, M., McCormack, P. C., et al. (2022). *Safeguarding marine life: conservation of biodiversity and ecosystems*. Springer International Publishing doi:10.1007/s11160-022-09700-3.

Xiao, X., Wang, P., Lin, W. Z., Jia, J. H., and Chou, K. C. (2013). IAMP-2L: A two-level multi-label classifier for identifying antimicrobial peptides and their functional types. *Anal. Biochem.* 436, 168–177. doi:10.1016/j.ab.2013.01.019.

Yin, J., Chen, J. C., Wu, Q., and Chen, G. Q. (2015). Halophiles, coming stars for industrial biotechnology. *Biotechnol. Adv.* 33, 1433–1442. doi:10.1016/j.biotechadv.2014.10.008.

Zaremba-Niedzwiedzka, K., Caceres, E. F., Saw, J. H., Bäckström, Di., Juzokaite, L., Vancaester, E., et al. (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541, 353–358. doi:10.1038/nature21031.

Streszczenie

Streszczenie w języku polskim

Halofile są niezwykle ciekawą i zróżnicowaną grupą mikroorganizmów. Można wśród nich znaleźć przedstawicieli każdej z domen życia. Dzięki rozwojowi technik biologii molekularnej oraz metod sekwencjonowania, w ostatnich latach poczyniono znaczny postęp w badaniach nad bioróżnorodnością halofili. Szczególne znaczenie w tym względzie mają metody metagenomiczne pozwalające na lepsze poznanie mikroorganizmów, dla których nie udało się dotychczas opracować metod hodowli w warunkach laboratoryjnych. Przykładem grup taksonomicznych poznanych dzięki tym metodom mogą być nadtypy Archaea takie jak Asgardarchaeota i DPANN.

W związku z koniecznością adaptacji do życia w warunkach wysokiego zasolenia halofile wytworzyły wiele unikatowych szlaków metabolicznych. Szlaki te są związane na przykład z biosyntezą osmoprotektantów takich jak ektoina lub hydroksyektoina czy też barwników o działaniu antyoksydacyjnych takich jak karotenoidy. Ponadto w genomach halofili znajdują się geny kodujące enzymy aktywne w warunkach wysokiego zasolenia a często również innych niesprzyjających warunkach takich jak wysoka temperatura czy skrajne pH. Dzięki tym właściwościom enzymy te znajdują zastosowanie w różnych gałęziach przemysłu takich jak na przykład przetwórstwo żywności, przemysł papierniczy czy tekstylny. Mikroorganizmy halofilne są również producentami wielu substancji o potencjalne farmaceutycznym. Można do nich zaliczyć zarówno wtórne metabolity jak i peptydy wykazujące aktywność przeciwdrobnoustrojową.

Celem przedstawionej pracy doktorskiej było scharakteryzowanie bioróżnorodności mikrobiomów wybranych środowisk o wysokim zasoleniu oraz ocena potencjału biotechnologicznego bytujących w nich mikroorganizmów w oparciu o dane genomiczne i metagenomiczne. Wygenerowane dane oraz przeprowadzone analizy pozwoliły na ocenę składu taksonomicznego mikrobiomu Kopalni Soli Bochnia oraz na identyfikację klastrów genów związanych z biosyntezą wtórnych metabolitów. Ponadto w analizowanych metagenomach zidentyfikowano 16 unikatowych MAGs, z których 15 nie ma swoich odpowiedników wśród dotychczas znanych sekwencji. W toku analiz zsekwencjonowano również genomy pięciu szczepów Eubakterii oraz trzech szczepów Archaea wyizolowanych z solanek pobranych na obszarze Przedgórza Karpackiego. Część z genomów należy najprawdopodobniej do nowych, nieopisanych dotychczas gatunków. W dalszej kolejności na podstawie danych metagenomicznych wytypowano trzy sekwencje, które potencjalnie mogły kodować peptydy przeciwdrobnoustrojowe. Zostały one zsyntezowane a ich właściwości zostały ocenione w warunkach laboratoryjnych. Wykazano, że peptyd P1 był aktywny przeciw *E. faecalis* 29212 a peptyd P3 przeciw *E. faecalis* 29212 i *S. aureus* 43300. Ponadto udowodniono, że oba analizowane peptydy nie mają właściwości cytotoksycznych ani nie wykazują aktywności hemolitycznej.

Podsumowując, w toku prac scharakteryzowano mikrobiom Kopalni Soli Bochnia oraz genomy szczepów mikroorganizmów wyizolowanych z solanek pobranych na obszarze Przedgórza Karpackiego. Ponadto wykazano, że wybrane peptydy kodowane w analizowanych metagenomach wykazują aktywność przeciwdrobnoustrojową i mogą być obiektem dalszych badań mających na celu ocenę możliwości wykorzystania ich w praktyce.

Streszczenie w języku angielskim

Halophiles are an extremely interesting and diverse group of microorganisms. Among them can be found representatives of each domain of life. Thanks to the development of molecular biology techniques and sequencing methods, significant progress has been made in research on the biodiversity of halophiles in recent years. Of particular importance in this field are metagenomic methods that allow for a better understanding of microorganisms for which methods of cultivation in laboratory conditions have not yet been developed. Archaeal superphyla such as Asgardarchaeota and DPANN can serve as an example of taxonomic groups discovered thanks to these methods.

Due to the need to adapt to life in conditions of high salinity, halophiles have developed many unique metabolic pathways. These pathways are related, for example, to the biosynthesis of osmoprotectants such as ectoine or hydroxyectoine, or pigments with antioxidant activity such as carotenoids. In addition, the genomes of halophiles contain genes encoding enzymes active in conditions of high salinity and often also in other unfavourable conditions such as high temperature or extreme pH. Thanks to these properties, halophilic enzymes are used in various industries, such as food processing, paper and textile industries. Halophilic microorganisms are also producers of substances with pharmaceutical potential. These include both secondary metabolites and peptides with antimicrobial activity.

The aim of the presented doctoral thesis was to characterize the biodiversity of microbiomes in selected environments with high salinity and to assess the biotechnological potential of microorganisms living in them based on genomic and metagenomic data. The generated data and the analyses made it possible to assess the taxonomic composition of the microbiome of the Bochnia Salt Mine and to identify biosynthetic gene clusters. In addition, 16 unique MAGs were identified in the analysed metagenomes, 15 of which have no equivalents among the previously known sequences. In the course of the analyses, the genomes of five strains of Bacteria and three strains of Archaea isolated from brines collected in the area of the Carpathian Foreland were also sequenced. Some of the genomes most likely belong to new, undescribed species. Subsequently, three sequences that could potentially encode antimicrobial peptides were selected based on metagenomic data. They were synthesized and their properties were evaluated under laboratory conditions. It was shown that the P1 peptide was active against *E. faecalis* 29212 and the P3 peptide against *E. faecalis* 29212 and *S. aureus* 43300. Moreover, both analysed peptides proved to be neither cytotoxic nor haemolytic.

In conclusion, in the course of the work, the microbiome of the Bochnia Salt Mine and the genomes of the strains of microorganisms isolated from brines collected in the area of the Carpathian Foothills were characterized. In addition, it has been shown that selected peptides encoded in the analysed metagenomes

exhibit antimicrobial activity and may be the subject of further research aimed at assessing the possibility of using them in practice.

Załączniki

genes

MDPI

Review

The Methods of Digging for "Gold" within the Salt: Characterization of Halophilic Prokaryotes and Identification of Their Valuable Biological Products Using Sequencing and Genome Mining Tools

Jakub Lach [1,2,*], Paulina Jęcz [1], Dominik Strapagiel [2], Agnieszka Matera-Witkiewicz [3] and Paweł Stączek [1]

1 Department of Molecular Microbiology, Faculty of Biology and Environmental Protection, University of Lodz, 93-338 Lodz, Poland; paulina.jecz@biol.uni.lodz.pl (P.J.); pawel.staczek@biol.uni.lodz.pl (P.S.)
2 Biobank Lab, Department of Molecular Biophysics, Faculty of Environmental Protection, University of Lodz, 93-338 Lodz, Poland; dominik.strapagiel@biol.uni.lodz.pl
3 Screening Laboratory of Biological Activity Tests and Collection of Biological Material, Faculty of Pharmacy, Wroclaw Medical University, 50-368 Wroclaw, Poland; agnieszka.matera-witkiewicz@umed.wroc.pl
* Correspondence: Jakub.lach@biol.uni.lodz.pl

Abstract: Halophiles, the salt-loving organisms, have been investigated for at least a hundred years. They are found in all three domains of life, namely Archaea, Bacteria, and Eukarya, and occur in saline and hypersaline environments worldwide. They are already a valuable source of various biomolecules for biotechnological, pharmaceutical, cosmetological and industrial applications. In the present era of multidrug-resistant bacteria, cancer expansion, and extreme environmental pollution, the demand for new, effective compounds is higher and more urgent than ever before. Thus, the unique metabolism of halophilic microorganisms, their low nutritional requirements and their ability to adapt to harsh conditions (high salinity, high pressure and UV radiation, low oxygen concentration, hydrophobic conditions, extreme temperatures and pH, toxic compounds and heavy metals) make them promising candidates as a fruitful source of bioactive compounds. The main aim of this review is to highlight the nucleic acid sequencing experimental strategies used in halophile studies in concert with the presentation of recent examples of bioproducts and functions discovered in silico in the halophile's genomes. We point out methodological gaps and solutions based on in silico methods that are helpful in the identification of valuable bioproducts synthesized by halophiles. We also show the potential of an increasing number of publicly available genomic and metagenomic data for halophilic organisms that can be analysed to identify such new bioproducts and their producers.

Keywords: halophiles; biomolecules; metagenomics; bioinformatics; genome mining; biodiversity; hypersaline environments

Citation: Lach, J.; Jęcz, P.; Strapagiel, D.; Matera-Witkiewicz, A.; Stączek, P. The Methods of Digging for "Gold" within the Salt: Characterization of Halophilic Prokaryotes and Identification of Their Valuable Biological Products Using Sequencing and Genome Mining Tools. *Genes* **2021**, *12*, 1756. https://doi.org/10.3390/genes12111756

Academic Editors: Célia M. Manaia, Olga Nunes, Balbina Nogales and Silvia Turroni

Received: 18 August 2021
Accepted: 30 October 2021
Published: 1 November 2021

1. Introduction

Halophiles are a highly miscellaneous class of extremophilic organisms characterised by their requirements for high salinity and comprise entities from all three domains of life, namely Bacteria, Archaea, and Eukarya [1,2]. Owing to their phylogenetic origin and the nourishment acquisition manner, halophilic microorganisms can be grouped as follows: (1) heterotrophic, phototrophic or methanogenic archaea; (2) heterotrophic, lithotrophic or photosynthetic bacteria, and (3) heterotrophic or photosynthetic eukaryotes [2–4]. In the following parts of this article, we will focus only on the halophiles representing the first two mentioned groups.

Due to the salt concentration requirements (specifically and commonly, sodium cations and chloride anions), halophiles can be classified as slight, with optimal growth at 0.2–0.85M (1–5%) NaCl, moderate thriving in 0.85–3.4M (5–20%) NaCl, and extreme growing optimally at 3.4–5.1M (20–30%) [5]. On the contrary, non-halophiles do not grow

in the environment containing above 0.2M (1%) NaCl, and halotolerant organisms are viable in the presence or absence of highly saline conditions, but it is not necessary for their optimal growth [5]. Moreover, halophilic and halotolerant organisms are able to adapt to a broad range of salt concentrations, occurring seasonally, annually, or irregularly in their natural environments [4,6].

Halophilic microorganisms are forced to efficiently prevent osmosis due to the high external salinity, and thus they have evolved two types of strategies to struggle with cellular water loss—"salt-out" ("low-salt-in") and "salt-in" [4,7]. The first strategy is based on biosynthesis (*de novo* or from the storage substances) or absorption from the environment compatible solutes (osmolytes or osmoprotectants) and is utilized mainly by moderate halophiles, halotolerant bacteria, and eukaryotes. Polyols, sugars, amino acids, betaines, ectoines, N-acetylated diamino acids, and N-derivatized carboxamides of glutamine are commonly used. The second strategy relies on the accumulation of salt, predominantly potassium chloride, to provide intracellular osmotic pressure comparable to the external one and is typical for extremely halophilic Archaea and a few representatives of Bacteria (genus *Salinibacter* and members of the order *Halanaerobiales*) [2]. This mechanism requires specific adaptation of enzymes and other proteins, e.g., by elevating a level of negatively charged amino acids, leading to the formation of an acidic proteome as observed in *Halobacterium* sp. NRC-1 [7–9]. However, the evidence provided by Elevi Bardavid and Oren (2012) suggests that this may not be a strict rule, and other mechanisms must be involved in osmoregulation in halophiles [10]. Some halophiles (especially from the archaeal class of *Halobacteria*) have applied a mix of these strategies to cope mainly with periodic fluctuation of salinity [11,12].

2. Global Distribution of Hypersaline Environments

Although the oceans and seas (average salinity–0.6 M, 3.5% or 35 parts per thousand) come to mind first, the term "hypersaline environments" refers to conditions where the salt concentration exceeds that present in marine basins (even ten times up to or above salt saturation) [13,14]. Hypersaline environments are generally classified as thalassic (thalassohaline) when originating from seawater with its characteristic ionic composition (dominated by Cl^-—49% and Na^+—42% of the total molarity) and as athalassic (athalassohaline, also inland or epicontinental) not directly associated with a marine source and dominated by divalent ions mainly Mg^{2+} and Ca^{2+} [5]. Some authors also distinguish the third type—artificial reservoirs employed for salt production (saltern crystalliser ponds) [15].

Nevertheless, hypersaline ecosystems and their habitats are widely explored mainly due to their utilisation in mineral processing—salt mines, solar salterns and salt flat [16–18], aquaculture (e.g., brine-shrimps predominantly in the Great Salt Lake, commercial lakes in China, Russia, and Kazakhstan) [19,20], biotechnical applications (biomolecules like enzymes, pigments, antimicrobial agents, nanoparticles) [21–24]; in the role of microbial cell factors [25]; environmental and protection studies as niches for eukaryotes, prokaryotes, and archaea [15,17,26,27], biodegradation of contaminants [28–31]; astrobiological signification and early Earth connotations [32,33]. On the other hand, issues related to the anthropogenic impact on hypersaline environments have become more and more significant in recent years. To name only some: climate alteration, overexploitation of mining and mineral extraction, overflow of agriculture, water diversion and salinity enlargement, urban overdevelopment, industrial sewage and contamination with ultimate examples of the Dead Sea, the Caspian Sea, the Aral Sea, and the Great Salt Lake [15]. And as it turns out, these activities have a tremendous influence on the (bio)diversity of the hypersaline ecosystems.

3. Biodiversity of Hypersaline Environments

The Dutch microbiologist and botanist investigating various saline and hypersaline lakes worldwide Lourens G. M. Baas Becking (1895-1963) claimed that "*everything is ev-*

erywhere: but, *the environment selects*" [34]. This statement is highly relevant to the hypersaline ecosystems broadly distributed around the world, from the Antarctic to the Himalayas, from Australia to the USA, from Africa to South America, and thus are much dissimilar in terms of salt concentration, chemical composition, and presence of additional stress conditions designated by geological attributes [5,35–37]. Therefore, they are not only characterised by high-salt content but other environmental physicochemical extrema like high pressure and UV radiation, low oxygen concentration, hydrophobic conditions, extreme temperatures and pH, high concentrations of toxic compounds and heavy metals [27,36,38–42].

The most frequently identified bacterial phyla in saline and hypersaline environments are Actinobacteria, Bacteroidetes, Cyanobacteria, Proteobacteria (Alpha, Beta, Gamma, and Delta), and Firmicutes [36,41,43–48]. Halophilic Archaea are typically represented by Halobacteria and methylotrophic methanogens class members, both belonging to the phylum Euryarchaeota. The former includes about 70 genera and 260 species, classified in three orders and six families: the *Halobacteriales* (families *Halobacteriaceae*, *Haloarculaceae*, *Halococcaceae*), the *Haloferacales* (families *Haloferacaceae*, *Halorubraceae*), and the *Natrialbales* (family *Natrialbaceae*), and the latter comprises of 4 classes: *Methanomicrobia*, *Methanobacteria*, *Methanopyri* and *Methanococci* [49,50]. Furthermore, it was demonstrated that Archaea tend to dominate Bacteria as salinity increases, which is illustrated by an excellent example of the two arms of the Great Salt Lake, significantly different in the salt content, and thus in a taxonomy of their inhabitants [36,37,51,52]. Moreover, the composition and structure of halophilic communities in saline and hypersaline ecosystems are considerably influenced by the salinity fluctuation in time or geographical location and may differ between the places of sampling within the same setting [53–57].

Saline soils are other fascinating and valuable from the ecological, economical, and biotechnological points of view examined environments with abundance and high diversity of their inhabitants, taxonomically comparable to aqueous ones (phylum level) [58,59]. In addition, it has been established that salinity, along with pH and electrical conductivity (EC), are the pivotal factors determining the variety and arrangement of halophiles and haloalkaliphiles in saline soils [60–62]. Intriguingly, these microorganisms are gaining special attention due to progressing global soil salinisation, and thus their potential applicability as plant symbionts enabling and increasing crop productivity in saline soils [63,64]. It is noteworthy that a successful attempt was currently done to employ halophilic microorganisms as bioindicators of the soil salt contamination caused by extensive de-icing of roads during harsh winters in Baltimore, Maryland, USA. It became possible since halophiles become persistent members of microbial communities as a result of salting roads for their de-icing during winter months [65].

In addition to these environmental species, there is a constantly extending group of human or human-related halophiles, both bacterial and archaeal [66–69]. Brining, i.e., treating, food with dry salt or a salt solution, is one of the oldest methods to preserve and season the eatables in food processing. There are numerous and continual scientific reports on isolating new halophilic microorganisms, the diversity and properties of halophilic Bacteria and Archaea, as well as genomic analyses from commercial salt [70–73], cheeses [74–77], table olives [78], kimchi (Asian fermented vegetables) [79–81], and shrimp paste [82,83]. Recent years have also brought interest in the halophilic and halotolerant prokaryotes contributing to the human gut microbiota [73,84–86]. This attention results in part from observing a hazardous tendency to consume increasing amounts of salt delivered with food and its tremendous consequences on human health, including obesity, hypertension, cardiovascular disorders, and stomach cancers [69].

Finally, halophilic prokaryotes are the established producers of multiple biomolecules and chemical substances, predominantly osmolytes, hydrolytic enzymes, and pigments (e.g., carotenoids) [22,87–90]. The increasing interest in compounds and proteins of halophilic origin results mainly from the fact that they remain active under harsh conditions like high salinity, extreme temperatures, and ultimate pH [91]. Moreover, halophilic

enzymes retain solubility and solvation in low water activity [11], and as has been shown recently, they demonstrate anti-desiccation and antifreeze properties, so desirable in food processing and preservation [88]. Halophiles also produce biodegradable polysaccharides and polymers, potentially replacing environmentally hazardous plastics; glycoproteins are considered promising candidates in nanoparticles synthesis, and gas vesicles are examined in terms of an effective drug delivery system as described thoroughly in a review released by Singh and Singh, 2017 [90]. Despite that, bacterial and particularly archaeal halophiles for decades have been underestimated and unexplored in terms of the ability to produce various bioactive compounds, especially of antimicrobial and anticancer potential. However, due to the rapid development of molecular techniques described in the following part of this review, they turned out to be a promising and rich source of diverse biomolecules of great importance in the ongoing post-antibiotic era that is additionally characterised by the galloping increase of cancer cases [22,91–95]. Due to methodological difficulties, time-consuming and expensive procedures that require frequent optimization and the increasing availability of sequencing, research on halophilic biomolecules are moving more and more towards genomic and metagenomic-based bioinformatics analyses [96–101].

4. Experiment Strategies Used in Halophiles Research

As pointed out, microorganisms living in saline environments are characterized by high diversity, like other extremophiles. A wide range of methods is used to better understand this diversity, including culture-based experiment strategies and culture-independent approaches for direct testing of environmental samples. Studies involving cultivation under laboratory conditions have identified and tested most of the currently known halophiles to obtain their exact characteristics [16,80,102,103]. Through the use of laboratory cultivation of pure bacterial cultures, new bioagents produced by halophiles can also be discovered [95,104,105]. However, these traditional cultivation-based methods do not always work well for halophiles. In the saline ecological niches, numerous species of microorganisms are encountered that cannot be easily cultivated in laboratory conditions. Often, only less than 1% of microorganisms from a given site can be successfully grown in the laboratory [106]. Thus, new methods of optimal and high-throughput culturing of microorganisms, based on the use of modern approaches such as culturomics [107,108] or optimal culture enrichment [109], are developed. However, there are still many microorganisms that cannot be tested through common-based laboratory analysis. It is also worth noting that the methods based on the culturing of environmental microorganisms usually provide not only complete information about the taxonomic composition of a specific microbiome but also about the proportions in an abundance of individual microorganisms. For precise inquiry regarding the biodiversity of microbiomes inhabiting specific ecological niches, it is necessary to use methods that provide insight into the studied environment without culturing the organisms living in it [110,111].

Culture-independent methods are the primary investigation tools for looking into the microbial "dark matter" based on the sequencing of nucleic acids for both entire microbial communities (metagenomics and metatranscriptomics) and single cells (single-cell sequencing) [110,112]. Furthermore, functional metagenomics can be used, where it is possible to screen environmental samples for the production of substances of interest with potential industrial applications [113,114]. Finally, metatranscriptomic analyses enable the assessment of the gene expression in the studied environment. Thanks to their use, it is possible to identify active metabolic pathways in the environmental niche, which gives a unique insight into its characteristics. The best results are achieved by combining metagenomics with metatranscriptomics analysis, where the former one identifies genes in the studied microbiome and allows the creation of MAGs (metagenome-assembled genomes), while the latter one allows assessing the expression level of the detected genes [115,116]. Despite the great potential of obtaining significantly valuable data in this way, still relatively unpopular experiments due to the great difficulties in receiving the appropriate quality material for research from saline environments due to the low available biomass

and a high diversity of the samples are observed [117]. Proteomics also is used in the study of halophilic microorganisms, where gene ontology analysis of some up-regulated proteins determining molecular functions, cellular components or a biological process are identified, but it is usually applied when individual strains are grown under laboratory conditions [118–120]. Recently, there have been attempts to use proteomics to directly analyse environmental samples, but this is difficult and still rarely applied. The main reason for the difficulties in using proteomics is the low biomass of proteins in the analyzed samples [115]. In the following part of this article, we will focus on characterising the methods based on nucleic acid sequencing and their role in understanding the biodiversity and biotechnological potential of microorganisms living in saline environments.

5. In Silico Methods for Identification of Novel Halophiles Bioproducts

For many years, the availability of methods based on NGS (Next Generation Sequencing) has increased. New sequencing methods such as SMRT (Single Molecule Real-Time) from PacBio or nanopore sequencing from Oxford Nanopore Technologies (ONT) have been also invented and developed. All of them have changed the research insight concerning microbiomes of saline environments as well as the analysis of individual strains thanks to the quality of genomic and transcriptomic data improvement [121,122]. Due to the development of the new and improvement of the already existing sequencing methods that prove obtaining much longer reads is possible, where the maximum length of 300 bp (base pairs) for Illumina, 400 bp for MGI and Ion Torrent, 1200 bp for Roche, and in the case of ONT and PacBio, the length of a single read can reach several hundred thousand bp with a median length up to 100 kbp [71,123–126]. However, with the ONT technology, the obstacles regarding a relatively low accuracy of reads compared to the second-generation sequencing platforms and even SMRT is observed. In the case of ONT and SMRT, very high requirements are defined in terms of the quantity and quality of the DNA used to prepare the libraries. To avoid the problem of low accuracy, manufacturers are constantly perfecting their technologies both in terms of equipment, reagents and algorithms used in the base-calling process. These actions lead also to a decrease in the cost of ONT sequencing and help to improve the quality of raw data [127–129]. Data obtained from sequencing are often used for in silico analyses, where the identification of interesting bioproducts in the form of secondary metabolites, AMP, or enzymes can be performed. For this intended purpose, some bioinformatics methods are used, which are characterised in the following part of this section.

One of the most explored areas in the field of bioproducts prediction from sequencing data is the identification of biosynthetic genes clusters (BGCs). This approach permits the identification of the pathways responsible for the production of secondary metabolites. BGC identification methods are continuously improved. One of the most intensively developed tools in this area is the antiSMASH. The currently released version 6.0 helps to identify 71 types of BGCs including non-ribosomal peptide synthetases (NRPSs), type I and type II polyketide synthases (PKSs), lanthipeptides, lasso peptides, sactipeptides, thiopeptides, bacteriocins, terpenes and much more [130]. For identification of BGCs, the antiSMASH uses the set of "rules" which define the core biosynthetic functions that need to exist in a genomic region to constitute a BGC. Additionally, a rule-independent method based on the ClusterFinder prediction algorithm was implemented to this platform [130]. The antiSMASH is widely used in halophiles research for the identification of BGCs in genomic and metagenomic data with satisfying results. The combination of rule-based and rule-independent methods in a single platform is one of the biggest advantages of antiSMASH in halophiles research. Many BGCs responsible for the synthesis of valuable halophilic bioagents already have been described, which allows the identification of their variants using the rule-based approach with high specificity. However, halophile genomes most likely also contain a large number of unknown BGCs, in the identification of which a rule-independent approach can be very helpful.

As an example of BGCs identification, *Planococcus maritimus* SAMP MCC 3013 may be mentioned, where a BGC for biosurfactant synthesis was identified [131]. In this genome, researchers discovered a complete gene cluster comprising genes *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, responsible for the biosynthesis of terpenes. Based on genomic and functional analysis, the authors coined the term "terpene containing biosurfactant" for the surfactant produced by *Planococcus maritimus* [131]. Analysis of *Planococcus maritimus* SAMP MCC 3013 biosurfactant shows that the compound is active against *Mycobacterium tuberculosis* (IC50 64.11 ± 1.64 µg/mL and MIC 160.8 ± 1.64 µg/mL), *Plasmodium falciparum* (EC50 34 ± 0.26 µM), and cancer cell lines HeLa (IC50 41.41 ± 4.21 µg/mL), MCF-7 (IC50 42.79 ± 6.07 µg/mL) and HCT (IC50 31.233 ± 5.08 µg/mL). This activity suggests its potential as a lead candidate for drug development investigations [132]. Also, other halophilic microorganisms were analyzed similarly and were identified as potential new producers of highly valuable biosurfactants. *Bacillus amyloliquefaciens* was identified as a producer of surfactin and fengycin. Approximately 37.7 kb fengycin biosynthetic gene cluster with 90% similarity to fengycin cluster of DSM 23117 containing *FenA, FenB, FenC, FenD, and FenE* genes was identified in the genome of *Bacillus amyloliquefaciens* [133]. Interesting BGCs were also identified in metagenomes such as the Shark Bay Microbial Mats metagenome where BGC responsible for syringomycin production was identified. In this study, 17 of 28 condensation domains from MAGs belonging to Chloroflexi phylum, aligned closely with syringomycin [134]. Searching for novel halophilic biosurfactants is a promising field of research due to their potential as antimicrobial agents, bio-detoxifiers, and emulsifiers [105,135–138]. Halophiles producing biosurfactants can also find application in bioremediation, especially in high-salinity conditions [138,139]. Therefore, the identification and in-depth characterisation of halophiles capable of producing biosurfactants are crucial.

Other important compounds produced by halophiles are ectoine and hydroxyectoine. Compatible solutes such as ectoine and hydroxyectoine are used as protective factors for cells, DNA, and proteins, which makes them highly valuable for cosmetics and medicine sectors [140,141]. Ectoine is commonly used in the cosmetics industry, but clinical trials are also conducted to indicate its importance in the treatment of eye and respiratory diseases [142–145]. In silico studies for new microorganisms producing ectoine and hydroxyectoine can be improved by using BGC identification methods like antiSMASH [146–148]. Microorganisms identified in this way can be an alternative to those currently used. They can also combine the production of ectoine with other bioproducts, such as polyhydroxyalkanoates (PHAs) [147]. As an example of this type of co-production *Salinivibrio proteolyticus* M318 isolated from fermented shrimp paste can be mentioned [147]. The authors identified the first strain of *Salinivibrio proteolyticus*, where the complete phaCAB and teaABCD operons were present in the genome. The phaCAB operon is associated with the production of PHAs. It consists of three genes: polyhydroxyalkanoate synthase (*phaC-orf00667*), acetyl-CoA C-acyltransferases (*phaA-orf00669*), and acetoacetyl-CoA reductase (*phaB-orf00670*), while teaABCD operon is associated with the synthesis of ectoine. It includes the *ectA*, *ectB*, *ectC*, and *lysC* genes responsible for synthesising ectoine and ectoine hydroxylase gene (*ectD-orf00133*) accountable for producing hydroxyectoine from ectoine [147]. Thanks to the combined production of PHAs with ectoine and hydroxyectoine, *Salinivibrio proteolyticus* M318 is a very valuable strain that can be used broadly in the industrial production of these substances.

BGCs identification can be useful also in the prediction of non-ribosomal produced antimicrobial peptides. The case with streptomonomicin (STM) produced by *Streptomonospora alba* can be proposed. STM is a lasso peptide with antimicrobial properties. It belongs to the class of ribosomally synthesized and post-translationally modified peptides (RiPPs) [149]. STM inhibits the growth of Bacillus anthracis, the causative agent of anthrax, and its MIC was in the range of 4-8 µg/mL. Genomes analysis of bacteria isolated from salt-pans with Plant Growth Promoting Features (microorganisms naturally capable of enhancing plant growth and protecting crops from pests) also determine antimicrobial agent encoding

BGCs. Here, identified BGCs were similar to bacteriocin encoding BGCs [150] AntiSMASH was also utilized in the search for other BGCs, such as those responsible for the production of exopolysaccharide and pigment source [151], Persiamycin A [152], or carotenoids [153]. Due to its versatility, antiSMASH performs well in halophilic analysis, but it is usually combined with other rule-based or rule-independent tools for better predictions.

The use of in silico methods for BGCs identification helps in more effective quest of valuable compounds new producers. It is particularly important given the rapid increase in the number of publicly available halophilic genomes and metagenomes. Such data can be screened for the presence of BGCs and then analysed more thoroughly if show significant biotechnological or pharmaceutical potential [154]. It is worth to point out that rule-based methods for identifying BGCs require high-quality input assemblies [155]. When low-quality assemblies or metagenomic data, characterized by low continuity and significant fragmentation, are obtained, tools based on the identification of individual core domains, such as NaPDos or eSNaPD [156,157], are preferable. Also, biosyntheticSPAdes to generate better quality input data for BGCs analysis based on de novo assembly graph analysis can be an alternative solution [158]. The growing number and quality of databases containing information on BGCs also have a positive effect on the prospects for the development of BGC identification methods [159–161].

Identification of novel valuable enzymes can be also achieved using simple homology-based approaches. In this case, query protein sequences are aligned to enzyme reference databases such as Expasy ENZYME [162], BRENDA [163], KEGG [164] or protein domains and motifs databases like Pfam [165], TIGRFAM [166] or SMART [167]. These approaches can be used for in silico screening and enzymes sequence characterisation identified in laboratory conditions. Many halophilic microorganisms produce hydrolytic enzymes highly tolerant to salinity, thermostable and stable in a broad pH spectrum [168]. Moreover, halophilic enzymes have not only a significant economic value but are also utilized in eco-friendly industrial processes, enhancing their role in sustainable development [24,169,170]. Using simple homology-based comparison can help understand enzymes structure, properties and similarity to enzymes from other taxon's. As an example, Hypersaline lake "Acıgöl" esterase (hAGEst) and GH11 xylanase Xyn22 can be mentioned [113,114]. Both them were identified and purified based on functional metagenomics methods. Newly discovered hAGEst esterase is similar to the alpha/beta hydrolase from *Halomonas gudaonensis* (WP_089686035.1) with a 91% amino acids sequence identity. It is characterised by high activity at low temperatures, high tolerance to DMSO, and metal ions (majority of analysed metal ions in a concentration of 1 mM improve the activity of the esterase where in the case of other esterases, the presence of these ions cause a reduction in the activity of an enzyme). These features can be important in various industrial applications such as detergent formulations, pharmaceutical production and other fine chemicals [113]. On the other hand, GH11 xylanase Xyn22 has very high halotolerance and thermal stability. GH11 xylanase coding sequence was identified in the metagenomic DNA of a saline-alkaline soil. Based on the amino acid sequence analysis and site-directed mutagenesis, it was shown that acidic amino acid residues E137 and E139 are responsible for halotolerance, while the aromatic interaction between Y48 and F53 is responsible for thermostability. Thanks to their properties, specific enzyme can be used in various fields, like seafood processing, paper industry, or biofuel production, and is a good foundation for further work on modified enzymes of this type [114]. Another example is a novel metagenome-derived halotolerant cellulase PersiCel3 obtained from rumen microbiota can be mentioned [171]. In this case, for new enzyme identification, a multi-stage in silico screening pipeline was employed. Analyses were based on NCBI BLAST alignment of predicted ORFs to sequences from the custom database containing experimentally validated thermostable and/or halotolerant cellulase sequences selected by literature mining. The existence of cellulase domain in predicted genes was confirmed by alignment to sequences from NCBI Conserved Domains Database (CDD) [172]. Maximum activity of the PersiCel3 could be seen in the concentration of 3 M NaCl for both free (132.46%) and immobilized (197.47%) enzymes. Applying

both the free and immobilized enzyme during the degradation of the rice straw in saline conditions leads to an increase in the production of reducing sugars [171]. In the field of microbial enzymes identification, more sophisticated, machine-learning-based tools were also developed, but they have not gained much popularity in halophilic research yet [173–175]. Most likely, it is related to the simplicity and clarity of the standard analytical pipelines. Another group of in silico methods is associated with ribosomally synthesised antimicrobial peptides' (AMPs) identification. Few databases and tools dedicated to AMPs identification and prediction can be found. Some general antimicrobial peptides databases like DRAMP [176], CRAMPR [177], APD [178], LAMP [179], and numerous specific for different types of AMPs databases [180] are publicly available for researchers. AMPs identification can usually be performed as a simple sequence alignment process [181], but also more sophisticated methods based on machine learning can be used [182–185]. The direct identification of AMPs is used rather rarely in halophile genomic and metagenomic sequences. Only a few examples of identification of genes coding AMPs after their wet lab detection have been described [181,186]. Unfortunately, screening of genomic and metagenomic data for AMPs genes identification is rarely performed in the analysis of halophiles. It is a significant gap, especially because the halophiles are well-known producers of AMPs [186–190], and learning more of their products through in silico analysis could lead to the discovery of new effective antibiotics. Performing an extensive screening of AMPs may allow the identification of new valuable bioproducts and halophilic strains that may be their producers, as is the case with other groups of microorganisms [191]. In the context of the rising problem with antibiotic resistance, the identification of novel AMPs to which bacteria do not easily develop resistance may be the last hope [192].

The summary of information about selected bioinformatics tools used to identify new bioproducts in genomes of halophilic microorganisms and saline environment metagenomes has been presented in Table 1.

Table 1. Strengths and weaknesses of methods used in the in silico identification of bioproducts.

Target	Class of Methods	Bioinformatic Tool Examples	Advantages	Disadvantages
Biosynthetic Gene Clusters	Rule-based BGCs identification	antiSMASH [130]	High number of identified BGC classes (71 in antiSMASH 6.0 version) Manually curated rules User-friendly pipelines (e.g., web interface of antiSMASH)	Requires high-quality assemblies Limited to BGCs for which rules have been implemented
		PRISM [193]		
		SMURF [194]		
		BAGEL [195]		
	Rule-independent BGCs identification	eSNaPD [157]	Can be used for highly fragmented data (e.g., metagenomes) A potential to identify novel BGCs (e.g., EvoMining)	Less specific for known BGCs then rule-based methods
		NaPDos [156]		
		EvoMining [196]		
		ClusterFinder [197]		
Halophilic enzymes	Classic alignment based approach	Expasy Enzyme [162]	The user is not limited by a predefined set of databases and comparison parameters User-friendly web interface and stand-alone versions Very flexible	Requires a combination of tools for the best effect In some cases, a choice of optimal pipeline can be challenging
		Pfam [165]		
		BRENDA [163]		
		KEGG [164]		
		BLAST [198]		

Table 1. *Cont.*

Target	Class of Methods	Bioinformatic Tool Examples	Advantages	Disadvantages
Ribosomally produced AMP	Automated pipelines	Anastasia [173]	Provides high reproducibility Usually better optimized than classical analytical pipelines Does not require detailed knowledge of all analysis parameters (the default parameters should be appropriate for most common cases)	Limited customization options
		MCIC [174]		
		FINDER [175]		
	Classic alignment-based approach	DRAMP [176]	The user is not limited by a predefined set of databases and comparison parameters Very flexible Any combination of databases for analysis can be selected	Requires a combination of tools for the best effect In some cases a choice of optimal pipeline can be challenging Cannot identify the AMP types that are not included in the database
		CAMPR [177]		
		LAMP [179]		
		APD [178]		
	Automated pipelines	Macrel [182]	Provides high reproducibility Usually better optimized than classical analytical pipelines Does not require detailed knowledge of all analysis parameters (the default parameters should be appropriate for most common cases)	Limited databases for individual tools Cannot identify the AMP types that are not included in the database
		AMAP [184]		
		AmPEP [185]		

6. Conclusions

In recent years, there has been an intensive development of research methods that allow for better insight into the biodiversity of extreme environments, including high salinity environments. New niches such as brine graduation towers or food products have started to be investigated as profoundly as, so far the most analysed, salt environments, such as salt lakes, brines, or the Dead Sea. The use of modern research techniques based both on the culture-dependent and culture-independent methods enables a better understanding of the potential of halophiles and setting new directions for their applications. The research on halophiles is also important since these microorganisms are often polyextremophiles simultaneously, which is associated with their unique properties and may be significant in the context of astrobiological research, where halophiles are often considered organisms that could be detected on distant planets or moons [199].

The intensive development of sequencing methods and related bioinformatics tools opens new doors for identifying compounds of great importance for biotechnology and pharmacy. The falling costs of NGS and the development of third-generation sequencing methods allow us to look at saline environments' previously inaccessible microbial dark matter. However, much remains to be done in the area of developing new analytical tools. It is necessary to conduct further research on the methods of identifying new compounds that may be of significant importance for biotechnology and pharmacy. An important niche is also the identification of absolutely new enzymes, BGCs, or AMPs that do not show substantial similarity to those that have already been known. It is also worth noting that halophilic prokaryotes are rarely analyzed for the presence of ribosomal AMPs in

their genomes. The main focus of genome mining activities is on the identification of BGC. On the other hand, the identification of ribosomal AMPs may be an interesting niche for research due to the fact that there are several examples of AMPs produced by halophiles, and it is quite probable that other compounds of this type will be found in other strains.

It should be noted, however, that despite the need for further development of genome mining methods, so far, thanks to their use, it has been possible to identify many compounds important for humanity, and one can hope that further research will allow for identification of even more of them, which will significantly support green economy and pharmacy. Saline environments and halophiles, thanks to their characteristics, are an important source of new biomolecules identified through the use of the aforementioned methods. Halophiles are known as biotechnologically preferable microorganisms capable of producing heterogenous bioproducts [200]. They have been identified as a relevant source of stable enzymatic proteins, antioxidants, stabilising agents, or antimicrobial compounds such as halocins with additional myocardial protection activity [24]. Research is also being carried out to improve biotechnological applications by utilising genetic manipulations and modifications based on mutagenesis, the introduction of recombinant plasmids, and heterologous gene expression.

In summary, in silico identification of halophile bioproducts is crucial for optimal characterisation of this group of microorganism and their use in the industry. Growing databases containing the genome sequences of the halophilic organisms and metagenomes from saline environments offer the opportunity to perform extensive in silico analyses to identify their so far hidden potential. More efforts should be made to identify new AMPs and their producers among halophiles. This area has been greatly neglected and should be considered more frequently for in silico characterization of the biotechnological potential of halophiles. A positive boost is given by the further development of sequencing technology, leading to improved sequencing data quality. This is particularly important for metagenomic studies, which are a key component in halophile research and facilitate the characterization of the biotechnological potential of uncultivated halophiles.

Author Contributions: All authors have made a substantial, direct, and intellectual contribution to the work and approved the final version of the manuscript. J.L. and P.J. drafted the manuscript, prepared tables, and wrote the final version of the manuscript. A.M.-W., D.S. and P.S. comprehensively edited and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Publication's printing cost was co-financed by the European Union from the European Social Fund under the "InterDOC-STARt" project (POWR.03.02.00-00-I033/16-00).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The presented study was supported by the "InterDOC-STARt" project (POWR.03.02.00-00-I033/16-00) and the MINI InterDOC-STARt grant. Publication's printing cost was co-financed by the European Union from the European Social Fund under the "InterDOC-STARt" project (POWR.03.02.00-00-I033/16-00). The publication was also supported by the subsidy for scientific activity provided to the Department of Molecular Microbiology by the University of Łódź (B2011000000038.01).

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflict of interest.

References

1. Ma, Y.; Galinski, E.A.; Grant, W.D.; Oren, A.; Ventosa, A. Halophiles 2010: Life in Saline Environments. *Appl. Environ. Microbiol.* **2010**, *76*, 6971–6981. [CrossRef] [PubMed]
2. Oren, A. Extremophiles Handbook. *Extrem. Handb.* **2011**, *31*, 1–26. [CrossRef]
3. Andrei, A.-S.; Banciu, H.L.; Oren, A. Living with salt: Metabolic and phylogenetic diversity of archaea inhabiting saline ecosystems. *FEMS Microbiol. Lett.* **2012**, *330*, 1–9. [CrossRef]
4. DasSarma, S.; DasSarma, P. Halophiles. *eLS* **2017**, 1–13. [CrossRef]

5. Ventosa, A.; Arahal, D.R. Physico-chemical characteristics of hypersaline environments and their biodiversity. In *Extremophiles*; CRC Press: Boca Raton, FL, USA, 2009.
6. Vauclare, P.; Natali, F.; Kleman, J.-P.; Zaccai, G.; Franzetti, B. Surviving salt fluctuations: Stress and recovery in Halobacterium salinarum, an extreme halophilic Archaeon. *Sci. Rep.* **2020**, *10*, 3298. [CrossRef]
7. Gunde-Cimerman, N.; Plemenitaš, A.; Oren, A. Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. *FEMS Microbiol. Rev.* **2018**, *42*, 353–375. [CrossRef]
8. Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.R.; Baliga, N.; Thorsson, V.; Sbrogna, J.; et al. Genome sequence of Halobacterium species NRC-1. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 12176–12181. [CrossRef]
9. Fukuchi, S.; Yoshimune, K.; Wakayama, M.; Moriguchi, M.; Nishikawa, K. Unique Amino Acid Composition of Proteins in Halophilic Bacteria. *J. Mol. Biol.* **2003**, *327*, 347–357. [CrossRef]
10. Bardavid, R.E.; Oren, A. Acid-shifted isoelectric point profiles of the proteins in a hypersaline microbial mat: An adaptation to life at high salt concentrations? *Extremophiles* **2012**, *16*, 787–792. [CrossRef] [PubMed]
11. DasSarma, S.; DasSarma, P. Halophiles and their enzymes: Negativity put to good use. *Curr. Opin. Microbiol.* **2015**, *25*, 120–126. [CrossRef]
12. Purdy, K.J.; Cresswell-Maynard, T.D.; Nedwell, D.B.; McGenity, T.J.; Grant, W.D.; Timmis, K.N.; Embley, T.M. Isolation of haloarchaea that grow at low salinities. *Environ. Microbiol.* **2004**, *6*, 591–595. [CrossRef] [PubMed]
13. Green, W.J.; Lyons, W.B. The Saline Lakes of the McMurdo Dry Valleys, Antarctica. *Aquat. Geochem.* **2008**, *15*, 321–348. [CrossRef]
14. Castro, P.; Huber, M.E. *Marine Biology*, 11th ed.; Mcgraw-Hill Education: New York City, NY, USA, 2018; ISBN 978-1-260-08510-5.
15. Paul, V.G.; Mormile, M.R. A case for the protection of saline and hypersaline environments: A microbiological perspective. *FEMS Microbiol. Ecol.* **2017**, *93*, fix091. [CrossRef]
16. Albuquerque, L.; Kowalewicz-Kulbat, M.; Drzewiecka, D.; Stączek, P.; D'Auria, G.; Rosselló-Móra, R.; da Costa, M.S. Halorhabdus rudnickae sp. nov., a halophilic archaeon isolated from a salt mine borehole in Poland. *Syst. Appl. Microbiol.* **2016**, *39*, 100–105. [CrossRef] [PubMed]
17. Gajardo, G.; Redón, S. Andean hypersaline lakes in the Atacama Desert, northern Chile: Between lithium exploitation and unique biodiversity conservation. *Conserv. Sci. Pract.* **2019**, *1*, e94. [CrossRef]
18. Pal, S.; Biswas, R.; Misra, A.; Sar, A.; Banerjee, S.; Mukherjee, P.; Dam, B. Poorly known microbial taxa dominate the microbiome of hypersaline Sambhar Lake salterns in India. *Extremophiles* **2020**, *24*, 875–885. [CrossRef] [PubMed]
19. Anufriieva, E.V. How can saline and hypersaline lakes contribute to aquaculture development? A review. *J. Oceanol. Limnol.* **2018**, *36*, 2002–2009. [CrossRef]
20. Litvinenko, L.I.; Litvinenko, A.I.; Boiko, E.G.; Kutsanov, K. Artemia cyst production in Russia. *Chin. J. Oceanol. Limnol.* **2015**, *33*, 1436–1450. [CrossRef]
21. Abdollahnia, M.; Makhdoumi, A.; Mashreghi, M.; Eshghi, H. Exploring the potentials of halophilic prokaryotes from a solar saltern for synthesizing nanoparticles: The case of silver and selenium. *PLoS ONE* **2020**, *15*, e0229886. [CrossRef]
22. Corral, P.; Amoozegar, M.A.; Ventosa, A. Halophiles and Their Biomolecules: Recent Advances and Future Applications in Biomedicine. *Mar. Drugs* **2019**, *18*, 33. [CrossRef]
23. Yin, J.; Chen, J.C.; Wu, Q.; Chen, G.Q. *Halophiles, Coming Stars for Industrial Biotechnology*; Elsevier B.V.: Amsterdam, The Netherlands, 2015; Volume 33, ISBN 8610627942.
24. Liu, C.; Baffoe, D.K.; Zhan, Y.; Zhang, M.; Li, Y.; Zhang, G. Halophile, an essential platform for bioproduction. *J. Microbiol. Methods* **2019**, *166*, 105704. [CrossRef]
25. Mitra, R.; Xu, T.; Xiang, H.; Han, J. Current developments on polyhydroxyalkanoates synthesis by using halophiles as a promising cell factory. *Microb. Cell Factories* **2020**, *19*, 86. [CrossRef]
26. Oren, A. Biodiversity in highly saline environments. In *Physiology and Biochemistry of Extremophiles*; Wiley: Hoboken, NJ, USA, 2007; pp. 223–231.
27. Ventosa, A. Unusual micro-organisms from unusual habitats: Hypersaline environments. *Prokaryotic Divers. Mech. Significance Publ. Soc. Gen. Microbiol.* **2010**, *66*, 223–254. [CrossRef]
28. Edbeib, M.F.; Wahab, R.A.; Huyop, F. Halophiles: Biology, adaptation, and their role in decontamination of hypersaline environments. *World J. Microbiol. Biotechnol.* **2016**, *32*, 135. [CrossRef]
29. Ibrahim, I.M.; Konnova, S.; Sigida, E.N.; Lyubun, E.V.; Muratova, A.; Fedonenko, Y.P.; Elbanna, K. Bioremediation potential of a halophilic Halobacillus sp. strain, EG1HP4QL: Exopolysaccharide production, crude oil degradation, and heavy metal tolerance. *Extremophiles* **2019**, *24*, 157–166. [CrossRef]
30. John, J.; Siva, V.; Kumari, R.; Arya, A.; Kumar, A. Unveiling Cultivable and Uncultivable Halophilic Bacteria Inhabiting Marakkanam Saltpan, India and Their Potential for Biotechnological Applications. *Geomicrobiol. J.* **2020**, *37*, 691–701. [CrossRef]
31. Kiadehi, M.S.H.; Amoozegar, M.A.; Asad, S.; Siroosi, M. Exploring the potential of halophilic archaea for the decolorization of azo dyes. *Water Sci. Technol.* **2018**, *77*, 1602–1611. [CrossRef] [PubMed]
32. Bryanskaya, A.V.; Berezhnoy, A.A.; Rozanov, A.S.; Serdyukov, D.S.; Malup, T.K.; Peltek, S.E. Survival of halophiles of Altai lakes under extreme environmental conditions: Implications for the search for Martian life. *Int. J. Astrobiol.* **2019**, *19*, 1–15. [CrossRef]
33. DasSarma, P.; Laye, V.; Harvey, J.; Reid, C.; Shultz, J.; Yarborough, A.; Lamb, A.; Koske-Phillips, A.; Herbst, A.; Molina, F.; et al. Survival of halophilic Archaea in Earth's cold stratosphere. *Int. J. Astrobiol.* **2016**, *16*, 321–327. [CrossRef]
34. Becking, L.B. *Geobiologie of Inleiding tot de Milieukunde*; WP Van Stockum & Zoon: Amsterdam, The Netherlands, 1934.

35. Oren, A. Halophilic microbial communities and their environments. *Curr. Opin. Biotechnol.* **2015**, *33*, 119–124. [CrossRef]
36. Shurigin, V.; Hakobyan, A.; Panosyan, H.; Egamberdieva, D.; Davranov, K.; Birkeland, N.-K. A glimpse of the prokaryotic diversity of the Large Aral Sea reveals novel extremophilic bacterial and archaeal groups. *Microbiol. Open* **2019**, *8*, e00850. [CrossRef]
37. Simachew, A.; Lanzén, A.; Gessesse, A.; Øvreås, L. Prokaryotic Community Diversity Along an Increasing Salt Gradient in a Soda Ash Concentration Pond. *Microb. Ecol.* **2015**, *71*, 326–338. [CrossRef]
38. Belilla, J.; Moreira, D.; Jardillier, L.; Reboul, G.; Benzerara, K.; López-García, J.M.; Bertolino, P.; López-Archilla, A.I.; López-García, P. Hyperdiverse archaea near life limits at the polyextreme geothermal Dallol area. *Nat. Ecol. Evol.* **2019**, *3*, 1552–1561. [CrossRef] [PubMed]
39. Edwardson, C.; Hollibaugh, J.T. Composition and Activity of Microbial Communities along the Redox Gradient of an Alkaline, Hypersaline, Lake. *Front. Microbiol.* **2018**, *9*, 14. [CrossRef]
40. Merino, N.; Aronson, H.S.; Bojanova, D.P.; Feyhl-Buska, J.; Wong, M.L.; Zhang, S.; Giovannelli, D. Living at the Extremes: Extremophiles and the Limits of Life in a Planetary Context. *Front. Microbiol.* **2019**, *10*, 780. [CrossRef]
41. Pecher, W.T.; Martínez, F.L.; DasSarma, P.; Guzman, D.; DasSarma, S. 16S rRNA Gene Diversity in the Salt Crust of Salar de Uyuni, Bolivia, the World's Largest Salt Flat. *Microbiol. Resour. Announc.* **2020**, *9*, 16–18. [CrossRef]
42. Williams, T.J.; Allen, M.A.; DeMaere, M.; Kyrpides, N.; Tringe, S.; Woyke, T.; Cavicchioli, R. Microbial ecology of an Antarctic hypersaline lake: Genomic assessment of ecophysiology among dominant haloarchaea. *ISME J.* **2014**, *8*, 1645–1658. [CrossRef] [PubMed]
43. AlBataineh, H.; Stevens, D.C. Marine Myxobacteria: A Few Good Halophiles. *Mar. Drugs* **2018**, *16*, 209. [CrossRef]
44. Gupta, S.; Sharma, P.; Dev, K.; Srivastava, M.; Sourirajan, A. A diverse group of halophilic bacteria exist in Lunsu, a natural salt water body of Himachal Pradesh, India. *SpringerPlus* **2015**, *4*, 274. [CrossRef] [PubMed]
45. Leboulanger, C.; Agogué, H.; Bernard, C.; Bouvy, M.; Carré, C.; Cellamare, M.; Duval, C.; Fouilland, E.; Got, P.; Intertaglia, L.; et al. Microbial Diversity and Cyanobacterial Production in Dziani Dzaha Crater Lake, a Unique Tropical Thalassohaline Environment. *PLoS ONE* **2017**, *12*, e0168879. [CrossRef]
46. Mora-Ruiz, M.R.; Cifuentes, A.; Font-Verdera, F.; Pérez-Fernández, C.; Farias, M.E.; González, B.; Orfila, A.; Rosselló-Móra, R. Biogeographical patterns of bacterial and archaeal communities from distant hypersaline environments. *Syst. Appl. Microbiol.* **2018**, *41*, 139–150. [CrossRef]
47. Ul-Hasan, S.; Bowers, R.M.; Figueroa-Montiel, A.; Licea-Navarro, A.F.; Beman, J.M.; Woyke, T.; Nobile, C.J. Community ecology across bacteria, archaea and microbial eukaryotes in the sediment and seawater of coastal Puerto Nuevo, Baja California. *PLoS ONE* **2019**, *14*, e0212355. [CrossRef] [PubMed]
48. Sorokin, D.Y.; Berben, T.; Melton, E.D.; Overmars, L.; Vavourakis, C.; Muyzer, G. Microbial diversity and biogeochemical cycling in soda lakes. *Extremophiles* **2014**, *18*, 791–809. [CrossRef] [PubMed]
49. McGenity, T. Methanogens and Methanogenesis in Hypersaline Environments. In *Handbook of Hydrocarbon and Lipid Microbiology*; Springer: Berlin, Germany, 2010.
50. Oren, A. The microbiology of red brines. In *Advances in Applied Microbiology*; Elsevier Inc.: Amsterdam, The Netherlands, 2020; Volume 113, pp. 57–110. ISBN 9780128207093.
51. Baxter, B.K.; Zalar, P. The extremophiles of Great Salt Lake: Complex microbiology in a dynamic hypersaline ecosystem. In *Model Ecosystems in Extreme Environments*; Elsevier Inc.: Amsterdam, The Netherlands, 2019; pp. 57–99. [CrossRef]
52. Ben Abdallah, M.; Karray, F.; Kallel, N.; Armougom, F.; Mhiri, N.; Quemeneur, M.; Cayol, J.-L.; Erauso, G.; Sayadi, S. Abundance and diversity of prokaryotes in ephemeral hypersaline lake Chott El Jerid using Illumina Miseq sequencing, DGGE and qPCR assays. *Extremophiles* **2018**, *22*, 811–823. [CrossRef]
53. Almeida-Dalmet, S.; Sikaroodi, M.; Gillevet, P.M.; Litchfield, C.D.; Baxter, B.K. Temporal Study of the Microbial Diversity of the North Arm of Great Salt Lake, Utah, U.S. *Microorganisms* **2015**, *3*, 310–326. [CrossRef]
54. Chen, S.; Xu, Y.; Helfant, L. Geographical Isolation, Buried Depth, and Physicochemical Traits Drive the Variation of Species Diversity and Prokaryotic Community in Three Typical Hypersaline Environments. *Microorganisms* **2020**, *8*, 120. [CrossRef]
55. Di Meglio, L.; Santos, F.; Gomariz, M.; Almansa, C.; López, C.; Anton, J.; Nercessian, D. Seasonal dynamics of extremely halophilic microbial communities in three Argentinian salterns. *FEMS Microbiol. Ecol.* **2016**, *92*. [CrossRef]
56. Haferburg, G.; Gröning, J.A.D.; Schmidt, N.; Kummer, N.-A.; Erquicia, J.C.; Schlömann, M. Microbial diversity of the hypersaline and lithium-rich Salar de Uyuni, Bolivia. *Microbiol. Res.* **2017**, *199*, 19–28. [CrossRef]
57. Kalwasińska, A.; Deja-Sikora, E.; Burkowska-But, A.; Szabó, A.; Felföldi, T.; Kosobucki, P.; Krawiec, A.; Walczak, M. Changes in bacterial and archaeal communities during the concentration of brine at the graduation towers in Ciechocinek spa (Poland). *Extremophiles* **2017**, *22*, 233–246. [CrossRef]
58. Canfora, L.; Bacci, G.; Pinzari, F.; Papa, G.L.; Dazzi, C.; Benedetti, A. Salinity and Bacterial Diversity: To What Extent Does the Concentration of Salt Affect the Bacterial Community in a Saline Soil? *PLoS ONE* **2014**, *9*, e106662. [CrossRef] [PubMed]
59. Delgado-García, M.; Contreras-Ramos, S.M.; Rodríguez, J.A.; Mateos-Díaz, J.C.; Aguilar, C.N.; Camacho-Ruíz, R.M. Isolation of halophilic bacteria associated with saline and alkaline-sodic soils by culture dependent approach. *Heliyon* **2018**, *4*, e00954. [CrossRef] [PubMed]
60. Mukhtar, S.; Mehnaz, S.; Malik, K.A. Microbial diversity in the rhizosphere of plants growing under extreme environments and its impact on crop improvement. *Environ. Sustain.* **2019**, *2*, 329–338. [CrossRef]

61. Wang, S.; Sun, L.; Ling, N.; Zhu, C.; Chi, F.; Li, W.; Hao, X.; Zhang, W.; Bian, J.; Chen, L.; et al. Exploring Soil Factors Determining Composition and Structure of the Bacterial Communities in Saline-Alkali Soils of Songnen Plain. *Front. Microbiol.* **2020**, *10*, 2902. [CrossRef]
62. Zhao, R.; Feng, J.; Yin, X.; Liu, J.; Fu, W.; Berendonk, T.U.; Zhang, T.; Li, X.; Li, B. Antibiotic resistome in landfill leachate from different cities of China deciphered by metagenomic analysis. *Water Res.* **2018**, *134*, 126–139. [CrossRef]
63. Acuña-Rodríguez, I.S.; Hansen, H.; Gallardo-Cerda, J.; Atala, C.; Molina-Montenegro, M.A. Antarctic Extremophiles: Biotechnological Alternative to Crop Productivity in Saline Soils. *Front. Bioeng. Biotechnol.* **2019**, *7*, 1–13. [CrossRef]
64. Otlewska, A.; Migliore, M.; Dybka-Stępień, K.; Manfredini, A.; Struszczyk-Świta, K.; Napoli, R.; Białkowska, A.; Canfora, L.; Pinzari, F. When Salt Meddles Between Plant, Soil, and Microorganisms. *Front. Plant Sci.* **2020**, *11*, 553087. [CrossRef]
65. Pecher, W.T.; Al Madadha, M.E.; DasSarma, P.; Ekulona, F.; Schott, E.J.; Crowe, K.; Gut, B.S.; DasSarma, S. Effects of road salt on microbial communities: Halophiles as biomarkers of road salt pollution. *PLoS ONE* **2019**, *14*, e0221355. [CrossRef]
66. Kim, J.Y.; Whon, T.W.; Lim, M.Y.; Kim, Y.B.; Kim, N.; Kwon, M.-S.; Kim, J.; Lee, S.H.; Choi, H.-J.; Nam, I.-H.; et al. The human gut archaeome: Identification of diverse haloarchaea in Korean subjects. *Microbiome* **2020**, *8*, 114. [CrossRef]
67. Lagier, J.-C.; Khelaifia, S.; Azhar, E.; Croce, O.; Bibi, F.; Jiman-Fatani, A.A.; Yasir, M.; Ben Helaby, H.; Robert, C.; Fournier, P.-E.; et al. Genome sequence of Oceanobacillus picturae strain S1, an halophilic bacterium first isolated in human gut. *Stand. Genom. Sci.* **2015**, *10*, 91. [CrossRef] [PubMed]
68. Rodriguez-Medina, J.; Kim, H.G.; Castro, J.; Contreras, C.M.; Glon, C.L.; Goyal, A.; Guo, B.Y.; Knowles, S.; Lin, J.C.; McGuiness, C.L.; et al. Draft Genome Sequences of 16 Halophilic Prokaryotes Isolated from Diverse Environments. *Microbiol. Resour. Announc.* **2020**, *9*, 20–22. [CrossRef] [PubMed]
69. Seck, E.H.; Senghor, B.; Merhej, V.; Bachar, D.; Cadoret, F.; Robert, C.; Azhar, E.; Yasir, M.; Bibi, F.; Jiman-Fatani, A.A.; et al. Salt in stools is associated with obesity, gut halophilic microbiota and Akkermansia muciniphila depletion in humans. *Int. J. Obes.* **2018**, *43*, 862–871. [CrossRef] [PubMed]
70. Enomoto, S.; Shimane, Y.; Ihara, K.; Kamekura, M.; Itoh, T.; Ohkuma, M.; Takahashi-Ando, N.; Fukushima, Y.; Yoshida, Y.; Usami, R.; et al. Haloarcula mannanilytica sp. nov., a galactomannan-degrading haloarchaeon isolated from commercial salt. *Int. J. Syst. Evol. Microbiol.* **2020**, *70*, 6331–6337. [CrossRef] [PubMed]
71. Gibtan, A.; Park, K.; Woo, M.; Shin, J.-K.; Lee, D.-W.; Sohn, J.H.; Song, M.; Roh, S.W.; Lee, S.-J.; Lee, H.-S. Diversity of Extremely Halophilic Archaeal and Bacterial Communities from Commercial Salts. *Front. Microbiol.* **2017**, *8*, 799. [CrossRef]
72. Minegishi, H.; Echigo, A.; Kuwahara, A.; Shimane, Y.; Kamekura, M.; Itoh, T.; Ohkuma, M.; Usami, R. Halocalculus aciditolerans gen. nov., sp. nov., an acid-tolerant haloarchaeon isolated from commercial salt. *Int. J. Syst. Evol. Microbiol.* **2015**, *65*, 1640–1645. [CrossRef]
73. Seck, E.H.; Dufour, J.-C.; Raoult, D.; Lagier, J.-C. Halophilic & halotolerant prokaryotes in humans. *Futur. Microbiol.* **2018**, *13*, 799–812. [CrossRef]
74. Anast, J.M.; Dzieciol, M.; Schultz, D.L.; Wagner, M.; Mann, E.; Schmitz-Esser, S. Brevibacterium from Austrian hard cheese harbor a putative histamine catabolism pathway and a plasmid for adaptation to the cheese environment. *Sci. Rep.* **2019**, *9*, 6164. [CrossRef]
75. Ishikawa, M.; Kodama, K.; Yasuda, H.; Okamoto-Kainuma, A.; Koizumi, K.; Yamasato, K. Presence of halophilic and alkaliphilic lactic acid bacteria in various cheeses. *Lett. Appl. Microbiol.* **2006**, *44*, 308–313. [CrossRef] [PubMed]
76. Suzuki, T.; Matsutani, M.; Matsuyama, M.; Unno, R.; Matsushita, H.; Sugiyama, M.; Yamasato, K.; Koizumi, Y.; Ishikawa, M. Growth and metabolic properties of halophilic and alkaliphilic lactic acid bacterial strains of Marinilactibacillus psychrotolerans isolated from surface-ripened soft cheese. *Int. Dairy J.* **2020**, *112*, 104840. [CrossRef]
77. Jonnala, B.R.Y.; McSweeney, P.L.H.; Sheehan, J.J.; Cotter, P.D. Sequencing of the Cheese Microbiome and Its Relevance to Industry. *Front. Microbiol.* **2018**, *9*, 1020. [CrossRef]
78. Yalçınkaya, S.; Kılıç, G.B. Isolation, identification and determination of technological properties of the halophilic lactic acid bacteria isolated from table olives. *J. Food Sci. Technol.* **2019**, *56*, 2027–2037. [CrossRef] [PubMed]
79. Jang, J.-Y.; Oh, Y.J.; Lim, S.K.; Park, H.K.; Lee, C.; Kim, J.Y.; Lee, M.-A.; Choi, H.-J. Salicibibacter kimchii gen. nov., sp. nov., a moderately halophilic and alkalitolerant bacterium in the family Bacillaceae, isolated from kimchi. *J. Microbiol.* **2018**, *56*, 880–885. [CrossRef] [PubMed]
80. Oh, Y.J.; Kim, J.Y.; Park, H.K.; Jang, J.-Y.; Lim, S.K.; Kwon, M.-S.; Choi, H.-J. Salicibibacter halophilus sp. nov., a moderately halophilic bacterium isolated from kimchi. *J. Microbiol.* **2019**, *57*, 997–1002. [CrossRef] [PubMed]
81. Oh, Y.J.; Kim, J.Y.; Jo, H.E.; Park, H.K.; Lim, S.K.; Kwon, M.-S.; Choi, H.-J. Lentibacillus cibarius sp. nov., isolated from kimchi, a Korean fermented food. *J. Microbiol.* **2020**, *58*, 387–394. [CrossRef]
82. Booncharoen, A.; Visessanguan, W.; Kuncharoen, N.; Yiamsombut, S.; Santiyanont, P.; Mhuantong, W.; Charoensri, S.; Rojsitthisak, P.; Tanasupawat, S. Lentibacillus lipolyticus sp. nov., a moderately halophilic bacterium isolated from shrimp paste (Ka-pi). *Int. J. Syst. Evol. Microbiol.* **2019**, *69*, 3529–3536. [CrossRef]
83. Li, K.; Sang, X.; Zhu, Y.; Zhang, G.; Bi, J.; Hao, H.; Hou, H.; Qian, F. Lentibacillus panjinensis sp. nov., Isolated from Shrimp Paste, a Traditional Chinese Fermented Seafood. *Curr. Microbiol.* **2020**, *77*, 1997–2001. [CrossRef]
84. Diop, A.; Seck, E.H.; Dubourg, G.; Armstrong, N.; Blanc-Tailleur, C.; Raoult, D.; Fournier, P. Genome sequence and description ofGracilibacillus timonensissp. nov. strain Marseille-P2481T, a moderate halophilic bacterium isolated from the human gut microflora. *Microbiol. Open* **2018**, *8*, e00638. [CrossRef] [PubMed]

85. Khelaifia, S.; Lagier, J.-C.; Bibi, F.; Azhar, E.I.; Croce, O.; Padmanabhan, R.; Jiman-Fatani, A.A.; Yasir, M.; Robert, C.; Andrieu, C.; et al. Microbial Culturomics to Map Halophilic Bacterium in Human Gut: Genome Sequence and Description of Oceanobacillus jeddahense sp. nov. *OMICS A J. Integr. Biol.* **2016**, *20*, 248–258. [CrossRef]
86. Ngom, I.I.; Hasni, I.; Senghor, B.; Lo, C.I.; Armstrong, N.; Sokhna, C.; Raoult, D.; Fournier, P.-E.; Lagier, J.-C. Description of Gracilibacillus phocaeensis sp. nov., a new halophilic bacterium isolated from Senegalian human stool. *New Microbes New Infect.* **2020**, *38*, 100799. [CrossRef]
87. Amoozegar, M.A.; Siroosi, M.; Atashgahi, S.; Smidt, H.; Ventosa, A. Systematics of haloarchaea and biotechnological potential of their hydrolytic enzymes. *Microbiology* **2017**, *163*, 623–645. [CrossRef]
88. Desai, C.; Patel, P.; Markande, A.R.; Kamala, K.; Sivaperumal, P. Exploration of haloarchaea for their potential applications in food industry. *Int. J. Environ. Sci. Technol.* **2020**, *17*, 4455–4464. [CrossRef]
89. Rodrigo-Baños, M.; Garbayo, I.; Vílchez, C.; Bonete, M.-J.; Martínez-Espinosa, R.M. Carotenoids from Haloarchaea and Their Potential in Biotechnology. *Mar. Drugs* **2015**, *13*, 5508–5532. [CrossRef]
90. Singh, A.; Singh, A.K. Haloarchaea: Worth exploring for their biotechnological potential. *Biotechnol. Lett.* **2017**, *39*, 1793–1800. [CrossRef]
91. Gómez-Villegas, P.; Vigara, J.; Vila, M.; Varela, J.; Barreira, L.; Léon, R. Antioxidant, Antimicrobial, and Bioactive Potential of Two New Haloarchaeal Strains Isolated from Odiel Salterns (Southwest Spain). *Biology* **2020**, *9*, 298. [CrossRef] [PubMed]
92. Ghanmi, F.; Carré-Mlouka, A.; Zarai, Z.; Mejdoub, H.; Peduzzi, J.; Maalej, S.; Rebuffat, S. The extremely halophilic archaeon Halobacterium salinarum ETD5 from the solar saltern of Sfax (Tunisia) produces multiple halocins. *Res. Microbiol.* **2019**, *171*, 80–90. [CrossRef]
93. Makarova, K.S.; Wolf, Y.; Karamycheva, S.; Zhang, D.; Aravind, L.; Koonin, E.V. Antimicrobial Peptides, Polymorphic Toxins, and Self-Nonself Recognition Systems in Archaea: An Untapped Armory for Intermicrobial Conflicts. *mBio* **2019**, *10*, e00715-19. [CrossRef]
94. Shirazian, P.; Asad, S.; Amoozegar, M.A. The potential of halophilic and halotolerant bacteria for the production of antineoplastic enzymes: L-asparaginase and L-glutaminase. *EXCLI J.* **2016**, *15*, 268–279. [CrossRef] [PubMed]
95. Zolfaghar, M.; Amoozegar, M.A.; Khajeh, K.; Babavalian, H.; Tebyanian, H. Isolation and screening of extracellular anticancer enzymes from halophilic and halotolerant bacteria from different saline environments in Iran. *Mol. Biol. Rep.* **2019**, *46*, 3275–3286. [CrossRef] [PubMed]
96. De Castro, I.; Mendo, S.; Caetano, T. Antibiotics from Haloarchaea: What Can We Learn from Comparative Genomics? *Mar. Biotechnol.* **2020**, *22*, 308–316. [CrossRef] [PubMed]
97. Othoum, G.; Bougouffa, S.; Razali, R.; Bokhari, A.; Alamoudi, S.; Antunes, A.; Gao, X.; Hoehndorf, R.; Arold, S.T.; Gojobori, T.; et al. In silico exploration of Red Sea Bacillus genomes for natural product biosynthetic gene clusters. *BMC Genom.* **2018**, *19*, 382. [CrossRef]
98. Othoum, G.; Bougouffa, S.; Bokhari, A.; Lafi, F.F.; Gojobori, T.; Hirt, H.; Mijakovic, I.; Bajic, V.B.; Essack, M. Mining biosynthetic gene clusters in Virgibacillus genomes. *BMC Genom.* **2019**, *20*, 1–10. [CrossRef]
99. Othoum, G.; Prigent, S.; Derouiche, A.; Shi, L.; Bokhari, A.; Alamoudi, S.; Bougouffa, S.; Gao, X.; Hoehndorf, R.; Arold, S.T.; et al. Comparative genomics study reveals Red Sea Bacillus with characteristics associated with potential microbial cell factories (MCFs). *Sci. Rep.* **2019**, *9*, 19254. [CrossRef]
100. Ziko, L.; Adel, M.; Malash, M.; Siam, R. Insights into Red Sea Brine Pool Specialized Metabolism Gene Clusters Encoding Potential Metabolites for Biotechnological Applications and Extremophile Survival. *Mar. Drugs* **2019**, *17*, 273. [CrossRef]
101. Wang, H.; Fewer, D.; Holm, L.; Rouhiainen, L.; Sivonen, K. Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9259–9264. [CrossRef] [PubMed]
102. Oren, A. Diversity of halophilic microorganisms: Environments, phylogeny, physiology, and applications. *J. Ind. Microbiol. Biotechnol.* **2002**, *28*, 56–63. [CrossRef]
103. Choi, J.; Lee, D.; Jang, J.H.; Cha, S.; Seo, T. Aestuariibaculum marinum sp. nov., a marine bacterium isolated from seawater in South Korea. *J. Microbiol.* **2018**, *56*, 614–618. [CrossRef]
104. Singh, S.; Sran, K.S.; Pinnaka, A.K.; Choudhury, A.R. Purification, characterization and functional properties of exopolysaccharide from a novel halophilic Natronotalea sambharensis sp. nov. *Int. J. Biol. Macromol.* **2019**, *136*, 547–558. [CrossRef] [PubMed]
105. Fariq, A.; Yasmin, A. Production, characterization and bioactivities of biosurfactants from newly isolated strictly halophilic bacteria. *Process. Biochem.* **2020**, *98*, 1–10. [CrossRef]
106. Solden, L.; Lloyd, K.; Wrighton, K. The bright side of microbial dark matter: Lessons learned from the uncultivated majority. *Curr. Opin. Microbiol.* **2016**, *31*, 217–226. [CrossRef] [PubMed]
107. Lagier, J.-C.; Khelaifia, S.; Alou, M.T.; Ndongo, S.; Dione, N.; Hugon, P.; Caputo, A.; Cadoret, F.; Traore, S.I.; Seck, E.H.; et al. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol.* **2016**, *1*, 16203. [CrossRef]
108. Nowrotek, M.; Jałowiecki, Ł.; Harnisz, M.; Płaza, G.A. Culturomics and metagenomics: In understanding of environmental resistome. *Front. Environ. Sci. Eng.* **2019**, *13*, 40. [CrossRef]
109. Peng, P.; Lu, Y.; Bosma, T.N.P.; Nijenhuis, I.; Nijsse, B.; Shetty, S.A.; Ruecker, A.; Umanets, A.; Ramiro-Garcia, J.; Kappler, A.; et al. Metagenomic- and Cultivation-Based Exploration of Anaerobic Chloroform Biotransformation in Hypersaline Sediments as Natural Source of Chloromethanes. *Microorganisms* **2020**, *8*, 665. [CrossRef]

110. Laudadio, I.; Fulci, V.; Stronati, L.; Carissimi, C. Next-Generation Metagenomics: Methodological Challenges and Opportunities. *OMICS A J. Integr. Biol.* **2019**, *23*, 327–333. [CrossRef] [PubMed]
111. Cowan, D.A.; Ramond, J.-B.; Makhalanyane, T.P.; De Maayer, P. Metagenomics of extreme environments. *Curr. Opin. Microbiol.* **2015**, *25*, 97–102. [CrossRef] [PubMed]
112. Uritskiy, G.; DiRuggiero, J. Applying Genome-Resolved Metagenomics to Deconvolute the Halophilic Microbiome. *Genes* **2019**, *10*, 220. [CrossRef]
113. Tutuncu, H.E.; Balci, N.; Tuter, M.; Karaguler, N.G. Recombinant production and characterization of a novel esterase from a hypersaline lake, Acıgöl, by metagenomic approach. *Extremophiles* **2019**, *23*, 507–520. [CrossRef]
114. Li, Z.; Li, X.; Liu, T.; Chen, S.; Liu, H.; Wang, H.; Li, K.; Song, Y.; Luo, X.; Zhao, J.; et al. The critical roles of exposed surface residues for the thermostability and halotolerance of a novel GH11 xylanase from the metagenomic library of a saline-alkaline soil. *Int. J. Biol. Macromol.* **2019**, *133*, 316–323. [CrossRef]
115. Magnuson, E.; Mykytczuk, N.C.; Pellerin, A.; Goordial, J.; Twine, S.M.; Wing, B.; Foote, S.J.; Fulton, K.; Whyte, L.G. Thiomicrorhabdus streamers and sulfur cycling in perennial hypersaline cold springs in the Canadian high Arctic. *Environ. Microbiol.* **2020**, *23*, 3384–3400. [CrossRef]
116. Uritskiy, G.; Tisza, M.J.; Gelsinger, D.R.; Munn, A.; Taylor, J.; DiRuggiero, J. Cellular life from the three domains and viruses are transcriptionally active in a hypersaline desert community. *Environ. Microbiol.* **2020**, *23*, 3401–3417. [CrossRef] [PubMed]
117. Ramos-Barbero, M.D.; Martin-Cuadrado, A.-B.; Viver, T.; Santos, F.; Martinez-Garcia, M.; Antón, J. Recovering microbial genomes from metagenomes in hypersaline environments: The Good, the Bad and the Ugly. *Syst. Appl. Microbiol.* **2018**, *42*, 30–40. [CrossRef]
118. Vogel, C.; Marcotte, E.M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **2012**, *13*, 227–232. [CrossRef]
119. Ceylan, S.; Yilan, G.; Akbulut, B.S.; Poli, A.; Kazan, D. Interplay of adaptive capabilities of Halomonas sp. AAD12 under salt stress. *J. Biosci. Bioeng.* **2012**, *114*, 45–52. [CrossRef]
120. Bidle, K.A.; Kirkland, P.A.; Nannen, J.L.; Maupin-Furlow, J. Proteomic analysis of Haloferax volcanii reveals salinity-mediated regulation of the stress response protein PspA. *Microbiology* **2008**, *154*, 1436–1443. [CrossRef] [PubMed]
121. Sevim, V.; Lee, J.; Egan, R.; Clum, A.; Hundley, H.; Lee, J.; Everroad, R.C.; Detweiler, A.M.; Bebout, B.M.; Pett-Ridge, J.; et al. Shotgun metagenome data of a defined mock community using Oxford Nanopore, PacBio and Illumina technologies. *Sci. Data* **2019**, *6*, 285. [CrossRef]
122. Sato, M.P.; Ogura, Y.; Nakamura, K.; Nishida, R.; Gotoh, Y.; Hayashi, M.; Hisatsune, J.; Sugai, M.; Takehiko, I.; Hayashi, T. Comparison of the sequencing bias of currently available library preparation kits for Illumina sequencing of bacterial genomes and metagenomes. *DNA Res.* **2019**, *26*, 391–398. [CrossRef] [PubMed]
123. Jain, M.; Koren, S.; Miga, K.H.; Quick, J.; Rand, A.C.; Sasani, T.A.; Tyson, J.R.; Beggs, A.D.; Dilthey, A.T.; Fiddes, I.T.; et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat. Biotechnol.* **2018**, *36*, 338–345. [CrossRef]
124. Lim, A.; Naidenov, B.; Bates, H.; Willyerd, K.; Snider, T.; Couger, M.B.; Chen, C.; Ramachandran, A. Nanopore ultra-long read sequencing technology for antimicrobial resistance detection in Mannheimia haemolytica. *J. Microbiol. Methods* **2019**, *159*, 138–147. [CrossRef] [PubMed]
125. Omachi, H.; Terahara, T.; Futami, K.; Kawato, S.; Imada, C.; Kamei, K.; Waku, T.; Kondo, A.; Naganuma, T.; Agustini, T.W.; et al. Distribution of class IId bacteriocin-producing Virgibacillus salexigens in various environments. *World J. Microbiol. Biotechnol.* **2021**, *37*, 1–12. [CrossRef]
126. Tafer, H.; Poyntner, C.; Lopandic, K.; Sterflinger, K.; Piñar, G. Back to the Salt Mines: Genome and Transcriptome Comparisons of the Halophilic Fungus Aspergillus salisburgensis and Its Halotolerant Relative Aspergillus sclerotialis. *Genes* **2019**, *10*, 381. [CrossRef] [PubMed]
127. Loman, N.J.; Quick, J.; Simpson, J.T. A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat. Methods* **2015**, *12*, 733–735. [CrossRef]
128. Wick, R.R.; Judd, L.M.; Holt, K.E. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol.* **2019**, *20*, 1–10. [CrossRef]
129. Cali, D.S.; Kim, J.S.; Ghose, S.; Alkan, C.; Mutlu, O. Nanopore sequencing technology and tools for genome assembly: Computational analysis of the current state, bottlenecks and future directions. *Briefings Bioinform.* **2018**, *20*, 1542–1559. [CrossRef]
130. Blin, K.; Shaw, S.; Kloosterman, A.M.; Charlop-Powers, Z.; van Wezel, G.P.; Medema, M.H.; Weber, T. antiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* **2021**, *49*, W29–W35. [CrossRef]
131. Waghmode, S.; Suryavanshi, M.; Dama, L.; Kansara, S.; Ghattargi, V.; Das, P.; Banpurkar, A.; Satpute, S.K. Genomic Insights of Halophilic Planococcus maritimus SAMP MCC 3013 and Detail Investigation of Its Biosurfactant Production. *Front. Microbiol.* **2019**, *10*, 235. [CrossRef]
132. Waghmode, S.; Swami, S.; Sarkar, D.; Suryavanshi, M.; Roachlani, S.; Choudhari, P.; Satpute, S. Exploring the Pharmacological Potentials of Biosurfactant Derived from Planococcus maritimus SAMP MCC 3013. *Curr. Microbiol.* **2020**, *77*, 452–459. [CrossRef] [PubMed]
133. Daas, M.S.; Acedo, J.; Rosana, A.R.; Orata, F.; Reiz, B.; Zheng, J.; Nateche, F.; Case, R.J.; Kebbouche-Gana, S.; Vederas, J.C. Bacillus amyloliquefaciens ssp. plantarum F11 isolated from Algerian salty lake as a source of biosurfactants and bioactive lipopeptides. *FEMS Microbiol. Lett.* **2017**, *365*, fnx248. [CrossRef]

134. Chen, R.; Wong, A.H.L.; Kindler, G.S.; MacLeod, F.I.; Benaud, N.; Ferrari, B.C.; Burns, B.P. Discovery of an Abundance of Biosynthetic Gene Clusters in Shark Bay Microbial Mats. *Front. Microbiol.* **2020**, *11*, 1950. [CrossRef] [PubMed]
135. Donio, M.B.S.; Ronica, S.F.A.; Viji, V.T.; Velmurugan, S.; Jenifer, J.A.; Michaelbabu, M.; Citarasu, T. Isolation and characterization of halophilic Bacillus sp. BS3 able to produce pharmacologically important biosurfactants. *Asian Pac. J. Trop. Med.* **2013**, *6*, 876–883. [CrossRef]
136. Llamas, I.; Amjres, H.; Mata, J.A.; Quesada, E.; Béjar, V. The Potential Biotechnological Applications of the Exopolysaccharide Produced by the Halophilic Bacterium Halomonas almeriensis. *Molecules* **2012**, *17*, 7103–7120. [CrossRef]
137. Haque, E.; Bin Riyaz, M.A.; Shankar, S.; Hassan, S. Physiochemical and Structural Characterization of Biosurfactant Produced by Halophilic Pseudomonas aeruginosa ENO14 Isolated from Seawater. *Int. J. Pharm. Investig.* **2020**, *10*, 437–444. [CrossRef]
138. Schultz, J.; Rosado, A.S. Extreme environments: A source of biosurfactants for biotechnological applications. *Extremophiles* **2019**, *24*, 189–206. [CrossRef] [PubMed]
139. Gomes, M.B.; Gonzales-Limache, E.E.; Sousa, S.; Dellagnezze, B.; Sartoratto, A.; Silva, L.C.F.; Gieg, L.M.; Valoni, E.; Souza, R.S.; Torres, A.P.R.; et al. Exploring the potential of halophilic bacteria from oil terminal environments for biosurfactant production and hydrocarbon degradation under high-salinity conditions. *Int. Biodeterior. Biodegrad.* **2018**, *126*, 231–242. [CrossRef]
140. Pastor, J.M.; Salvador, M.; Argandoña, M.; Bernal, V.; Reina-Bueno, M.; Csonka, L.N.; Iborra, J.L.; Vargas, C.; Nieto, J.J.; Cánovas, M. Ectoines in cell stress protection: Uses and biotechnological production. *Biotechnol. Adv.* **2010**, *28*, 782–801. [CrossRef] [PubMed]
141. Bownik, A.; Stępniewska, Z. Ectoine as a promising protective agent in humans and animals. *Arch. Ind. Hyg. Toxicol.* **2016**, *67*, 260–265. [CrossRef]
142. Nosch, D.S.; Joos, R.E.; Job, M. Prospective randomized study to evaluate the efficacy and tolerability of Ectoin®containing Eye Spray (EES09) and comparison to the liposomal Eye Spray Tears Again®(TA) in the treatment of dry eye disease. *Contact Lens Anterior Eye* **2021**, *44*, 101318. [CrossRef]
143. Bilstein, A.; Heinrich, A.; Rybachuk, A.; Mösges, R. Ectoine in the Treatment of Irritations and Inflammations of the Eye Surface. *BioMed Res. Int.* **2021**, *2021*, 8885032. [CrossRef] [PubMed]
144. Tran, B.-H.; Dao, V.-A.; Bilstein, A.; Unfried, K.; Shah-Hosseini, K.; Mösges, R. Ectoine-Containing Inhalation Solution versus Saline Inhalation Solution in the Treatment of Acute Bronchitis and Acute Respiratory Infections: A Prospective, Controlled, Observational Study. *BioMed Res. Int.* **2019**, *2019*, 7945091. [CrossRef]
145. Bilstein, A.; Werkhäuser, N.; Rybachuk, A.; Mösges, R. The Effectiveness of the Bacteria Derived Extremolyte Ectoine for the Treatment of Allergic Rhinitis. *BioMed Res. Int.* **2021**, *2021*, 5562623. [CrossRef]
146. Moh, T.H.; Lau, N.S.; Furusawa, G.; Amirul, A.-A.A. Complete genome sequence of Microbulbifer sp. CCB-MM1, a halophile isolated from Matang Mangrove Forest, Malaysia. *Stand. Genom. Sci.* **2017**, *12*, 36. [CrossRef] [PubMed]
147. Van Thuoc, D.; Loan, T.T.; Trung, T.A.; Van Quyen, N.; Tung, Q.N.; Tien, P.Q.; Sudesh, K. Genome Mining Reveals the Biosynthetic Pathways of Polyhydroxyalkanoate and Ectoines of the Halophilic Strain Salinivibrio proteolyticus M318 Isolated from Fermented Shrimp Paste. *Mar. Biotechnol.* **2020**, *22*, 651–660. [CrossRef]
148. Frikha-Dammak, D.; Ayadi, H.; Hakim-Rekik, I.; Belbahri, L.; Maalej, S. Genome analysis of the salt-resistant Paludifilum halophilum DSM 102817T reveals genes involved in flux-tuning of ectoines and unexplored bioactive secondary metabolites. *World J. Microbiol. Biotechnol.* **2021**, *37*, 1–15. [CrossRef]
149. Metelev, M.; Tietz, J.; Melby, J.O.; Blair, P.M.; Zhu, L.; Livnat, I.; Severinov, K.; Mitchell, D.A. Structure, Bioactivity, and Resistance Mechanism of Streptomonomicin, an Unusual Lasso Peptide from an Understudied Halophilic Actinomycete. *Chem. Biol.* **2015**, *22*, 241–250. [CrossRef] [PubMed]
150. Petrillo, C.; Castaldi, S.; Lanzilli, M.; Selci, M.; Cordone, A.; Giovannelli, D.; Isticato, R. Genomic and Physiological Characterization of Bacilli Isolated From Salt-Pans With Plant Growth Promoting Features. *Front. Microbiol.* **2021**, *12*, 715678. [CrossRef]
151. Ghate, S.D.; Arun, A.B.; Rao, S.S.; Kumar, S.T.A.; Kandiyil, M.K.; Saptami, K.; Rekha, P.D. Genome analysis of a halophilic bacterium Halomonas malpeensis YU-PRIM-29T reveals its exopolysaccharide and pigment producing capabilities. *Sci. Rep.* **2021**, *11*, 1749. [CrossRef]
152. Matroodi, S.; Siitonen, V.; Baral, B.; Yamada, K.; Akhgari, A.; Metsä-Ketelä, M. Genotyping-Guided Discovery of Persiamycin A From Sponge-Associated Halophilic Streptomonospora sp. PA3. *Front. Microbiol.* **2020**, *11*, 1237. [CrossRef] [PubMed]
153. Yang, H.W.; Song, J.Y.; Cho, S.M.; Kwon, H.C.; Pan, C.-H.; Park, Y.-I. Genomic Survey of Salt Acclimation-Related Genes in the Halophilic Cyanobacterium Euhalothece sp. Z-M001. *Sci. Rep.* **2020**, *10*, 676. [CrossRef] [PubMed]
154. Baltz, R.H. Genome mining for drug discovery: Progress at the front end. *J. Ind. Microbiol. Biotechnol.* **2021**. [CrossRef] [PubMed]
155. Blin, K.; Kim, H.U.; Medema, M.H.; Weber, T. Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. *Brief. Bioinform.* **2017**, *20*, 1103–1113. [CrossRef] [PubMed]
156. Ziemert, N.; Podell, S.; Penn, K.; Badger, J.H.; Allen, E.; Jensen, P.R. The Natural Product Domain Seeker NaPDoS: A Phylogeny Based Bioinformatic Tool to Classify Secondary Metabolite Gene Diversity. *PLoS ONE* **2012**, *7*, e34064. [CrossRef]
157. Reddy, B.V.B.; Milshteyn, A.; Charlop-Powers, Z.; Brady, S.F. eSNaPD: A Versatile, Web-Based Bioinformatics Platform for Surveying and Mining Natural Product Biosynthetic Diversity from Metagenomes. *Chem. Biol.* **2014**, *21*, 1023–1033. [CrossRef]

158. Meleshko, D.; Mohimani, H.; Tracanna, V.; Hajirasouliha, I.; Medema, M.H.; Korobeynikov, A.; Pevzner, P.A.; Traccana, V. BiosyntheticSPAdes: Reconstructing biosynthetic gene clusters from assembly graphs. *Genome Res.* **2019**, *29*, 1352–1362. [CrossRef]
159. Kautsar, S.A.; Blin, K.; Shaw, S.; Navarro-Muñoz, J.C.; Terlouw, B.R.; Van Der Hooft, J.J.J.; Van Santen, J.A.; Tracanna, V.; Duran, H.G.S.; Andreu, V.P.; et al. MIBiG 2.0: A repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* **2019**, *48*, D454–D458. [CrossRef]
160. Palaniappan, K.; Chen, I.-M.A.; Chu, K.; Ratner, A.; Seshadri, R.; Kyrpides, N.; Ivanova, N.N.; Mouncey, N.J. IMG-ABC v.5.0: An update to the IMG/Atlas of Biosynthetic Gene Clusters Knowledgebase. *Nucleic Acids Res.* **2019**, *48*, D422–D430. [CrossRef] [PubMed]
161. Blin, K.; Shaw, S.; Kautsar, S.A.; Medema, M.H.; Weber, T. The antiSMASH database version 3: Increased taxonomic coverage and new query features for modular enzymes. *Nucleic Acids Res.* **2020**, *49*, D639–D643. [CrossRef]
162. Bairoch, A. The ENZYME database in 2000. *Nucleic Acids Res.* **2000**, *28*, 304–305. [CrossRef] [PubMed]
163. Chang, A.; Jeske, L.; Ulbrich, S.; Hofmann, J.; Koblitz, J.; Schomburg, I.; Neumann-Schaal, M.; Jahn, D.; Schomburg, D. BRENDA, the ELIXIR core data resource in 2021: New developments and updates. *Nucleic Acids Res.* **2020**, *49*, D498–D508. [CrossRef]
164. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2016**, *45*, D353–D361. [CrossRef] [PubMed]
165. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.L.; Tosatto, S.C.E.; Paladin, L.; Raj, S.; Richardson, L.J.; et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2020**, *49*, D412–D419. [CrossRef]
166. Haft, D.H.; Selengut, J.D.; Richter, R.A.; Harkins, D.; Basu, M.K.; Beck, E. TIGRFAMs and Genome Properties in 2013. *Nucleic Acids Res.* **2012**, *41*, D387–D395. [CrossRef]
167. Letunic, I.; Khedkar, S.; Bork, P. SMART: Recent updates, new developments and status in 2020. *Nucleic Acids Res.* **2020**, *49*, D458–D460. [CrossRef]
168. Delgado-García, M.; Valdivia-Urdiales, B.; Aguilar, C.N.; Contreras-Esquivel, J.C.; Rodríguez-Herrera, R. Halophilic hydrolases as a new tool for the biotechnological industries. *J. Sci. Food Agric.* **2012**, *92*, 2575–2580. [CrossRef]
169. Dalmaso, G.Z.L.; Ferreira, D.; Vermelho, A.B. Marine Extremophiles: A Source of Hydrolases for Biotechnological Applications. *Mar. Drugs* **2015**, *13*, 1925–1965. [CrossRef] [PubMed]
170. Amoozegar, M.A.; Safarpour, A.; Noghabi, K.A.; Bakhtiary, T.; Ventosa, A. Halophiles and Their Vast Potential in Biofuel Production. *Front. Microbiol.* **2019**, *10*, 1895. [CrossRef]
171. Motamedi, E.; Motahar, S.F.S.; Maleki, M.; Kavousi, K.; Ariaeenejad, S.; Moosavi-Movahedi, A.A.; Salekdeh, G.H. Upgrading the enzymatic hydrolysis of lignocellulosic biomass by immobilization of metagenome-derived novel halotolerant cellulase on the carboxymethyl cellulose-based hydrogel. *Cellulose* **2021**, *28*, 3485–3503. [CrossRef]
172. Marchler-Bauer, A.; Bo, Y.; Han, L.; He, J.; Lanczycki, C.J.; Lu, S.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; et al. CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* **2016**, *45*, D200–D203. [CrossRef]
173. Koutsandreas, T.; Ladoukakis, E.; Pilalis, E.; Zarafeta, D.; Kolisis, F.N.; Skretas, G.; Chatziioannou, A.A. ANASTASIA: An Automated Metagenomic Analysis Pipeline for Novel Enzyme Discovery Exploiting Next Generation Sequencing Data. *Front. Genet.* **2019**, *10*, 469. [CrossRef]
174. Shahraki, M.F.; Ariaeenejad, S.; Atanaki, F.F.; Zolfaghari, B.; Koshiba, T.; Kavousi, K.; Salekdeh, G.H. MCIC: Automated Identification of Cellulases From Metagenomic Data and Characterization Based on Temperature and pH Dependence. *Front. Microbiol.* **2020**, *11*, 567863. [CrossRef]
175. Sung, J.-Y.; Lee, Y.-J.; Cho, Y.-J.; Shin, M.-N.; Lee, S.-J.; Lee, H.-S.; Koh, H.; Bae, J.-W.; Shin, J.-H.; Kim, H.J.; et al. A large-scale metagenomic study for enzyme profiles using the focused identification of the NGS-based definitive enzyme research (FINDER) strategy. *Biotechnol. Bioeng.* **2021**, *118*, 4360–4374. [CrossRef] [PubMed]
176. Kang, X.; Dong, F.; Shi, C.; Liu, S.; Sun, J.; Chen, J.; Li, H.; Xu, H.; Lao, X.; Zheng, H. DRAMP 2.0, an updated data repository of antimicrobial peptides. *Sci. Data* **2019**, *6*, 1–10. [CrossRef]
177. Waghu, F.H.; Barai, R.S.; Gurung, P.; Idicula-Thomas, S. CAMPR3: A database on sequences, structures and signatures of antimicrobial peptides: Table 1. *Nucleic Acids Res.* **2015**, *44*, D1094–D1097. [CrossRef]
178. Wang, G.; Li, X.; Wang, Z. APD3: The antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res.* **2015**, *44*, D1087–D1093. [CrossRef]
179. Zhao, X.; Wu, H.; Lu, H.; Li, G.; Huang, Q. LAMP: A Database Linking Antimicrobial Peptides. *PLoS ONE* **2013**, *8*, e66557. [CrossRef]
180. Liu, S.; Fan, L.; Sun, J.; Lao, X.; Zheng, H. Computational resources and tools for antimicrobial peptides. *J. Pept. Sci.* **2016**, *23*, 4–12. [CrossRef]
181. Besse, A.; Vandervennet, M.; Goulard, C.; Peduzzi, J.; Isaac, S.; Rebuffat, S.; Carré-Mlouka, A. Halocin C8: An antimicrobial peptide distributed among four halophilic archaeal genera: Natrinema, Haloterrigena, Haloferax, and Halobacterium. *Extremophiles* **2017**, *21*, 623–638. [CrossRef]
182. Santos-Júnior, C.D.; Pan, S.; Zhao, X.-M.; Coelho, L.P. Macrel: Antimicrobial peptide screening in genomes and metagenomes. *PeerJ* **2020**, *8*, e10555. [CrossRef] [PubMed]

183. Spänig, S.; Heider, D. Encodings and models for antimicrobial peptide classification for multi-resistant pathogens. *BioData Min.* **2019**, *12*, 7. [CrossRef]
184. Gull, S.; Shamim, N.; Minhas, F. AMAP: Hierarchical multi-label prediction of biologically active and antimicrobial peptides. *Comput. Biol. Med.* **2019**, *107*, 172–181. [CrossRef]
185. Bhadra, P.; Yan, J.; Li, J.; Fong, S.; Siu, S.W.I. AmPEP: Sequence-based prediction of antimicrobial peptides using distribution patterns of amino acid properties and random forest. *Sci. Rep.* **2018**, *8*, 1697. [CrossRef] [PubMed]
186. Manikandan, P.; Moopantakath, J.; Imchen, M.; Kumavath, R.; SenthilKumar, P.K. Identification of Multi-Potent Protein Subtilisin A from halophilic bacterium Bacillus firmus VE2. *Microb. Pathog.* **2021**, *157*, 105007. [CrossRef]
187. Platas, G.; Meseguer, I.; Amils, R. Purification and biological characterization of halocin H1 from Haloferax mediterranei M2a. *Int. Microbiol.* **2002**, *5*, 15–19. [CrossRef]
188. Rodriguez-Valera, F.; Juez, G.; Kushner, D.J. Halocins: Salt-dependent bacteriocins produced by extremely halophilic rods. *Can. J. Microbiol.* **1982**, *28*, 151–154. [CrossRef]
189. Meseguer, I.; Rodriguez-Valera, F. Production and purification of halocin H4. *FEMS Microbiol. Lett.* **1985**, *28*, 177–182. [CrossRef]
190. Karthikeyan, P.; Bhat, S.G.; Chandrasekaran, M. Halocin SH10 production by an extreme haloarchaeon Natrinema sp. BTSH10 isolated from salt pans of South India. *Saudi J. Biol. Sci.* **2013**, *20*, 205–212. [CrossRef] [PubMed]
191. Majumder, A.; Biswal, M.R.; Prakash, M.K. Computational screening of antimicrobial peptides for Acinetobacter baumannii. *PLoS ONE* **2019**, *14*, e0219693. [CrossRef] [PubMed]
192. Patch, J.; Barron, A. Mini Review on Antimicrobial Peptides, Sources, Mechanism and Recent Applications. *Protein Pept. Lett.* **2019**, *27*, 4–16. [CrossRef]
193. Skinnider, M.A.; Johnston, C.W.; Gunabalasingam, M.; Merwin, N.J.; Kieliszek, A.M.; MacLellan, R.J.; Li, H.; Ranieri, M.R.M.; Webster, A.L.H.; Cao, M.P.T.; et al. Comprehensive prediction of secondary metabolite structure and biological activity from microbial genome sequences. *Nat. Commun.* **2020**, *11*, 6058. [CrossRef]
194. Khaldi, N.; Seifuddin, F.T.; Turner, G.; Haft, D.; Nierman, W.C.; Wolfe, K.; Fedorova, N.D. SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* **2010**, *47*, 736–741. [CrossRef]
195. Van Heel, A.J.; De Jong, A.; Montalbán-López, M.; Kok, J.; Kuipers, O.P. BAGEL3: Automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Res.* **2013**, *41*, W448–W453. [CrossRef]
196. Sélem-Mojica, N.; Aguilar, C.; Gutiérrez-García, K.; Martínez-Guerrero, C.E.; Barona-Gómez, F. EvoMining reveals the origin and fate of natural product biosynthetic enzymes. *Microb. Genom.* **2019**, *5*, e000260. [CrossRef]
197. Cimermancic, P.; Medema, M.H.; Claesen, J.; Kurita, K.; Brown, L.C.W.; Mavrommatis, K.; Pati, A.; Godfrey, P.A.; Koehrsen, M.; Clardy, J.; et al. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* **2014**, *158*, 412–421. [CrossRef]
198. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
199. DasSarma, S.; DasSarma, P.; Laye, V.J.; Schwieterman, E.W. Extremophilic models for astrobiology: Haloarchaeal survival strategies and pigments for remote sensing. *Extremophiles* **2019**, *24*, 31–41. [CrossRef] [PubMed]
200. Waditee-Sirisattha, R.; Kageyama, H.; Takabe, T. Halophilic microorganism resources and their applications in industrial and environmental biotechnology. *AIMS Microbiol.* **2016**, *2*, 42–54. [CrossRef]

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-023-25770-7

RESEARCH ARTICLE

A first insight into the Polish Bochnia Salt Mine metagenome

Jakub Lach[1,2] · Klaudyna Królikowska[2,3] · Monika Baranowska[2,3] · Magdalena Krupińska[4] · Dominik Strapagiel[2] · Agnieszka Matera-Witkiewicz[4] · Paweł Stączek[1]

Received: 14 August 2022 / Accepted: 2 February 2023

Abstract

The Bochnia Salt Mine is one of the oldest mines in Europe. It was established in the thirteenth century, and actively operated until 1990. The mine has been placed on the UNESCO World Heritage List. Previous research describing Polish salt mines has been focused on bioaerosol characteristics and the identification of microorganisms potentially important for human health. The use of Polish salt mines as inhalation chambers for patients of health resorts has also been investigated. Nevertheless, the biodiversity of salt mines associated with biotechnological potential has not been well characterized. The present study paper examines the biodiversity of microorganisms in the Bochnia Salt Mine based on 16S rRNA gene and shotgun sequencing. Biodiversity studies revealed a significantly higher relative abundance of *Chlamydiae* at the first level of the mine (3.5%) compared to the other levels (< 0.1%). *Patescibacteria* microorganisms constituted a high percentage (21.6%) in the sample from site RA6. Shotgun sequencing identified 16 unique metagenome-assembled genomes (MAGs). Although one was identified as *Halobacterium bonnevillei*, the others have not yet been assigned to any species; it is possible that these species may be undescribed. Preliminary analyses of the biotechnological and pharmaceutical potential of microorganisms inhabiting the mine were also performed, and the biosynthetic gene cluster (BGC) profiles and antimicrobial peptide (AMP) coding genes in individual samples were characterized. Hundreds of BGCs and dozens of AMP coding genes were identified in metagenomes. Our findings indicate that Polish salt mines are promising sites for further research aimed at identifying microorganisms that are producers of potentially important substances with biotechnological and pharmaceutical applications.

Keywords Biodiversity · Biosynthetic gene cluster · Metagenomics · Salt mine

Introduction

Halophiles are a highly diverse class of extremophilic organisms with high salinity requirements which comprises entities from all three domains of life (Ma et al. 2010). Due to their production of active biomolecules for biotechnological and pharmaceutical applications, and their unique metabolism, low nutritional requirements and adaptability to harsh conditions, halophilic microorganisms are attractive candidates for use in biomedicine and industry. Indeed, in the era of multidrug-resistant bacteria, a growing incidence of cancer, and extreme environmental pollution, research on the development of new active compounds is urgently needed. A number of valuable compounds extracted from halophiles have found applications in biotechnology, cosmetics, and medicine, such as ectoine, biosurfactants, and antimicrobial agents such as streptomonomicin (Amoozegar et al. 2017; Corral et al. 2020; Desai et al. 2020; Gómez-Villegas et al. 2020; Lach et al. 2021; Singh and Singh 2017).

Responsible Editor: Robert Duran

✉ Jakub Lach
jakub.lach@biol.uni.lodz.pl

[1] Department of Molecular Microbiology, Faculty of Biology and Environmental Protection, University of Lodz, Lodz, Poland

[2] Biobank Lab, Department of Oncobiology and Epigenetics, Faculty of Biology and Environmental Protection, University of Lodz, Lodz, Poland

[3] Department of Invertebrate Zoology and Hydrobiology, Faculty of Biology and Environmental Protection, University of Lodz, Lodz, Poland

[4] Screening of Biological Activity Assays and Collection of Biological Material Laboratory, Faculty of Pharmacy, Wroclaw Medical University Biobank, Wroclaw Medical University, Wroclaw, Poland

Hypersaline ecosystems and their habitats have been widely explored mainly because of their unique biodiversity, and studies are currently focussing on environments related to salt production, aquaculture, and food production (Delgado-García et al. 2018; Pal et al. 2020; Van Thuoc et al. 2020). Their findings have led to a better understanding of the ecology of recently discovered groups of microorganisms such as DPANN, a superphylum of Archaea living in salterns that is characterized by a smaller cell size than other archaea (Pal et al. 2020). Also, new strains of halophilic microorganisms have been isolated from food products. One example is the strain *Salinivibrio proteolyticus* M318 isolated from fermented shrimp paste, which can be used to produce polyhydroxyalkanoates and ectoines (Van Thuoc et al. 2020). Studies of saline soils, which serve as reservoirs for many unique halophile species, have been found to contain new strains of halophilic microorganisms belonging to *Halobacillus* spp., *Marinococcus* spp., and *Alkalibacillus* spp. (Delgado-García et al. 2018). Newer methods based on omics are increasingly being used to analyze the soil microbiome to obtain a better understanding of the rhizosphere microbiome; this may have important implications for agriculture (Mukhtar et al. 2019). However, there are still many hypersaline environments that have not been sufficiently studied.

One such unexplored environments is the salt mine located in Bochnia, Poland. It is one of the oldest, and longest operating, salt mines in Europe, established in the thirteenth century and actively exploited until 1990. Hence, Bochnia Salt Mine is a unique natural and historical treasure which has been placed on the UNESCO World Heritage List (UNESCO World Heritage Centr n.d.). It was first established in a fragment of the marine sediments of the Miocene salt-bearing formation. The salt deposit in the Bochnia area was formed as a result of tectonic enrichment of the initially thin salt rock layers. Its geological structure has been relatively well characterized in the literature (De Leeuw et al. 2010; Garlicki 2008; Maj 2017; Poborski 1952; Wagner et al. 2010). Currently, the historic mine consists of nine post-mining galleries reaching 350 m below the surface, and a significant part of the excavations is open to the public (Puławska et al. 2021). Since the cessation of salt mining activity, the Bochnia Salt Mine has mainly been focused on tourism, recreation, and health protection, with the number of visitors increasing from 1500 in 1990 to over 133,500 in 2011 following the adaptation of the mine for tourism activities and its recognition as a UNESCO World Heritage Site (Krupa and Dec 2014; Wiewiórka et al. 2009).

The microbiota of salt mines in Poland has been the subject of scientific interest during the last decade. However, due to the function of mines as inhalation chambers for patients of health resorts, most studies have focused on bioaerosols. Their findings indicate that the air in salt chambers is characterized by a reduced level of biological contamination, including pathogen contents (Gębarowska et al. 2018; Górny et al. 2020; Myszkowska et al. 2019). Studies on the origin and health significance of airborne dusts in the Bochnia Salt Mine have also failed to detect any potentially toxic substances in the air. In addition, the presence of salt particles and salty spray in the atmosphere of salt mine may also have a beneficial effect on health (Puławska et al. 2021). However, none of the previous studies attempted a detailed characterization of the microbial community of the environment. In the only study of Polish salt mines to use next generation sequencing (NGS) to date, Goraj et al. (2021) describe the structure of the microbial community inhabiting halite and its importance for carbon transformation. In contrast, no studies have examined the brines from seepages and outflows in salt mines, nor have they assessed the biotechnological and pharmaceutical potential of the microbiota of Polish salt mines.

It has been shown that new species of halophilic microorganisms can be found in salt mines. In 2016, a new species of halophilic microorganism *Halorhabdus rudnickae* sp. nov. was isolated from a brine sample from a borehole in the unexploited Barycz mining area near Bochnia (Albuquerque et al. 2016). Research conducted in other parts of the world has also shown that salt mines are environments where unique organisms can be found, such as *Halobellus captivus* sp. nov., *Halococcus salifodinae* sp. nov., and *Halorubrum trueperi* sp. nov. (Chen et al. 2020, 2017; Denner et al. 1994). Studies conducted on entire microbial communities, such as in the Karak Salt Mine, Pakistan, found them to be dominated by bacteria, which accounted for more than 65% of the community (mostly *Bacteroidetes* and *Proteobacteria*). Only less than 20% of the microbial population of this mine consisted of archaea (Cycil et al. 2020). In contrast, a study of the microbial biodiversity of the Kilroot Salt Mine in Carrickfergus, Northern Ireland, showed a higher abundance of haloarchaea in an NaCl stalactite than samples from the nearby brine pools and soils; in addition, the metagenomes from NaCl stalactite also contained more genes related to adaptation to extreme environments (Thompson et al. 2021).

The effect of halite entrapment on halophilic microbial community composition has also been analyzed. The results indicate that halite-trapped archaeal communities were resilient to entrapment durations of up to 21 weeks and that such a halite entrapment may be an effective survival strategy for haloarchaea (Huby et al. 2021). In addition, ancient microorganisms trapped in halite fluid inclusions were identified in the USA, China, and Australia (Chen et al. 2021; Vreeland et al. 2000). Moreover, 830-million-year-old cell forms in fluid inclusions were observed in halite from Browne Formation (Australia). If this observation is confirmed by further biological analysis, it will be possible to identify sequences of genomes of the oldest halophilic

microorganisms discovered so far (Schreder-Gomes et al. 2022). Some of the isolated microbes also encoded potentially valuable enzymes, such as a novel ω-transaminase (Kelly et al. 2018).

Due to the scarcity of biodiversity studies of brines from salt mines and the lack of assessment of the biotechnological potential of halophilic microbial communities, the aim of this study was to estimate the biodiversity of the microbiome of the Bochnia Salt Mine and to assess its biological potential, with particular emphasis on the detection of biosynthetic gene clusters (BGCs) and genes encoding antimicrobial peptides (AMPs). The historical significance of the Bochnia Salt Mine and its isolation from other saline environments characterizing this particular microbiome can be considered as a valuable supplement to the knowledge of the halophiles inhabiting salt mines in Europe and elsewhere.

Materials and methods

Sample collection and processing

Samples were collected at the Bochnia Salt Mine located in the southern part of the Carpathian Foreland, near Kraków, Poland (49°58′09″N 20°25′03″E). Rock salt mining took place in the mine from the thirteenth century until 1990 (UNESCO World Heritage Centre n.d.). The inner tectonics of the Bochnia deposits reveal a unique accumulation of steep folds of large amplitude. Apart from the halite deposits, the mine contains layers of various rocks, including anhydrite, shaley marly claystone, zuber, and anhydritic claystone (Garlicki 2008). The current structure of the mine consists of nine post-mining galleries extending down to 350 m below the surface (Puławska et al. 2021). The sampling sites are presented in Fig. 1.

Brine samples collected from seven sampling sites in the Bochnia Salt Mine and analyzed by the Hydrogeochemical Laboratory of the Department of Hydrogeology and Engineering Geology at the AGH University of Science and Technology, Krakow, showed that the brines were characterized by similar salinity and pH throughout the mine (Interewicz 2018). The brines were nearly saturated with NaCl, with the salinity ranging from 31.3 to 33.3% and the pH was close to neutral, ranging from 6.6 to 7.4. More significant differences between sampling sites were observed in ion concentration levels. Brines with the highest mineralization level (based on Ca^{2+}, Mg^{2+}, and K^{+} ion concentrations) were collected from sampling sites POP27 and RA6. Samples from the first level of the mine were generally characterized by lower mineralization than those from the third and fourth levels. Brine collected from site POP27 (third level) contained the highest level of SO_4^{2-} (111 g/l), with this value being over ten times higher than other samples (Interewicz 2018).

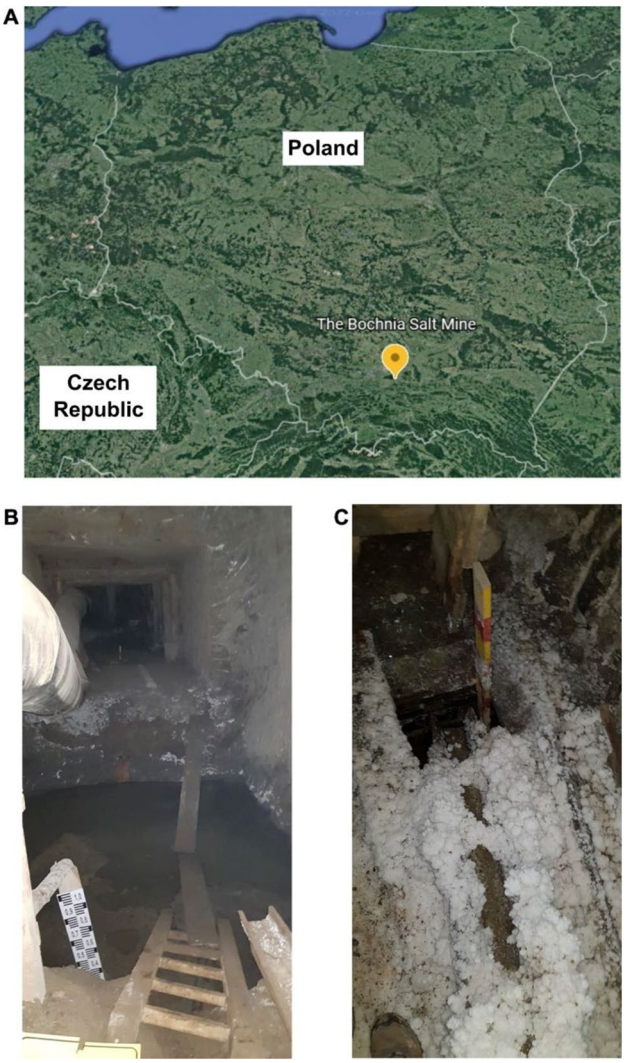


Fig. 1 **A** The location of the Bochnia Salt Mine on the map of Poland. Examples of sampling sites: **B** RDB and **C** RA10

Twenty-four water samples were collected from brine wells at three levels of the mine in 2017 and 2019 and used for 16S rRNA gene sequencing. In 2017, 17 samples were collected: ten at level 1 of the mine (Danielowiec), four at level 3 (Wernier), and three at level 4 (August); in 2019, seven samples were collected: four at level 1 and three at level 4. In addition, one sample each was taken from levels 1, 2 (Sobieski), 3, and 6 (Sienkiewicz) for shotgun sequencing. The choice of sampling sites was related to the presence of brine reservoirs and flows in the mine, and the safety of access; due to the age of the mine and the state of its maintenance and ventilation, some levels and a significant number of locations were inaccessible for sampling.

The brine samples were collected into sterile 50-ml Falcon tubes and stored at 4 °C. To prepare the material for

DNA isolation, 2 ml of saline was transferred into a new sterile Eppendorf tube and centrifuged for 10 min at 14 000 × *g*. After centrifugation, 1.8 ml brine was withdrawn and 1.8 ml of new brine was added. The centrifugation was then repeated. This procedure was repeated five times for each sample. Samples prepared in this way were then used for DNA isolation using the PowerSoil DNA Isolation Kit (Mo Bio Laboratories Inc., Carlsbad, CA, USA). The samples were eluted with nuclease-free water in a volume of 50 µl. DNA concentration was determined using the Qubit high-sensitivity (HS) assay kit (ThermoFisher, Waltham, USA).

Sequencing and data availability

16S rRNA gene amplicon sequencing libraries were prepared according to the 16S Metagenomic Sequencing Library Preparation Illumina protocol (https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf). Shotgun sequencing libraries were prepared using the Vazyme TruePrep DNA Library Prep Kit V2 for Illumina (Nanjing, China). Throughout the DNA library preparation, negative controls without DNA were used to monitor for contamination. 16S rRNA metagenomic libraries were sequenced on an Illumina MiSeq with paired end reads. The libraries prepared for shotgun sequencing were sequenced on an Illumina NextSeq 500 with paired end reads.

All sequencing data are publicly available from the National Institutes of Health under BioProject accession PRJNA848445.

Data analysis

At the first stage of the analysis, the quality of reads from both the 16S rRNA gene and shotgun sequencing protocols were checked using FastQC (Andrews 2022). In the next step, adaptors and low-quality sequences were removed from the reads with Trim Galore! v. 0.6.4 with default parameters (Krueger 2022).

Further analysis of the 16S rRNA gene samples was performed with QIIME 2 2021.8 (Bolyen et al. 2019). DADA2 was used for denoising data and ASVs (amplicon sequence variants) table generation. Both forward and reverse reads were trimmed at nucleotide 20 on the 5′ side, and at nucleotides 245 and 230, respectively, on the 3′ side. Alpha and beta diversity metrics were generated with the core-metrics-phylogenetic plugin with a sampling depth of 11,542, i.e. the number of reads that all samples were downsampled by to normalize them and reduce bias associated with differences in sequencing depths between individual samples.

Alpha diversity comparisons were performed between groups by the QIIME 2 "diversity alpha-group-significance" plugin with the Kruskal–Wallis test for both "all groups" and "pairwise" tests (Kruskal and Wallis 1952). For beta diversity comparisons, the QIIME 2 "diversity beta-group-significance" plugin with the PERMANOVA test was used (Anderson 2001). Taxonomic classification was performed using the feature-classifier classify-consensus-vsearch plugin based on pre-formatted SILVA reference sequence and taxonomy files from QIIME 2 data resources. Sequences were also functionally annotated using PICRUSt2 v. 2.4.1 (Douglas et al. 2020). Analysis of composition of microbiomes (ANCOM) was used to identify features that differed in abundance between groups (Mandal et al. 2015). Before ANOCOM, the taxa present only in one sample were removed.

For shotgun sequencing data, de novo assembly was performed using MEGAHIT v.1.2.9 with a minimum length of 1000 bp of contigs obtained (Li et al. 2015). The quality of the assemblies was checked using metaQUAST (Mikheenko et al. 2016). The contigs were binned to the metagenome-assembled genomes (MAGs) with metaBAT2 v. 2:2.15 (Kang et al. 2019). For dereplication of MAGs, dRep v3.2.2 was used (Olm et al. 2017). MAGs with a minimum length 50,000 bp, contamination lower than 25%, and completeness higher than 75% were taken for further analysis. Metagenome gene annotations were performed using DFAST version 1.2.14 (Tanizawa et al. 2018). BGC prediction was performed with antiSMASH v.6.0.1 (Blin et al. 2021) and deepBGC v. 0.1.27 (Hannigan et al. 2019). Taxonomic annotation of MAGs was performed with gtdbtk version 1.5.1 (Chaumeil et al. 2020) and at the read level using Kaiju 1.8.2 (Menzel et al. 2016). Identification of AMP sequences was performed using marcel v. 0.3.1 on default parameters (Santos-Júnior et al. 2020). AMP sequences with haemolytic properties were verified and classified into functional types by using iAMP-2L platform (Xiao et al. 2013).

Results

Sequencing statistics

NGS was used across the V3-V4 region of the 16S rRNA gene to sequence 24 environmental samples. The mean number of ASVs per sample obtained after DADA2 was 72,050 ± 31,480, ranging from 22,568 to 143,267. From a total of 1,729,209 contigs remaining after quality control and removal of chimeras, a total of 3655 unique ASVs were obtained. These ASVs represented 46 phyla, 93 classes, 209 orders, 352 families, and 635 genera within the domains of bacteria and archaea.

Sampling season comparison

To compare possible changes between sampling seasons, the results of 16S rRNA amplicon sequencing are presented for 17 samples that were selected from sites analyzed in both seasons. The Bray–Curtis distance was used for Principal Coordinate Analysis (PCoA); the results indicate that the samples were clustered based on sampling sites (Fig. 2). Only samples from the RDB site (purple) were not placed in a homogeneous cluster. However, clustering by season was not demonstrated. A comparison of Bray–Curtis distances showed no statistically significant difference between seasons (pseudo-$F = 1.020224$, $p = 0.381$). The difference between locations was globally significant (pseudo-$F = 3.67932$, $p = 0.001$). The pairwise tests identified significant differences between RDG and RA10 (pseudo-$F = 9.047075$, $p = 0.043$), RDG and RDB (pseudo-$F = 3.361130$, $p = 0.023$), and RDB and RD10 (pseudo-$F = 3.245796$, $p = 0.029$). No statistically significant differences were found between locations based on the Shannon index. The results suggest that the microbiome of each site was stable over time, allowing samples from both seasons to be used to compare the microbiome of the mine sampling sites in the further analyses.

Sampling site and mine level comparison

The microbiomes of all 24 brine samples from individual sampling sites were compared. PCoA based on Bray–Curtis distances confirmed partial clustering of samples depending on the sampling site (Figure S1). Separate clusters were formed by samples from the RDG and RD10/Wernier sites. The results also indicate that samples from other sampling sites formed one common, mixed cluster. No statistically significant differences were found between individual sites or mine levels based on the Shannon index. However, analysis based on the Bray–Curtis distance revealed a statistically significant difference at the global level (pseudo-$F = 2.707996$, $p = 0.001$), pairwise differences between Level 1 and Level 3 (pseudo-$F = 3.633673$, $p = 0.002$) and between Level 1 and Level 4 (pseudo-$F = 2.268532$, $p = 0.015$). Figure 3 shows the biplot projecting the taxon abundance onto the principal component matrix generated from the Bray–Curtis distance. The taxa most strongly influencing the location of samples on the PCoA were Halobacteria, Gammaproteobacteria, and Clostridia, at the class level, and *Halomicrobiaceae*, *Thiohalorhabdaceae*, and *Halomonadaceae* at the family level.

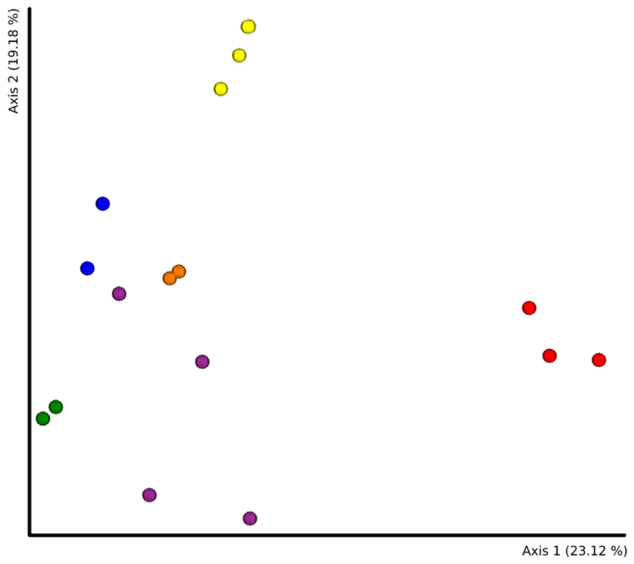


Fig. 2 PCoA calculated on Bray–Curtis distance for samples collected in a single location in both seasons. Yellow—RDG, blue—RA6, orange—RD10, purple—RDB, green—RD13, red—RA10

ANCOM analysis showed that microorganisms belonging to the Chlamydiae class of the phylum Verrucomicrobiota were significantly more abundant in the samples from level 1 of the mine than other levels, where the members of this class were almost absent. The mean content of the Chlamydiae sequence was 3.5% in samples from level 1 and 0.04% for level 4; the sequence was absent from level 3. The highest occurrence of this taxon was observed in samples from the RD13 site, where Chlamydiae accounted for 10.07% of all ASVs in the 2017 season, and 12.28% in the 2019 season. Samples from the RDB and RDG sites on level 1, like the other mine levels, had low levels of ASVs belonging to the Chlamydiae (< 1%). Within this class, most sequences belonged to the *Simkaniaceae*, but sequences belonging to families such as *Omnitrophaceae* or *Puniceicoccaceae* were also found at lower frequencies. Samples from locations RA3, RD13, and RDB had an increased proportion of *Desulfobacteraceae* representatives in the microbiota, and samples from RD13 and RDB sites additionally contained more *Thioalkalispiraceae* than other locations. The latter are microorganisms that require sulphur for growth, and their presence may be related to the structure of the mine rocks, especially the presence of anhydrite which may be a source of this element.

Samples from level 1 and level 4 of the mine contained mainly bacterial ASVs (74.8% ± 12.2% and 61.5% ± 14.8%, respectively), while samples from level 3 had a more balanced microbiome, consisting mainly of archaea (52.2% ± 7%) and bacteria (47.4% ± 7.1%). The two predominant phyla were Proteobacteria and Halobacterota. The samples from the first level of the mine contained the most Proteobacteria (31.7% ± 18.7%); the main contributors were the two samples from the RDB site, which contained 81.9% and 62.9% of Proteobacteria, respectively. In contrast, the samples from levels 3 and 4 contained 22.5% ± 10.1% and 29.7% ± 10.7% of Proteobacteria, respectively. The mean prevalence of Halobacterota was 19.5% ± 10.4% for samples

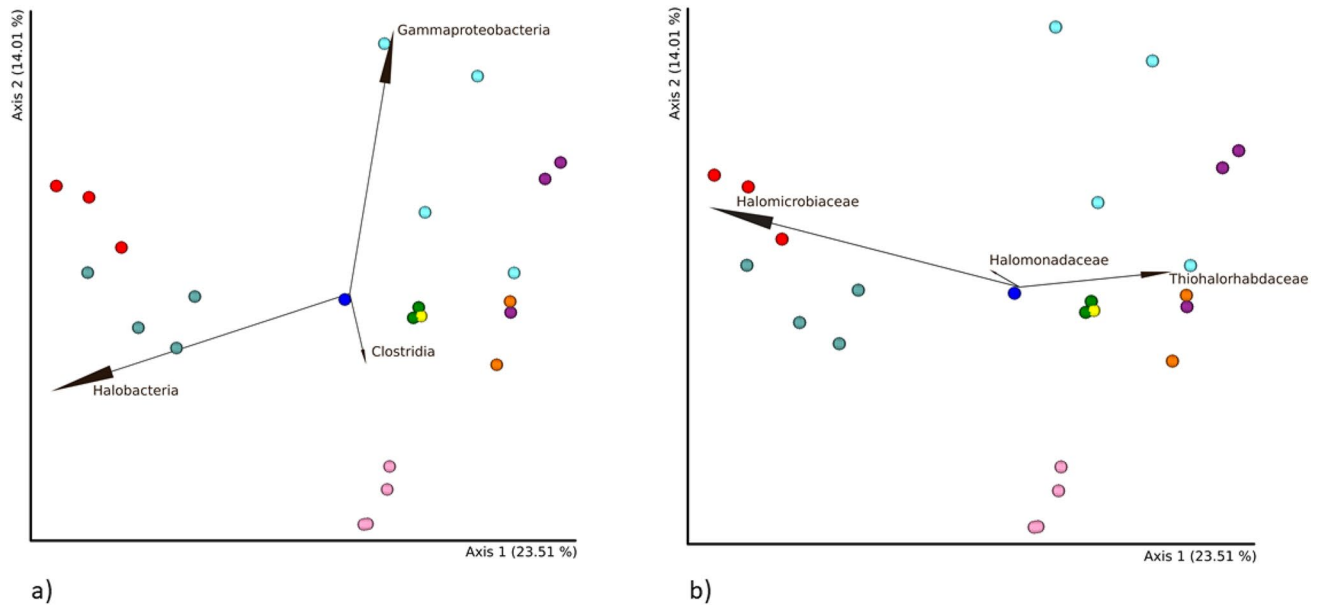


Fig. 3 Biplot projecting taxonomic abundance at (**a**) class and (**b**) family level onto a principal component matrix (PCM) generated based on the Bray–Curtis distance

from level 1, 42.3% ± 6.2% for samples from level 3, and 30.0% ± 13.8% for samples from level 4.

It is also worth noting that all samples contained bacteria belonging to the phylum Desulfobacterota, capable of metabolizing sulphur compounds. This taxon was most abundant in the first mine level, and particularly the sample from the RDB sampling site (31.5%), where the mean value for this level was 9.7%. Most of the Desulfobacterota in the RDB site were found to belong to the genera *Desulfovermiculus* and MSBL7; these were commonly accompanied by the genus *Bradymonadaceae* in other samples. The phyla Nanohaloarchaeota and Firmicutes were also present in all samples, both at an average level of 6.4%.

Samples from the RDG site were characterized by a high percentage of Halanaerobiaeota: its mean prevalence was 18.7%, compared to 3.7% for all samples. The main genera in the Halanaerobiaeota were *Halanaerobium* and *Halanaerobacter*, which were responsible for most of the ASVs assigned to this phylum. One interesting case was presented by the sample from the RA6 site, where an important element of the microbiome was the phylum Patescibacteria, comprising microorganisms with highly reduced genomes (< 1Mbp). In this sample, Patescibacteria accounted for 21.6% of the total ASVs, while the mean value for all samples was 3.5%.

In the RA6 sample, the dominant species within the Patescibacteria was *Candidatus Falkowbacteria*, which accounted for 18.6% of the microbial composition of the site. The taxonomic composition of the samples is shown in Fig. 4.

In total, 46 phyla were identified throughout all the 24 analyzed samples. Nine of these phyla were represented in all samples, making up the main part of their microbiome: Halobacterota, Proteobacteria, Bacteroidota, Desulfobacterota, Nanohaloarchaeota, Firmicutes, Patescibacteria, Halanaerobiaeota, and Acetothermia. Only two of these common phyla: Nanohaloarchaeota and Halobacterota, belonged to the Archaea. Moreover, the presence of Patescibacteria in all samples may be of interest since they are microorganisms with strongly reduced genomes, which is associated with limited possibilities of adaptation to stress conditions (Tian et al. 2020).

The taxa present in all samples comprised 14 of 96 classes, 20 of 209 orders, 22 of 352 families, and 26 of 635 genera. The genera found in all samples included *Natronomonas*, *Nanosalinaceae*, *Thiohalorhabdus*, *Salinisphaera*, *Halofilum*, *Haloplanus*, *Halobacterium*, *Halolamine*, *Desulfovermiculus*, *Bradymonadaceae*, *Alcanivorax*, *Halodesulfurarchaeum*, *Acetothermiia*, *Abscondimielaslaas*, *Pseudohlaxla*, and *Abscondimielaslaas*. The other genera were non-cultivable bacteria (designated unculturable in the SILVA database) belonging to the families *Balneolaceae*, *Halomonadaceae*, *Balneolaceae*, *Halomicrobiaceae*, and *Micavibrionaceae* or to the orders and classes (lowest level designated) Thermovenabulales, Chitinophagales, and Bacteroidia. At genus level, the most common taxa also belonged to the Bacteria (19 taxa) and only a small part were representatives of the Archaea (seven taxa); this is interesting as the very high salinity of the brine should favour the presence of extreme halophiles from the *Archaea* domain. The most representative phylum

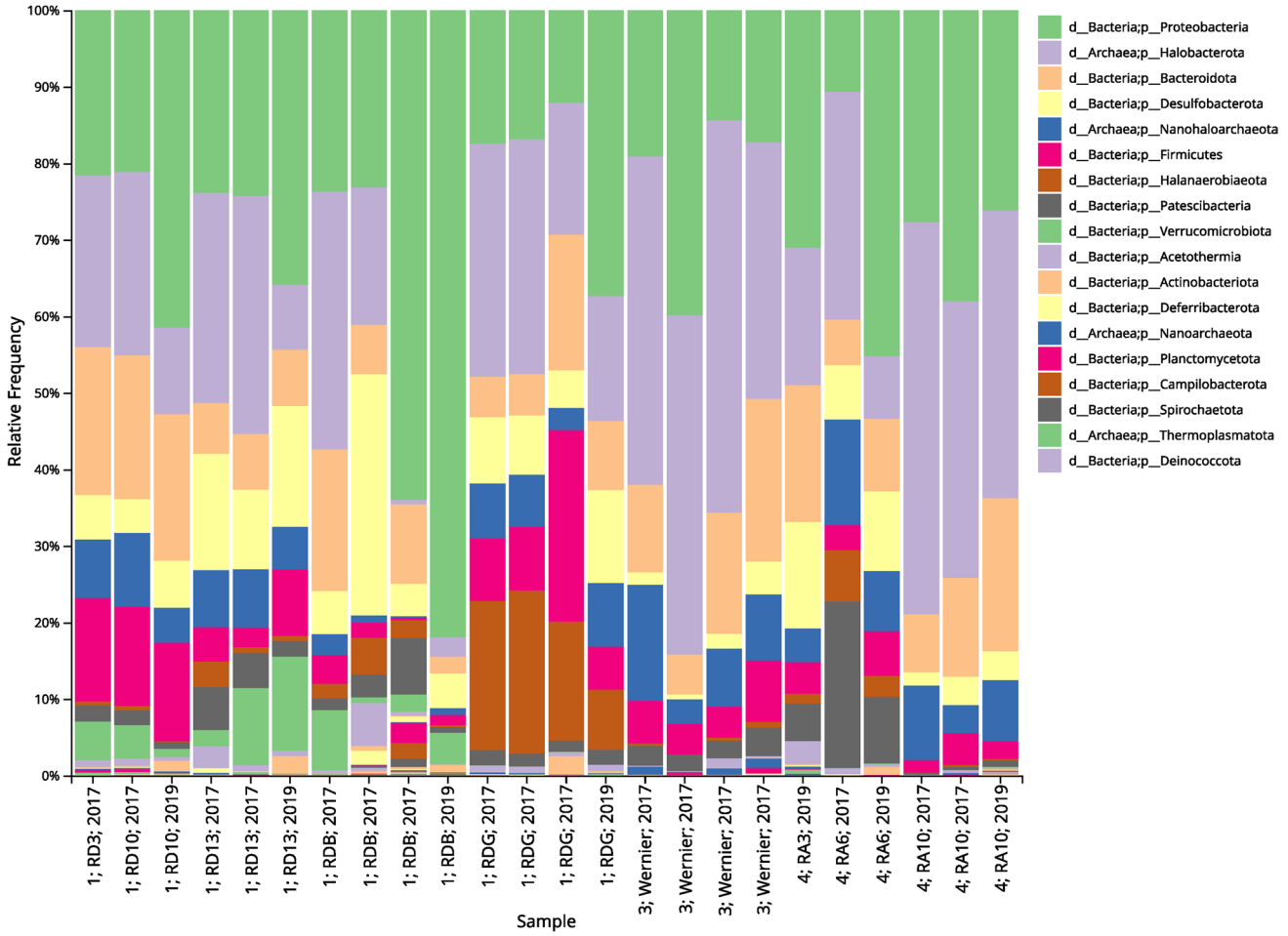


Fig. 4 Taxonomic composition of samples at phylum level. Samples are sorted by salt mine level and sampling sites

was *Proteobacteria*, which included 10 of the designated 26 genera. The presence of numerous taxa with sulphur-based metabolism may be associated with the occurrence of anhydrite in the mine (Garlicki 2008).

Metabolic pathway analysis

The ASVs were annotated with BioCyc ID metabolic pathways using the PICRUSt2 tool. PCoA based on Bray–Curtis distance was used to cluster samples in three main groups, as previously with the taxa: one group contained samples from the Wernier and RA10 sites; a second group from the RDG site, with one sample from the 2019 season being an outlier, unlike the analysis at the ASVs level; and a third group that included the remaining samples, although in this case a clear clustering pattern was observed (Figure S2).

Based on the biplot, five main pathways influencing the placement of samples on the PCoA were distinguished: PWY-5304 (sulphur oxidation superpathway (*Acidianus ambivalens*)), FAO-PWY (fatty acid β-oxidation I (generic)), PWY-7904 (tRNA- uridine 2-thiolation (thermophilic bacteria)), PWT-5971 (palmitate biosynthesis II (type II fatty acid synthase)), ANAEROFRUCAT-PWY (homolactic fermentation) (Fig. 5). The samples from the Wernier and RA10 sites demonstrated increased abundance of the PWY-5304, PWY-7904, and PWT-5971 pathways, while samples from the RDG site favoured the ANAEROFRUCAT-PWY. ANCOM analysis indicated that the samples from level 4 of the mine had higher participation of the P162-PWY, i.e. L-glutamate degradation V (via hydroxyglutarate), and PWY-5177 (glutaryl-CoA degradation) pathways than the other samples. The samples from the RA6 site demonstrated the highest share of the PWY-5177 (glutaryl-CoA degradation) pathway. In contrast, samples from RA3, RA6, RD13, RDG, RDB sites were characterized by increased levels of the P161-PWY, i.e. acetylene degradation (anaerobic) pathway. The only pathway present in samples from the RDB site was PWY-7046, i.e. the 4-coumarate degradation (anaerobic) pathways. In addition, samples from sites RA10, RD10, RD3, and Wernier had lower levels of the PWY-6749 (CMP-legionaminate biosynthesis I) pathway than the others.

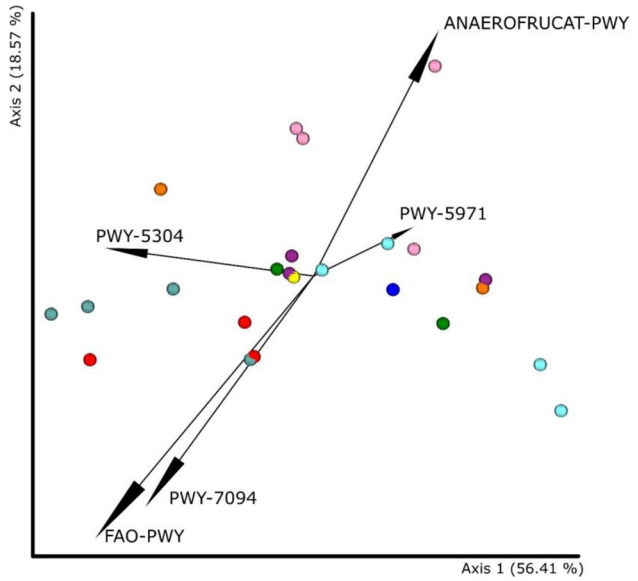


Fig. 5 Biplot projecting metabolic pathway abundance onto a principal component matrix (PCM) generated from the Bray–Curtis distance. Red—RA10 (level 4), cadet blue—Wernier (level 3), light blue—RDB (level 1), purple—RD13 (level 1), dark blue—RA3 (level 4), green—RD10 (level 1), yellow—RD3 (level 1), orange—RA6 (level 4), pink—RDG (level 1)

Core feature analysis showed that 305 of 415 (73.5%) of the identified metabolic pathways were common to all samples. At least half of the samples contained 364 of the 415 lanes (87.7%). Based on the Shannon index, a statistically significant difference was demonstrated between mine level 1 and mine level 3 (KW global test result $p = 0.007$ and $q = 0.008$ for this pair), with samples from level 3 showing significantly lower path diversity than samples from level 1. This observation was also confirmed by the Bray–Curtis distance ($p = 0.003$; $q = 0.003$). This relationship was also evident in PCoA.

Shotgun metagenomics analysis

Four samples were included in the analysis: W1 from the first level (Danielowiec), WSO4 from the second level (Sobieski), POP27 from the third level (Wernier), and W81 from the sixth level (Sienkiewicz). The total length of the metagenomes ranged from 25,265,906 to 71,844,617 bp. Each was composed of between 12,477 and 25,733 contigs longer than 1000 bp. The range of GC content was 52.0–61.7%. The number of CDS ranged from 14,004 to 49,587. The statistics concerning de novo assembly and binning are presented in Table S1, and the taxonomic assignment of reads at the phylum level is presented in Table 1.

The results indicate the dominance of Euryarchaeota with its prevalence ranging from 62.63% in W1 to 92.67% in W81. The second most represented phylum was Proteobacteria (1.79%—W81, 17.34%—POP27). Moreover, in sample W1, where the greatest diversity of taxa was observed, a significant proportion of Actinobacteria (2.89%) was found. At the genus level, all samples were characterized by a microbiome composition typical of high-salinity environments. Almost all identified genera belonged to the *Halobacteria* family, with *Halorubrum* being the most abundant genus: it

Table 1 Taxonomic composition of metagenomic samples at the phylum level. The category "Others" contains reads assigned to viruses and phyla with an abundance lower than 0.1% in each of the samples. Unassigned reads were excluded from taxa abundance calculation

Sample name	W1 (level 1) (% of reads)	WSO4 (level 2) (% of reads)	POP27 (level 3) (% of reads)	W81 (level 6) (% of reads)
Acidobacteria	0.20	1.07	1.49	1.12
Actinobacteria	2.89	0	0	0
Bacteroidetes	5.91	1.14	0.91	0.25
Balneolaeota	4.75	0.25	0.16	0
Chlamydiae	0.19	0	0	0
Chloroflexi	0.26	0	0	0
Cyanobacteria	0.35	0.12	0.17	0.27
Euryarchaeota	62.63	88.19	74.61	92.67
Firmicutes	3.22	0.68	0.77	0.68
Ignavibacteriae	0.10	0	0	0
Nitrospirae	0.12	0	0	0
Planctomycetes	0.25	0	0	0
Proteobacteria	12.40	4.83	17.34	1.79
Spirochaetes	0.15	0	0	0
Thaumarchaeota	0.10	0	0	0
Verrucomicrobia	0.12	0	0	0
Others	6.36	3.72	4.55	3.22

was the most abundant genus in POP27, WSO4, and W81 samples, with 12.1%, 14.5%, and 16.1%, respectively. In sample W1, the most common genus was *Natronomonas*, accounting for 4.7% of the microbial community, and *Halorubrum* was the second most common, with an abundance of 3.9%. *Natronomonas* was also present as an important part of the microbiomes of the other samples. Another important taxon was *Halobacterium*, which was detected in all samples, and accounted for 3.3 to 16.1%.

Sample W1 was found to demonstrate a unique microbiome composition. First, more than 25% of the microbial community in this sample comprised bacteria, which was not the case in the other samples. It was also the only sample where bacteria other than *Proteobacteria*, belonging to the Bacteroidetes, Balneolaeota, Firmicutes, and Actinobacteria, accounted for more than 1% of the microorganisms. In addition, as mentioned earlier, the dominant type of archaea in this sample was *Natronomonas*, which was not the case with other samples. These findings emphasize the slightly different nature of the W1 sample compared to the first level of the mine, thus confirming the observations made through 16S rRNA analysis. Krona charts for each sample at the genus level are shown in Fig. 6.

The dRep tool was used for dereplication of the obtained MAGs. Ultimately, 16 unique MAGs with less than 25% contamination and more than 75% completeness were selected. The bin pairs WSO4.1-W81.7, WSO4.13-POP27.10, WSO4.4-POP27.1, WSO4.22-POP27.15, and

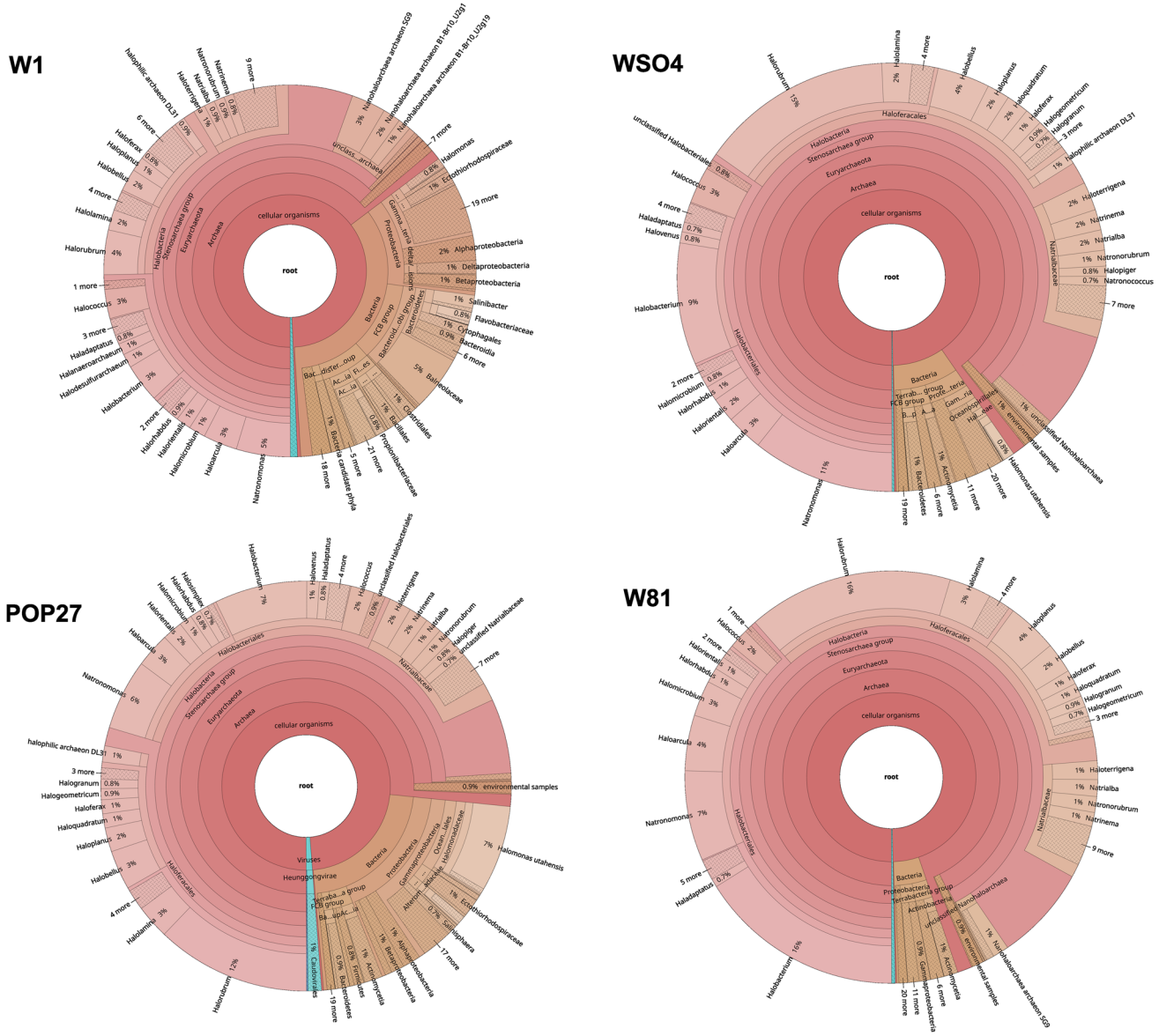


Fig. 6 Taxonomic composition at the genus level based on shotgun sequencing data

WSO4.8-POP27.16 had paired Average Nucleotide Identity (ANI) values above 99%, which indicated that they belonged to the same species and probably the same strain. Therefore, only one genome from each pair was selected for further analysis as representative of the group. Finally, the following number of MAGs was identified in each sample: three in POP27, two in W1, seven in W81, and four in WSO4. The highest number of MAGs was identified in W81; this sample demonstrated the least diversity in the microbial community, which could have resulted in a better recovery of contigs belonging to the identified MAGs. The completeness of MAGs ranged from 76.2 to 98.28%, and contamination from 1.4 to 24.61%. The resulting MAGs were between 826,353 and 5,341,881 bp in length. Detailed statistics on the quality of the bins are presented in Table S2.

Taxa were assigned to bins using the GTDB-Tk tool, as shown in Table 2. Classification to species level was only possible in bin WSO4.4; the results indicated the species *Halobacterium bonnevillei*. In the remaining cases, the classification ended at the genus or family level, which may indicate that the bins belong to unknown species. Only four of the obtained MAGs were identified as bacterial: MAG POP27.15, assigned to the genus *Halospina*, W1.7 assigned to the order T1Sed10-126, and two MAGs, POP27.16 and WSO4.8, assigned to the family *Salinisphaeraceae*. The other MAGs identified as archaeal belonged to the Halobacteria and Nanosalinia classes. Interestingly, only two MAGs belonged to the Nanosalinia, and both were identified in the W81 sample.

The BGCs were identified using two methods. The first one, antiSMASH, is based on the use of rules related to the construction of a cluster, and strongly focuses on the already existing BGCs. The second method, DeepBGC, is based on neural networks that give a better chance to identify previously undescribed BGCs, where functional domain analysis was applied. Metagenomes were also screened for the presence of AMP-encoding genes. The total number of BGCs detected in the metagenomes, the number of BGCs for each class, and the total number of detected AMPs are presented in Table 3.

It can be seen that DeepBGC enabled the detection of significantly more potential BGCs than AntiSMASH. Both methods identified the highest number of potential BGCs in the sample from the third level of the mine, and the lowest number in the sample from level 1. It was also noteworthy that both methods detected BGCs for terpenes in all samples. In sample POP27, antiSMASH identified the complete operon ectABC, related to the biosynthesis of ectoine, and the operon iucABCD, encoding aerobactin biosynthetic proteins. A BGC related to thiopeptide synthesis was also identified. The metagenome of POP27 contained numerous BGCs organized around genes coding phytoene synthase, which may suggest that these clusters were related to the biosynthesis of carotenoids. In contrast, antiSMASH identified only three clusters in sample W1; these appear to be incomplete and, therefore, cannot be properly characterized. In sample W81, two BGCs associated with the production of class II lanthipeptides were identified. One of the clusters can most

Table 2 Taxonomic classification of unique bins

Bin name	Domain	Phylum	Class	Order	Family	Genus	Species
POP27.15	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Oleiphilaceae*	*Halospina*	-
POP27.16	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Nevskiales*	*Salinisphaeraceae*	-	-
POP27.17	*Archaea*	*Halobacteriota*	*Halobacteria*	*Halobacteriales*	*Halobacteriaceae*	-	-
W1.4	*Archaea*	*Halobacteriota*	*Halobacteria*	*Halobacteriales*	*Haloarculaceae*	*Natronomonas*	-
W1.7	*Bacteria*	*T1Sed10-126*	*T1Sed10-126*	*T1Sed10-126*	-	-	-
W81.6	*Archaea*	*Nanohaloarchaeota*	*Nanosalinia*	*Nanosalinales*	*Nanosalinaceae*	*SG9*	-
W81.7	*Archaea*	*Halobacteriota*	*Halobacteria*	*Halobacteriales*	*Haloarculaceae*	*Natronomonas*	-
W81.9	*Archaea*	*Halobacteriota*	*Halobacteria*	*Halobacteriales*	*Halobacteriaceae*	*Halobacterium*	-
W81.14	*Archaea*	*Halobacteriota*	*Halobacteria*	*Halobacteriales*	*Haloferacaceae*	*Halorubrum*	-
W81.15	*Archaea*	*Nanohaloarchaeota*	*Nanosalinia*	*Halorubrum*	*Nanosalinaceae*	*M3-22*	-
W81.20	*Archaea*	*Halobacteriota*	*Halobacteria*	*Halobacteriales*	*Haloferacaceae*	*Haloplanus*	-
W81.21	*Archaea*	*Halobacteriota*	*Halobacteria*	*Halobacteriales*	*Haloarculaceae*	*Halomicrobium*	-
WSO4.1	*Archaea*	*Halobacteriota*	*Halobacteria*	*Halobacteriales*	*Haloferacaceae*	*Halolamina*	-
WSO4.4	*Archaea*	*Halobacteriota*	*Halobacteria*	*Halobacteriales*	*Halobacteriaceae*	*Halobacterium*	*Halobacterium bonnevillei PCN9*
WSO4.8	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Nevskiales*	*Salinisphaeraceae*	-	-
WSO4.13	*Archaea*	*Halobacteriota*	*Halobacteria*	*Halobacteriales*	*Haloferacaceae*	*Halorubrum*	-

Table 3 BGC and AMP statistics

Sample name	W1 (level 1)	WSO4 (level 2)	POP27 (level 3)	W81 (level 6)
AntiSMASH				
# of BGC	3	12	20	12
T3PKS	1	0	0	0
Arylpolyene	1	1	2	0
Betalactone	0	1	1	2
Ectoine	0	1	3	0
Lanthipeptide class II	0	0	0	2
RiPP-like	0	1	0	0
Siderophore	0	0	2	0
Terpene	1	8	11	8
Thiopeptide	0	0	1	0
DeepBGC				
# of BGC	136	753	856	558
Product class				
NRP	0	6	7	1
Other	0	19	17	18
Polyketide	16	86	109	52
RiPP	9	33	58	41
Saccharide	10	83	54	83
Terpene	4	24	20	15
Unclassified	99	516	604	350
Product activity				
Antibacterial	115	621	696	459
Antifungal	0	3	8	1
Cytotoxic	2	7	11	9
Inhibitor	1	10	9	5
Unclassified	18	115	140	85
AMPs				
# of AMPs	15	30	25	7
Haemolytic	12	20	14	6
Non-haemolytic	3	10	11	1

likely produce three lanthipeptides: MSVAIDNKAVIGG – RKKQFDAEFEDhbNDhbDDKDhaDDhbLGPAKLCIG-DhaFDhbCVLNDhaRVVVP, VCERDASRLVNVAVRSR-SQHSRFAVFNVSLSSSWLSIDTETNELLTSQTA – LDhbRQCRDhbNYCCNQDhbPIFDhbWLDhbV, LSIDTETNELLTSQTG – LDhbRQCLRAFRVRQE-LYRKPVCEALKHAVLADhbQQRFEINDVLRVD-haVAWKYALMDhbKWLYALRERF and the other is related to the production of four lanthipeptides: MSVAMS-DDAGATVNKQAYAEEFNQSAPSTDHDKGEDTVGA – NCYLNDhaIVCDhbFDhbDhbGGQ, VPIIRTFIFLK-TAATA – FERWGVRLCRARDhbPAEFLKLIIRLHLKNI-AFKHDhaDhaFR, MAMAAPRRSGTSNESRSRSPAA – ADhbNDhaADhbGADhbMPDhbNCADhbVLAIDha, VSHHRHDEVPSA – LVEIAQQDhbAQQERERHVRE-CAAPGADVV. In the case of the WSO4 sample, as in POP27, the complete ectABC operon was identified. BGCs associated with the synthesis of a RiPP-like protein, most likely similar to Linocin M18, were also identified in this sample.

Discussion

The salt mine in Bochnia is an excellent example of a hypersaline site, the analysis of which may contribute to a better understanding of the ecology and genomics of halophiles. Due to the great historical and natural value of the site, our exploration of the microbiota of the Bochnia Salt Mine is particularly important (UNESCO World Heritage Centr, n.d.). Our findings present a novel insight into the structure of the microbiota of the Bochnia Salt Mine using metagenomics, which provides a glimpse into the diversity, distribution, and functionality of the microbiome as well as its structure. As such, the present study is the first such investigation of the hitherto undescribed metagenomic community

profiles of the halophile community from the Bochnia Salt Mine.

The observed differences in microbial community composition were probably not associated with the basic physicochemical properties of analyzed brines. This is due to the fact that brines in the Bochnia Salt Mine are characterized by similar salinity, pH, and levels of major ions. Each of the brines was almost saturated with NaCl and had a pH close to neutral. Therefore, it can be assumed that any recorded differences in the microbial composition were related to the location within the mine, as well as other factors such as surface water supply, or differences in brine composition at the level of trace elements.

The 16S rRNA analysis revealed significant differences between microbiome composition at different mine levels. A statistically significant difference in biodiversity was found based on Bray–Curtis distance between the first mine level and the third and fourth levels. In addition, a difference in the diversity of metabolic pathways assigned to ASVs was also observed between the first and third mine levels. These differences could be related to the mineralization of brine. The Ca^{2+}, Mg^{2+}, and K^{+} ion concentration was lower in brines from the first level than those from the third and fourth levels, which could have affected the microbial community structure. This observation was confirmed by PCoA analysis, which found the samples from the first level to be widely scattered. Samples from the other mine levels were more clustered.

ANCOM analysis showed that microorganisms belonging to the Chlamydiae class of the Verrucomicrobiota were significantly more frequent in samples from level 1 of the mine than in those of other levels. The mean Chlamydiae sequence content was 3.5% in samples from level 1, only 0.04% in level 4, and absent from level 3. The highest occurrence of the this taxon was observed in samples from the RD13 site, where Chlamydiae accounted for 10.07% of all ASVs in the 2017 season, and 12.28% in the 2019 season. Within the Chlamydiae class, sequences belonging to the *Simkaniaceae* family predominated, but sequences belonging to other families, such as Omnitrophaceae and Puniceicoccaceae, were also found at low frequency. This is an interesting observation because the detected microorganisms are not characteristic for a hypersaline environment. Bacteria belonging to the class Chlamydiae are obligate intracellular parasites. Therefore, it can be expected that the RD13 brine contained a large number of eukaryotic cells, probably in the form of Protists, which can be hosts for Chlamydiae bacteria. It is possible that the brines of level 1 were characterized by a lower salinity, which together with the Chlamydiae, was also manifested by a reduced presence of halophilic Archaea; however, this supposition was not confirmed by the chemical analysis of brines. Another explanation for the higher abundance of *Chlamydiae* on the first level of mine could be that they come from surface waters. Pawlikowska-Warych and Deptuła (2019) report the presence of bacteria from the *Simkaniaceae* family in surface waters in Poland; this would suggest that microorganisms from the surface water supplying brine outflows may reach the upper levels of the mine. Although the observation of *Chlamydiae* in such numbers in a hypersaline environment was unusual, similar observations had been made previously (Collingro et al. 2020).

Another interesting example was the case of the sample from site RA6, in which the phylum Patescibacteria was an important component of the microbiome: 21.6% of all ASVs in the sample from site RA6 belonged to this taxon, compared to 3.5%, which was average for all samples. This taxon includes microorganisms with highly reduced genomes (< 1Mbp). In the sample from site RA6, the dominant species within the Patescibacteria phylum was *Candidatus Falkowbacteria*, which constituted 18.6% of the microbiota composition of this site (Brown et al. 2015). These organisms were common in aquifers where conditions fluctuate between aerobic and anoxic (Anantharaman et al. 2016). They can also be related to the dissimilatory sulphur cycle in the environment (Anantharaman et al. 2018).

Some samples were found to harbour a large number of taxa not typically identified by 16S rRNA sequencing as extremophiles, such as Proteobacteria, Verrucomicrobiota, or Patescibacteria. This may suggest contamination of the Bochnia Salt Mine by a large number of halotolerants commonly identified in other environments. This is possible considering the flow of tourists, and surface water, into the mine; however, due to the very high salinity of the brines, exceeding 30% NaCl, it seems that survival in such conditions over a prolonged period would require the halotolerants to develop mechanisms similar to those of strict halophiles.

Furthermore, metagenomic data from shotgun sequencing found the mine microbiome to have slightly different characteristics. In contrast to the 16S rRNA gene sequencing, shotgun metagenomics found *Archaea* to be dominant in all samples. It was also noticed that fewer taxa were found at particular taxonomic levels compared to the 16S samples. Using shotgun sequencing, 16 unique MAGs were obtained with contamination less than 25% and completeness greater than 75%. Twelve out of 16 MAGs were assigned to *Archaea*, indicating that they were one of the most common microorganisms in the studied habitats.

It was only possible to assign a species to a single MAG: WSO4.4, which was assigned to the species *Halobacterium bonnevillei* (Myers and King 2020). Genomic sequence similarity analysis of the other MAGs indicated that they belong to currently unknown microorganism species. Interestingly, two of the received MAGs belong to the *Nanosalinaceae* family which is a part of the phylum Nanohaloarchaeota, which is part of DPANN superphylum. This taxon includes non-cultured microorganisms living in symbiosis with Halobacteria,

characterized by small genomes and limited metabolic capabilities (Castelle et al. 2018; Xie et al. 2022). Due to the still incomplete characterization of microorganisms belonging to the Nanohaloarchaeota, the obtained MAGs may bring significant value to further research.

The results of the shotgun sequencing suggest that the high abundance of *Proteobacteria* indicated by the 16S rRNA results may not reflect the true presence of these bacteria in the environment: a methodological error may in fact favour the detection of this type of microorganism by 16S rRNA sequencing. Indeed, other researchers have also obtained high abundances of *Proteobacteria* in analyzed brines, example from the Karak Salt Mine or the graduation towers in the Ciechocinek city (Cycil et al. 2020; Kalwasińska et al. 2018).

The profiles of BGCs and AMP-encoding genes were also characterized in individual samples using AntiSMASH and DeepBGC. Only BGCs for terpenes were detected in all samples. The highest number of BGCs detected using both methods was noted in a sample from mine level 3; this was also the only sample in which BGCs for siderophore and thiopeptide production were detected. Siderophores are considered valuable products in agriculture, serving as plant growth–promoting factors or natural plant protection agents. They also have applications in medicine, as siderophore-antibiotic conjugates are effective against multi-drug-resistant pathogens, and can be used as imaging agents, and biosensors for metal ions and pathogens (Fan and Fang 2021; Ghosh et al. 2020). Thiopeptides are also useful in medicine where they can be applied as antibiotics (Wang et al. 2020). AMPs with haemolytic or antibacterial properties were detected in all samples, except for the one collected at level 6. Non-haemolytic, antibacterial properties are significant features supporting the use of AMP in clinical applications; however, even AMPs with haemolytic properties can find practical applications, such as nisin, which has been used in food preservation (Santos et al. 2018).

In conclusion, the presented study provides the first insight into the metagenome of the Bochnia Salt Mine. Our findings demonstrate that the microbiome of the mine is highly diverse and contains many potentially unknown species of microorganisms. It was also shown that the microorganisms inhabiting the mine contain numerous BGCs and AMP-encoding genes, which could be associated with valuable bioproducts.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11356-023-25770-7.

Acknowledgements The authors are grateful to the management of the Bochnia Salt Mine for granting permission for sampling; Mr. Tomasz Migdas, the Technical Director of the Bochnia Salt Mine, for his invaluable help in organizing sampling and for inspiring conversation; and to the mine geologists: Michał Flasza and Aleksandra Puławska for guiding the sampling tours.

Author contribution Conceptualization: Jakub Lach, Paweł Stączek, Agnieszka Matera-Witkiewicz, Dominik Strapagiel. Resources: Paweł Stączek, Dominik Strapagiel, Klaudyna Królikowska. Methodology: Jakub Lach, Klaudyna Królikowska, Magdalena Krupińska, Monika Baranowska. Formal analysis and investigation: Jakub Lach, Klaudyna Królikowska, Magdalena Krupińska, Monika Baranowska. Data curation: Jakub Lach, Paweł Stączek. Writing—original draft preparation: Jakub Lach. Writing—review and editing: Paweł Stączek, Dominik Strapagiel, Agnieszka Matera-Witkiewicz. Supervision: Paweł Stączek, Agnieszka Matera-Witkiewicz, Dominik Strapagiel.

Funding This research was financed from the following funds: "InterDOC-STARt" project (POWR.03.02.00–00-I033/16–00), the MINI InterDOC-STARt grant, Wroclaw Medical University SUB.D250.22.018 project, and WBIOŚ B1811000000038.01 project.

Data availability All sequencing data are publicly available from the National Institutes of Health under BioProject accession PRJNA848445. Other types of datasets used during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Competing of interest The authors declare no competing interests.

References

Albuquerque L, Kowalewicz-Kulbat M, Drzewiecka D, Staczek P, d'Auria G, Rosselló-Móra R, da Costa MS (2016) Halorhabdus rudnickae sp. nov., a halophilic archaeon isolated from a salt mine borehole in Poland. Syst Appl Microbiol 39:100–105. https://doi.org/10.1016/j.syapm.2015.12.004

Amoozegar MA, Siroosi M, Atashgahi S, Smidt H, Ventosa A (2017) Systematics of haloarchaea and biotechnological potential of their hydrolytic enzymes. Microbiol (united Kingdom) 163:623–645. https://doi.org/10.1099/mic.0.000463

Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, Thomas BC, Singh A, Wilkins MJ, Karaoz U, Brodie EL, Williams KH, Hubbard SS, Banfield JF (2016) Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Commun 7:1–11. https://doi.org/10.1038/ncomms13219

Anantharaman K, Hausmann B, Jungbluth SP, Kantor RS, Lavy A, Warren LA, Rappé MS, Pester M, Loy A, Thomas BC, Banfield

JF (2018) Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle. ISME J 12:1715–1728. https://doi.org/10.1038/s41396-018-0078-0

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26:32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Andrews S (2022) FastQC: a quality control tool for high throughput sequence data [WWW Document]. URL https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, Van Wezel GP, Medema MH, Weber T (2021) AntiSMASH 6.0: Improving cluster detection and comparison capabilities. Nucleic Acids Res 49:W29–W35. https://doi.org/10.1093/nar/gkab335

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852–857. https://doi.org/10.1038/s41587-019-0209-9

Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, Wilkins MJ, Wrighton KC, Williams KH, Banfield JF (2015) Unusual biology across a group comprising more than 15% of domain Bacteria. Nature 523:208–211. https://doi.org/10.1038/nature14486

Castelle CJ, Brown CT, Anantharaman K, Probst AJ, Huang RH, Banfield JF (2018) Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. Nat Rev Microbiol 16:629–645. https://doi.org/10.1038/s41579-018-0076-2

Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH (2020) GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848

Chen S, Sun S, Xu Y, Chen F, Liu J (2020) Halobellus captivus sp. nov., an extremely halophilic archaeon isolated from a subterranean salt mine. Antonie van Leeuwenhoek. Int J Gen Mol Microbiol 113:221–231. https://doi.org/10.1007/s10482-019-01332-1

Chen S, Xu Y, Ke LX (2017) Halorubrum trueperi sp. nov., a halophilic archaeon isolated from a salt mine. Int J Syst Evol Microbiol 67:1564–1570. https://doi.org/10.1099/ijsem.0.001762

Chen Y, Wang J, Zhou Z, Chen X, Zhao Z, Shi P, Han Y (2021) Ancient microorganisms and carotenoids preserved in fluid inclusions in halite from Chaka Salt Lake, Western China: evidence from micro-observation and in situ Raman spectroscopy. Acta Geol Sin 95:1008–1015. https://doi.org/10.1111/1755-6724.14618

Collingro A, Köstlbacher S, Horn M (2020) Chlamydiae in the environment. Trends Microbiol 28:877–888. https://doi.org/10.1016/j.tim.2020.05.020

Corral P, Amoozegar MA, Ventosa A (2020) Halophiles and their biomolecules: recent advances and future applications in biomedicine. Mar Drugs 18. https://doi.org/10.3390/md18010033

Cycil LM, DasSarma S, Pecher W, McDonald R, AbdulSalam M, Hasan F (2020) Metagenomic insights into the diversity of halophilic microorganisms indigenous to the Karak Salt Mine. Pakistan Front Microbiol 11:1–14. https://doi.org/10.3389/fmicb.2020.01567

De Leeuw A, Bukowski K, Krijgsman W, Kuiper KF (2010) Age of the Badenian salinity crisis; impact of Miocene climate variability on the circum-mediterranean region. Geology 38:715–718. https://doi.org/10.1130/G30982.1

Delgado-García M, Contreras-Ramos SM, Rodríguez JA, Mateos-Díaz JC, Aguilar CN, Camacho-Ruíz RM (2018) Isolation of halophilic bacteria associated with saline and alkaline-sodic soils by culture dependent approach. Heliyon 4. https://doi.org/10.1016/j.heliyon.2018.e00954

Denner EBM, McGenity TJ, Busse HJ, Grant WD, Wanner G, Stan-Lotter H (1994) Halococcus salifodinae sp. nov., an archaeal isolate from an Austrian salt mine. Int J Syst Bacteriol 44:774–780. https://doi.org/10.1099/00207713-44-4-774

Desai C, Patel P, Markande AR, Kamala K, Sivaperumal P (2020) Exploration of haloarchaea for their potential applications in food industry. Int J Environ Sci Technol 17:4455–4464. https://doi.org/10.1007/s13762-020-02773-2

Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI (2020) PICRUSt2 for prediction of metagenome functions. Nat Biotechnol 38:685–688. https://doi.org/10.1038/s41587-020-0548-6

Fan D, Fang Q (2021) Siderophores for medical applications: imaging, sensors, and therapeutics. Int J Pharm 597:120306. https://doi.org/10.1016/j.ijpharm.2021.120306

Garlicki A (2008) Salt mines at Bochnia and Wieliczka. Prz Geol 56:663–669

Gębarowska E, Pusz W, Kucińska J, Kita W (2018) Comparative analysis of airborne bacteria and fungi in two salt mines in Poland. Aerobiologia (bologna) 34:127–138. https://doi.org/10.1007/s10453-017-9502-6

Ghosh SK, Bera T, Chakrabarty AM (2020) Microbial siderophore – a boon to agricultural sciences. Biol Control 144:104214. https://doi.org/10.1016/j.biocontrol.2020.104214

Gómez-Villegas P, Vigara J, Vila M, Varela J, Barreira L, Léon R (2020) Antioxidant, antimicrobial, and bioactive potential of two new haloarchaeal strains isolated from odiel salterns (Southwest Spain). Biology (basel) 9:1–20. https://doi.org/10.3390/biology9090298

Goraj W, Safranek-Nakonieczna A, Grządziel J, Polakowski C, Słowakiewicz M, Zheng Y, Gałązka A, Stępniewska Z, Pytlak A (2021) Microbial involvement in carbon transformation via CH4 and CO2 in saline sedimentary pool. Biology (Basel) 10.8:792. https://doi.org/10.3390/biology10080792

Górny RL, Frączek K, Ropek DR (2020) Size distribution of microbial aerosols in overground and subterranean treatment chambers at health resorts. J Environ Heal Sci Eng 18:1437–1450. https://doi.org/10.1007/s40201-020-00559-9

Hannigan GD, Prihoda D, Palicka A, Soukup J, Klempir O, Rampula L, Durcak J, Wurst M, Kotowski J, Chang D, Wang R, Piizzi G, Temesi G, Hazuda DJ, Woelk CH, Bitton DA (2019) A deep learning genome-mining strategy for biosynthetic gene cluster prediction. Nucleic Acids Res 47:e110. https://doi.org/10.1093/nar/gkz654

Huby TJC, Clark DR, McKew BA, McGenity TJ (2021) Extremely halophilic archaeal communities are resilient to short-term entombment in halite. Environ Microbiol 23:3370–3383. https://doi.org/10.1111/1462-2920.14913

Interewicz E (2018) Izolacja drobnoustrojów halofilnych z naturalnych solanek i próbek powietrza z terenu Polski. (Isolation of halophilic microorganisms from natural brines and air samples from Poland). MSc thesis prepared in Polish at the Faculty of Biology and Environmental Protection, University of Lodz, Lodz, Poland

Kalwasińska A, Deja-Sikora E, Burkowska-But A, Szabó A, Felföldi T, Kosobucki P, Krawiec A, Walczak M (2018) Changes in bacterial and archaeal communities during the concentration of brine at the graduation towers in Ciechocinek spa (Poland). Extremophiles 22:233–246. https://doi.org/10.1007/s00792-017-0992-5

Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z (2019) MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 2019:1–13. https://doi.org/10.7717/peerj.7359

Kelly SA, Skvortsov T, Magill D, Quinn DJ, McGrath JW, Allen CCR, Moody TS, Gilmore BF (2018) Characterization of a novel ω-transaminase from a Triassic salt mine metagenome. Biochem Biophys Res Commun 503:2936–2942. https://doi.org/10.1016/j.bbrc.2018.08.073

Krueger F (2012) Trim Galore: a wrapper tool around Cutadapt and FastQC. Trim Galore! [WWW Document]. URL https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ Accessed 14 Aug 2022

Krupa J, Dec B (2014) Tourist and spa attractiveness of the salt mines in Poland. Tur i Rekreac 11:121–131

Kruskal W, Wallis A (1952) Use of ranks in one-criterion variance analysis. J Am Stat Assoc 47:583–621

Lach J, Jęcz P, Strapagiel D, Matera-Witkiewicz A, Stączek P (2021) The methods of digging for "gold" within the salt: characterization of halophilic prokaryotes and identification of their valuable biological products using sequencing and genome mining tools. Genes (Basel) 12. https://doi.org/10.3390/genes12111756

Li D, Liu CM, Luo R, Sadakane K, Lam TW (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Ma Y, Galinski EA, Grant WD, Oren A, Ventosa A (2010) Halophiles 2010: Life in saline environments. Appl Environ Microbiol 76:6971–6981. https://doi.org/10.1128/AEM.01868-10

Maj A (2017) Rock-mass movement monitoring system in historical salt mines, using the example of the Bochnia Salt Mine. Procedia Eng 191:496–503. https://doi.org/10.1016/j.proeng.2017.05.209

Mandal S, Van Treuren W, White RA, Eggesbø M, Knight R, Peddada SD (2015) Analysis of composition of microbiomes: a novel method for studying microbial composition. Microb Ecol Heal Dis 26. https://doi.org/10.3402/mehd.v26.27663

Menzel P, Ng KL, Krogh A, (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun 7. https://doi.org/10.1038/ncomms11257

Mikheenko A, Saveliev V, Gurevich A (2016) MetaQUAST: evaluation of metagenome assemblies. Bioinformatics 32:1088–1090. https://doi.org/10.1093/bioinformatics/btv697

Mukhtar S, Mehnaz S, Malik KA (2019) Microbial diversity in the rhizosphere of plants growing under extreme environments and its impact on crop improvement. Environ Sustain 2:329–338. https://doi.org/10.1007/s42398-019-00061-5

Myers MR, King GM (2020) Halobacterium bonnevillei sp. Nov., halobaculum saliterrae sp. nov. and halovenus carboxidivorans sp. nov., three novel carbon monoxide-oxidizing halobacteria from saline crusts and soils. Int J Syst Evol Microbiol 70:4261–4268. https://doi.org/10.1099/ijsem.0.004282

Myszkowska D, Kostrzon M, Dyga W, Kędzierska J, Namysł M, Stanisz A, Zagórska M, Ziemianin M, Obtułowicz K, Czarnobilska E (2019) Bioaerosol of salt chambers in the 'Wieliczka' Salt Mine. Poland Aerobiologia (bologna) 35:297–311. https://doi.org/10.1007/s10453-019-09561-7

Olm MR, Brown CT, Brooks B, Banfield JF (2017) DRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J 11:2864–2868. https://doi.org/10.1038/ismej.2017.126

Pal S, Biswas R, Misra A, Sar A, Banerjee S, Mukherjee P, Dam B (2020) Poorly known microbial taxa dominate the microbiome of hypersaline Sambhar Lake salterns in India. Extremophiles 24:875–885. https://doi.org/10.1007/s00792-020-01201-0

Pawlikowska-Warych M, Deptuła W (2019) The first detection of the sequence of bacteria from the Simkaniaceae family in surface waters in Poland. Pol J Vet Sci 22:61–65. https://doi.org/10.24425/pjvs.2018.125608

Poborski J (1952) The Bochnia salt deposit on the geological background of region. Biul Państwowego Inst Geol 78:3–160

Puławska A, Manecki M, Flasza M, Styszko K (2021) Origin, distribution, and perspective health benefits of particulate matter in the air of underground salt mine: a case study from Bochnia, Poland. Environ Geochem Health 43:3533–3556. https://doi.org/10.1007/s10653-021-00832-2

Santos-Júnior CD, Pan S, Zhao XM, Coelho LP (2020) Macrel: antimicrobial peptide screening in genomes and metagenomes. PeerJ 8:1–20. https://doi.org/10.7717/peerj.10555

Santos JCP, Sousa RCS, Otoni CG, Moraes ARF, Souza VGL, Medeiros EAA, Espitia PJP, Pires ACS, Coimbra JSR, Soares NFF (2018) Nisin and other antimicrobial peptides: production, mechanisms of action, and application in active food packaging. Innov Food Sci Emerg Technol 48:179–194. https://doi.org/10.1016/j.ifset.2018.06.008

Schreder-Gomes SI, Benison KC, Bernau JA (2022) 830-million-year-old microorganisms in primary fluid inclusions in halite. Geology XX:1–5. https://doi.org/10.1130/g49957.1

Singh A, Singh AK (2017) Haloarchaea: worth exploring for their biotechnological potential. Biotechnol Lett 39:1793–1800. https://doi.org/10.1007/s10529-017-2434-y

Tanizawa Y, Fujisawa T, Nakamura Y (2018) DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. Bioinformatics 34:1037–1039. https://doi.org/10.1093/bioinformatics/btx713

Thompson TP, Kelly SA, Skvortsov T, Plunkett G, Ruffell A, Hallsworth JE, Hopps J, Gilmore BF (2021) Microbiology of a NaCl stalactite 'salticle' in Triassic halite. Environ Microbiol 23:3881–3895. https://doi.org/10.1111/1462-2920.15524

Tian R, Ning D, He Z, Zhang P, Spencer SJ, Gao S, Shi W, Wu L, Zhang Y, Yang Y, Adams BG, Rocha AM, Detienne BL, Lowe KA, Joyner DC, Klingeman DM, Arkin AP, Fields MW, Hazen TC, Stahl DA, Alm EJ, Zhou J (2020) Small and mighty: adaptation of superphylum Patescibacteria to groundwater environment drives their genome simplicity. Microbiome 8:1–15. https://doi.org/10.1186/s40168-020-00825-w

Van Thuoc D, Loan TT, Trung TA, Van Quyen N, Tung QN, Tien PQ, Sudesh K (2020) Genome mining reveals the biosynthetic pathways of polyhydroxyalkanoate and ectoines of the halophilic strain Salinivibrio proteolyticus M318 isolated from fermented shrimp paste. Mar Biotechnol 22:651–660. https://doi.org/10.1007/s10126-020-09986-z

Vreeland RH, Rosenzweig WD, Powers DW (2000) Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. Nature 407:897–900

Wagner M, Bukowski K, Stochel B (2010) Petrological character of lignite (brown coal) from Badenian salts in the Bochnia Mine (Southern Poland). Geol Q 54:439–448

Wang W, Park KH, Lee J, Oh E, Park C, Kang E, Lee J, Kang H (2020) A new thiopeptide antibiotic, micrococcin P3, from a marine-derived strain of the bacterium bacillus stratosphericus. Molecules 25 https://doi.org/10.3390/molecules25194383

Wieliczka and Bochnia Royal Salt Mines. UNESCO World Heritage Convention [WWW Document]. https://whc.unesco.org/en/list/32/. Accessed 14 Aug 2022

Wiewiórka J, Dudek K, Charkot J, Gonera M (2009) Natural and historic heritage of the Bochnia Salt Mine (South Poland). Stud. Univ Babes-Bolyai, Geol 54:43–47. https://doi.org/10.5038/1937-8602.54.1.9

Xiao X, Wang P, Lin WZ, Jia JH, Chou KC (2013) IAMP-2L: A two-level multi-label classifier for identifying antimicrobial peptides and their functional types. Anal Biochem 436:168–177. https://doi.org/10.1016/j.ab.2013.01.019

Xie YG, Luo ZH, Fang BZ, Jiao JY, Xie QJ, Cao XR, Qu YN, Qi YL, Rao YZ, Li YX, Liu YH, Li A, Seymour C, Palmer M, Hedlund BP, Li WJ, Hua ZS (2022) Functional differentiation determines the molecular basis of the symbiotic lifestyle of Ca. Nanohaloarchaeota Microbiome 10:1–13. https://doi.org/10.1186/s40168-022-01376-y

Draft Genomes of Halophilic *Chromohalobacter* and *Halomonas* Strains Isolated from Brines of The Carpathian Foreland, Poland

Jakub Lach[a,b,#], Dominik Strapagiel[b], Agnieszka Matera-Witkiewicz[c], Paweł Stączek[a,#]

a Department of Molecular Microbiology, Faculty of Biology and Environmental Protection, University of Lodz, Lodz, Poland

b Biobank Lab, Department of Oncobiology and Epigenetics, Faculty of Biology and Environmental Protection, University of Lodz. Lodz, Poland

c Screening of Biological Activity Assays and Collection of Biological Material Laboratory, Wroclaw Medical University Biobank, Faculty of Pharmacy, Wroclaw Medical University, Wroclaw, Poland

#Address correspondence to Jakub Lach and Paweł Stączek: jakub.lach@biol.uni.lodz.pl, pawel.staczek@biol.uni.lodz.pl

Abstract:

Chromohalobacter and *Halomonas* are genera of bacterial microorganisms belonging to the group of halophiles. They are characterized by high diversity and the ability to produce bioproducts of biotechnological importance, such as ectoine, biosurfactants and carotenoids. Here, we report three draft genomes of *Chromohalobacter* and two draft genomes of *Halomonas* isolated from brines. The length of the genomes ranged from 3.6 Mbp to 3.8 Mbp, and GC content was in the 60.11%-66.46% range. None of the analysed genomes has been assigned to any previously known species of the genus *Chromohalobacter* or *Halomonas*. Phylogenetic analysis revealed that *Chromohalobacter* 296-RDG and *Chromohalobacter* 48-RD10 belonged to the same species, and *Chromohalobacter* 11-W is more distantly related to the other two analysed strains than to *Chromohalobacter canadensis*. *Halomonas* strains 11-S5 and 25-S5 were clustered together and located close to *Halomonas ventosae.* Functional analysis revealed BGCs related to ectoine production in all genomes analysed. This study increases our overall understanding of halophilic bacteria and is also consistent with the notion that members of this group have significant potential as useful natural product producers.

Introduction

Hypersaline ecosystems have been widely explored mainly because of their unique biodiversity and biotechnological potential. It is related to the adaptation of halophiles to the conditions of high osmotic pressure, limited availability of energy resources and other unfavourable environmental circumstances (1). Novel halophilic microorganisms have been isolated in recent years from habitats such as saline lakes, salt mines, saline soils or fermented foods (2–6). A great variety of saline environments related to factors other than salinity itself, such as pH, temperature or the availability of nutrients, results in a considerable diversity of inhabiting microorganisms, both in terms of genetics and metabolism.

Halomonadaceae is a family of halophilic *Gammaproteobacteria,* including *Chromohalobacter*, *Halomonas* and 12 other genera (7,8). Currently, *Chromohalobacter* genus includes eight validly published species isolated from salterns, seas and food products. The process of validating the publication of new species of prokaryotes is related to the fulfilment of all of the requirements set out in the "International Code of Nomenclature of Prokaryotes" (9). Moreover, the new species *Chromohalobacter moromii* sp. nov. isolated from lupine-based moromi fermentation has been described and is pending validation (6). Microorganisms belonging to the genus *Chromohalobacter*, for example, *Chromohalobacter salexigens*, have been identified as producers of biotechnologically valuable compounds and, due to their genomic characteristics, could become a useful metabolic engineering tool for the overproduction of ectoines (10).

On the other hand, the *Halomonas* genus includes 117 validly published species and new species like *Halomonas alkalisoli* sp. nov are waiting for validation (2). Most of these species are widely distributed in saline habitats, such as salt lakes, marine environments, and saline soils (11,12). *Halomonas* strains exhibit high metabolic and physiological versatility, thus a wide range of bioproducts, such as ectoine, glycine betaine and polyhydroxyalkanoates (PHA) can be produced (12,13).

This study presents the characteristics of five genomes of microorganisms isolated from brines, sources of which are located in the southern part of the Carpathian Foreland in Poland near Kraków city. Three genome sequences of *Chromohalobacter sp.* and two draft genome sequences of *Halomonas sp* are reported.

Materials and methods

Three strains of *Chromohalobacter sp.* and two strains of *Halomonas sp.* were isolated from brines. *Chromohalobacter* 11-W was isolated from the borehole of the former Barycz mining area (49°59'05" N 20°00'52" E), *Chromohalobacter* 296-RDG and *Chromohalobacter* 48-RD10 were isolated from the Bochnia Salt Mine (49°58′09″N 20°25′03″E) and *Halomonas* strains 11-S5 and 25-S5 were isolated from the brine source in Łapczyca (49°57'30"N 20°21'41"E). Strains were cultured in 28°C on plates containing halobacteria medium (DSMZ 372) with 15% NaCl addition for strains *Chromohalobacter* 11-W, *Chromohalobacter* 296-RDG and *Halomonas* 25-S5. Rest of the strains were cultured in the same conditions except NaCl concentration which was changed to 20%. Medium was solidified with 2% agar. For genomic DNA extraction QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) has been used.

Paired-end libraries were prepared from 1 ng of high-quality genomic DNA with the Nextera XT DNA sample preparation kit according to the manufacturer's instructions (Illumina Inc., San Diego, USA). The libraries were sequenced using a NextSeq 500 instrument (Illumina, San Diego, USA) at a read length of 2 × 150 bp in Biobank Lab, University of Lodz. The quality of reads was checked using FastQC (14). Furthermore, adaptors and low-quality sequences were removed from the reads with trim galore v. 0.6.4 on default parameters (15). *De novo* assembly was performed with SPAdes v3.15.0 (16). Contigs with coverage lower than 2, or lengths lower than 500 bp, were removed from the assembly. Contamination and completeness of assemblies were calculated using CheckM based on a reference database of marker genes (17). Overall statistics of assemblies quality parameters were tested using Quast (18). Genomes annotation was performed using Prokka v.1.14.0 (19) and the eggNOG-mapper website v.5.0.0 (20). Biosynthetic gene clusters (BGCs) prediction was conducted with antiSMASH v.6.0.1 (21). Taxonomic annotation of genomes was performed with gtdb-tk version 1.5.1 (22). Phylogenomic

analysis was accomplished on Type (Strain) Genome Server (TYGS) (23), and the whole-genome sequence-based phylogenetic tree was visualized using iTOL (24). For TYGS analysis, 27 reference strain genomes were used. They are available in the NCBI database under accessions: NC_007963, NZ_BNAE01000000, NZ_CAAHFN010000000, NZ_CABVOU010000000, NZ_CP018139, NZ_FNIV01000000, NZ_FOBC01000000, NZ_FPAQ01000000, NZ_JACHXQ010000000, NZ_JACHXR010000000, NZ_JACHZF010000000, NZ_JAGXFD010000000, NZ_JAKGAJ010000000, NZ_JAKGAK010000000, NZ_JAKGAL010000000, NZ_JAKGAM010000000, NZ_JAKGAN010000000, NZ_PYVX01000000, NZ_PZJV01000000, NZ_QPIJ01000000, NZ_RXNS01000000, NZ_SDMO01000000, NZ_SNZJ01000000, NZ_SOBR01000000, NZ_VBUI01000000, NZ_WUTS01000000, NZ_WUTT01000000. Phylogenetic tree for 16S rRNA gene sequences was prepared in MEGA X software (25).

All sequencing data are publicly available from the National Institutes of Health under BioProject accessions PRJNA899688, PRJNA899690, PRJNA899692, PRJNA899693, PRJNA899694.

Results and discussion

Each draft genome was composed of between 38 and 321 contigs, with genome sizes ranging 3.6-3.9 Mbps. The overall genome completeness was estimated at between 98.71-99.86%, with contamination in the range of 0.54-8.42 and GC content in the 60.11-66.46% range. The summary is presented in Table 1.

Preliminary taxonomic annotation of genomes using gtdb-tk, assigned isolates 11-W, 296-RDG, and 48-RD10 to *Chromohalobacter* genus, and isolates 11-S5, 25-S5 to *Halomonas* genus. The assignment to the species level was impossible because of too high differences in genome sequences between the analysed strains and the previously described genomes available in the databases. A phylogenetic analysis was performed in order to deepen knowledge about the relationship between the analysed isolates and other species. A phylogenetic tree, based on the 16S rRNA gene sequences was built. The resulting tree confirmed that the isolates 11-W, 296-RDG, and 48-RD10 fell within a cluster comprising members of the genus *Chromohalobacter* and the strains 11-S5, 25-S5 fell within a cluster including members of the genus *Halomonas* (Figure 1). However, in both cases, the analysed strains were separated from the other species included in the analysis. The closest species for *Chromohalobacter* strains 296-RDG and 48-RD10 was *Chromohalobacter canadensis*, and for strain 11-W, it was *Chromohalobacter sarecensis*. In the case of *Halomonas* strains, the closest taxon was *Halomonas sediminicola.* This initial phylogenetic analysis, using a comparison of 16S rRNA gene sequences, was then deepened through the construction of a further, whole-genome sequence-based phylogenetic tree build using Genome BLAST

Distance Phylogeny approach (GBDP) created on the TYGS platform (Figure 2). The obtained phylogenetic tree confirmed the observations made at an earlier stage. On the genome-wide scale, it was noticed that *Chromohalobacter* 11-W is more distant from the other two analysed strains than from *Chromohalobacter canadensis*. That may suggest genomes assignment to two different species of *Chromohalobacter*. This observation was confirmed by digital DNA-DNA hybridization (dDDH) evaluation, where the similarity between *Chromohalobacter* 296-RDG and 48-RD10 was 87.5% (d4 method), between 11-W and 296-RDG it was 42.7 % and between 11-W and 48-RD10 it reached 42.9%. *Halomonas* strains 11-S5 and 25-S5 were clustered together on the whole-genome tree. Allocation to the same species was confirmed by dDDH which was 89.7%. The closest related species to the analysed strains was *Halomonas ventosae*.

Functional annotation of genomes revealed that they all contained numerous genes involved in the biosynthesis of secondary metabolites (Table 2). However, both *Halomonas* strains, 11-S5 and 25-S5, had a higher number of genes belonging to this category (99 and 101 genes, respectively) than isolates belonging to the genus *Chromohalobacter*, which consisted of 69-89 such genes, depending on the strain. This observation is consistent with previous reports that *Halomonas* and *Chromohalobacter* have a high diversity of biosynthetic processes (10,12,13) Based on these results, the annotation of BGCs with antiSMASH was performed. Interestingly, as a result, more BGCs were identified in *Chromohalobacter* strains than *Halomonas*, despite a smaller number of genes associated with processes identified during the analysis. Moreover, BGCs related to ectoine production have been identified in all genomes. Most likely, it is related to the adaptation of the studied microorganisms to conditions of high salinity. Ectoine produced by the analysed strains is one of the most important compatible solutes that protects the cell against high osmotic pressure (26). In the 11-W, 296-RDG, 48-RD10, and 11-S5 strains, a complete operon *ectABC* was identified. In the 25-S5 strain, only the *ectC* gene, essential for ectoine production, was identified. In *Halomonas* isolates, BGCs associated with ectoine production were the only BGCs identified in the genomes. In *Chromohalobacter* strains, BGCs related to the production of siderophores, redox-cofactors, and arylopolyenes, were also identified. Table 3 summarizes the information on the identified BGCs in each of the strains.

To summarize, the draft genomes of three *Chromohalobacter* strains and two *Halomonas* strains expand the genomic representation in the tree of life. The strains analysed were isolated from the hitherto unexplored saline environment, which allows a deeper understanding of their biodiversity.

Acknowledgements:

This research was financed from the following funds: "InterDOC-STARt" project (POWR.03.02.00-00-I033/16-00), the MINI InterDOC-STARt grant and by the subsidy for scientific activity provided to the Department of Molecular Microbiology by the University of Lodz (B2111000000038.01). The authors are grateful to the management of the Bochnia Salt Mine for granting permission for sampling, Mr Tomasz Migdas, the Technical Director of the Bochnia Salt Mine, for his invaluable help in organizing sampling and for inspiring conversation, and to the mine geologists: Michał Flasza and Aleksandra Puławska for guiding the sampling tours. We would also like to thank Dr Magdalena Kowalewicz-Kulbat and Dr Dominika Drzewiecka for preparing pure strain cultures and Professor Łukasz Dziewit for a fruitful discussion and helpful advice regarding this publication.

Table 1. Genome features

	***Chromohalobacter* 11-W**	***Chromohalobacter* 296-RDG**	***Chromohalobacter* 48-RD10**	***Halomonas* 11-S5**	***Halomonas* 25-S5**
Genome length (bp)	3 660 465	3 763 726	3 785 492	3 634 207	3 872 356
Number of contigs	38	321	222	65	101
Largest contig (bp)	607 581	256 035	627 765	320 470	407 008
GC content (%)	60.83	60.11	60.3	66.46	65.9
N50 (bp)	318 640	58 943	228 913	105 985	122 550
Number of CDSs	3 360	3 576	3 572	3 357	3 551
Number of rRNAs	6	3	3	7	5
Number of tRNAs	63	63	65	60	58
Number of repeat regions	3	2	3	1	2
Completeness (%)	98.71	98.71	99.57	99.86	99.86
Contamination (%)	0.86	0.86	8.42	0.68	0.54

Table 2. eggNOG categories of coding proteins

Class	**Description**	***Chromohalobacter* 11-W [%]**	***Chromohalobacter* 296-RDG [%]**	***Chromohalobacter* 48-RD10 [%]**	***Halomonas* 11-S5 [%]**	***Halomonas* 25-S5 [%]**
Information storage and processing						
J	Translation, ribosomal structure, and biogenesis	187 [5.85]	184 [5.77]	193 [5.91]	196 [6.14]	194 [5.83]
A	RNA processing and modification	0	0	0	0	0
K	Transcription	265 [8.29]	259 [8.12]	269 [8.23]	223 [6.99]	238 [7.15]
L	Replication, recombination, and repair	129 [4.04]	187 [5.86]	211 [6.46]	178 [5.58]	232 [6.97]

B	Chromatin structure and dynamics	2 [0.06]	1 [0.03]	1 [0.03]	4 [0.13]	4 [0.12]
Cellular processes and signalling						
D	Cell cycle control, cell division, chromosome partitioning	50 [1.56]	47 [1.47]	52 [1.59]	48 [1.50]	50 [1.50]
Y	Nuclear structure	0	0	0	0	0
V	Defence mechanisms	34 [1.06]	41 [1.28]	34 [1.04]	50 [1.57]	53 [1.59]
T	Signal transduction mechanisms	135 [4.23]	119 [3.73]	124 [3.80]	156 [4.89]	155 [4.66]
M	Cell wall/membrane/envelope biogenesis	201 [6.29]	227 [7.11]	240 [7.35]	171 [5.36]	176 [5.29]
N	Cell motility	81 [2.54]	80 [2.51]	86 [2.63]	64 [2.01]	62 [1.86]
Z	Cytoskeleton	0	0	0	0	0
W	Extracellular structures	0	0	0	0	0
U	Intracellular trafficking, secretion, and vesicular transport	69 [2.16]	76 [2.38]	77 [2.36]	61 [1.91]	60 [1.80]
O	Posttranslational modification, protein turnover, chaperones	114 [3.57]	108 [3.38]	119 [3.64]	138 [4.33]	137 [4.12]
Metabolism						
C	Energy production and conversion	234 [7.32]	203 [6.36]	213 [6.52]	239 [7.49]	250 [7.51]
G	Carbohydrate transport and metabolism	223 [6.98]	224 [7.02]	219 [6.70]	169 [5.30]	171 [5.14]
E	Amino acid transport and metabolism	337 [10.55]	341 [10.69]	330 [10.10]	310 [9.72]	308 [9.26]
F	Nucleotide transport and metabolism	91 [2.85]	88 [2.76]	88 [2.69]	89 [2.79]	92 [2.77]

H	Coenzyme transport and metabolism	148 [4.63]	151 [4.73]	158 [4.84]	149 [4.67]	151 [4.54]
I	Lipid transport and metabolism	119 [3.72]	108 [3.38]	107 [3.28]	135 [4.23]	132 [3.97]
P	Inorganic ion transport and metabolism	231 [7.23]	224 [7.02]	213 [6.52]	233 [7.30]	237 [7.12]
Q	Secondary metabolites biosynthesis, transport, and catabolism	89 [2.79]	73 [2.29]	69 [2.11]	99 [3.10]	101 [3.04]
Poorly characterized						
R	General function prediction only	0	0	0	0	0
S	Function unknown	618 [19.34]	614 [19.24]	626 [19.16]	616 [19.31]	646 [19.42]
All proteins		3195	3191	3267	3190	3327

Table 3. BGCs identified with AntiSMASH in the analysed genomes

Sample name	***Chromohalobacter* 11-W**	***Chromohalobacter* 296-RDG**	***Chromohalobacter* 48-RD10**	***Halomonas* 11-S5**	***Halomonas* 25-S5**
AntiSMASH					
# of BGC	5	7	4	1	1
Arylpolyene	1	1	1	0	0
Betalactone	1	1	0	0	0
Butyrolactone	0	1	0	0	0
Ectoine	1	1	1	1	1
Siderophore	1	1	1	0	0
Phosphonate	0	1	0	0	0
Redox-cofactor	1	1	1	0	0

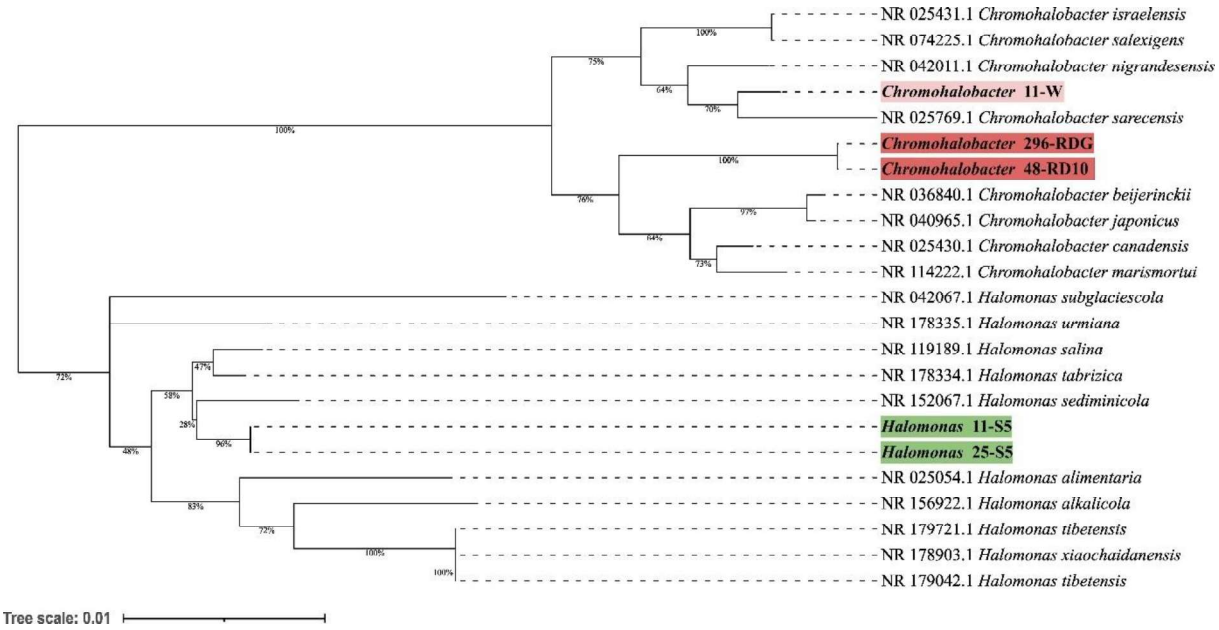


Figure 1. Phylogenetic tree prepared based on 16S rRNA gene sequences analysis using the Neighbor-Joining method and showing the relationships between analysed strains and other *Chromohalobacter* and *Halomonas* strains. The evolutionary distances were computed using the Kimura 2-parameter method.

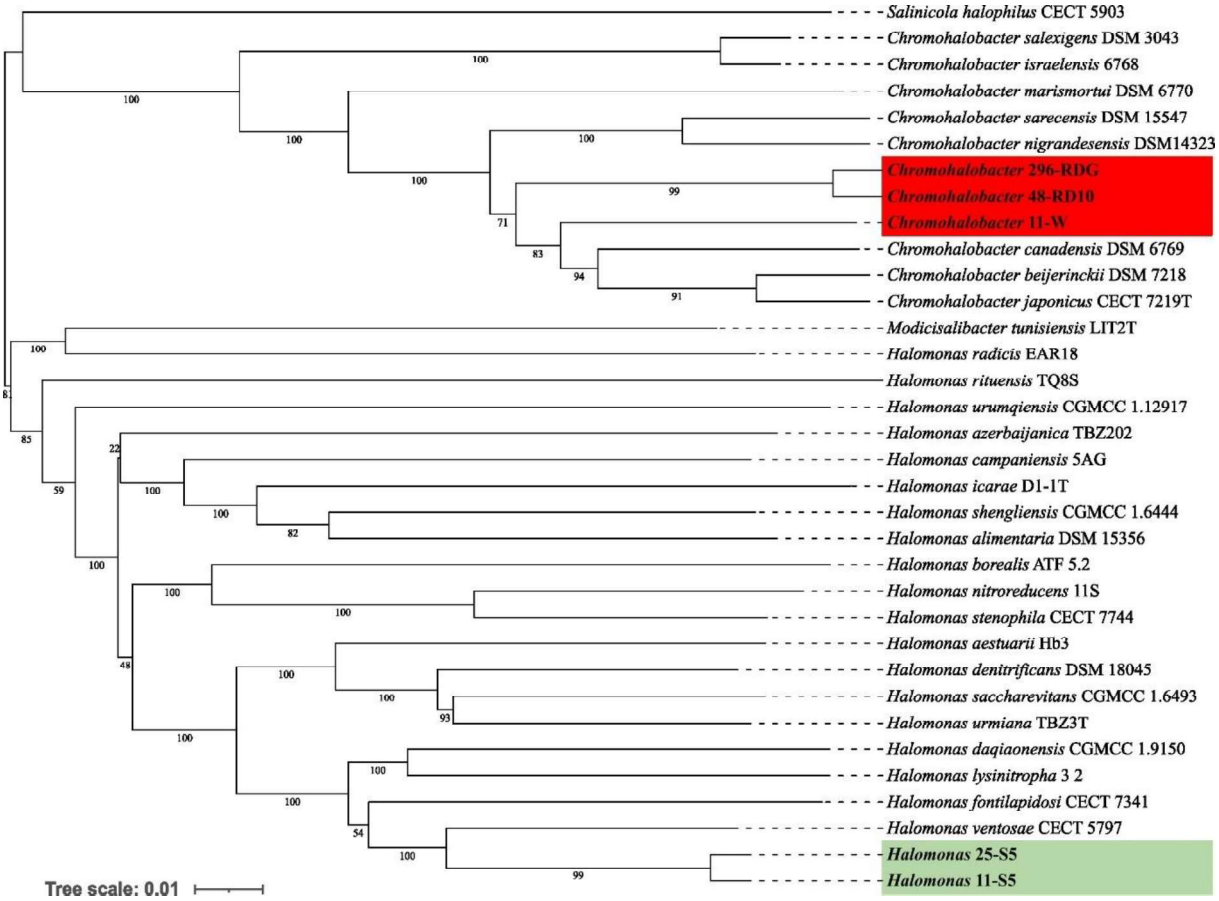


Figure 2. Whole-genome sequence-based phylogenetic tree build using Genome BLAST Distance Phylogeny approach (GBDP) on the Type (Strain) Genome Server (TYGS) platform showing the phylogenetic relationships between analysed Bacteria strains and other close related to them species.

References

1. Yin J, Chen JC, Wu Q, Chen GQ. Halophiles, coming stars for industrial biotechnology. Biotechnol Adv [Internet]. 2015;33(7):1433–42. Available from: http://dx.doi.org/10.1016/j.biotechadv.2014.10.008

2. Shang J, Xu L, Yang R, Zhao Y, Tang S, Sun J. Halomonas alkalisoli sp . nov ., a novel haloalkalophilic species from saline-alkaline soil , and reclassification of Halomonas daqingensis Wu et al . 2008 as a later heterotypic synonym of Halomonas desiderata. Syst Appl Microbiol [Internet]. 2022;45(5):126351. Available from: https://doi.org/10.1016/j.syapm.2022.126351

3. Gong Q, Zhao P, Miao S, Yi K, Ma C, Zhang G. Halorubrum salipaludis sp. nov., isolated from the saline–alkaline soil. Arch Microbiol [Internet]. 2022;204(1):1–6. Available from: https://doi.org/10.1007/s00203-021-02729-1

4. Zuo Z, Zhao D, Zhou J, Han J, Xiang H. Halalkalirubrum salinum gen. nov., sp. nov., a halophilic archaeon isolated from a saline lake. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol [Internet]. 2021;114(1):83–94. Available from: https://doi.org/10.1007/s10482-020-01502-6

5. Sun S, Chen F, Xu Y, Liu J, Chen S. Halorubrum amylolyticum sp. nov., a novel halophilic archaeon isolated from a salt mine. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol [Internet]. 2019;112(12):1849–61. Available from: https://doi.org/10.1007/s10482-019-01313-4

6. Lülf RH, Hilgarth M, Ehrmann MA. Chromohalobacter moromii sp. nov., a moderately halophilic bacterium isolated from lupine-based moromi fermentation. Syst Appl Microbiol. 2022;45(4).

7. Ventosa A, Gutierrez MC, Garcia MT, Ruiz-Berraquero F. Classification of "Chromobacterium marismortui" in a new genus, Chromohalobacter gen. nov., as Chromohalobacter marismortui comb. nov., nom. rev. Int J Syst Bacteriol. 1989;39(4):382–6.

8. Vreeland RH, Litchfield CD, Martin EL, Elliot E. Halomonas elongata, a new genus and species of extremely salt-tolerant bacteria. Int J Syst Bacteriol. 1980;30(2):485–95.

9. Parker CT, Tindall BJ, Garrity GM. International code of nomenclature of Prokaryotes. Int J Syst Evol Microbiol. 2019;69(1):S1.

10. Argandoña M, Piubeli F, Reina-Bueno M, Nieto JJ, Vargas C. New insights into hydroxyectoine synthesis and its transcriptional regulation in the broad-salt growing halophilic bacterium Chromohalobacter salexigens. Microb Biotechnol. 2021;14(4):1472–93.

11. Navarro-Torre S, Carro L, Rodríguez-Llorente ID, Pajuelo E, Caviedes MÁ, Igual JM, et al. Halomonas radicis sp. nov., isolated from Arthrocnemum macrostachyum growing in the Odiel marshes(Spain) and emended descriptions of Halomonas xinjiangensis and Halomonas zincidurans. Int J Syst Evol Microbiol. 2020;70(1):220–7.

12. Poli A, Nicolaus B, Denizci AA, Yavuzturk B, Kazan D. Halomonas smyrnensis sp. nov., a moderately halophilic, exopolysaccharide-producing bacterium. Int J Syst Evol Microbiol. 2013;63(1):10–8.

13. Thomas T, Elain A, Bazire A, Bruzaud S. Complete genome sequence of the halophilic PHA-producing bacterium Halomonas sp. SF2003: insights into its biotechnological potential. World J Microbiol Biotechnol [Internet]. 2019;35(3):1–14. Available from: http://dx.doi.org/10.1007/s11274-019-2627-8

14. Andrews S. FastQC: a quality control tool for high throughput sequence data [Internet]. 2022. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

15. Krueger F. Trim Galore [Internet]. 2022. Available from: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

16. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77.

17. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25(7):1043–55.

18. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics. 2018;34(13):i142–50.

19. Seemann T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068–9.

20. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 2019;47(D1):D309–14.

21. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, Van Wezel GP, Medema MH, et al. AntiSMASH 6.0: Improving cluster detection and comparison capabilities. Nucleic Acids Res. 2021;49(W1):W29–35.

22. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. Bioinformatics. 2020;36(6):1925–7.

23. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun [Internet]. 2019;10(1). Available from: http://dx.doi.org/10.1038/s41467-019-10210-3

24. Letunic I, Bork P. Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293–6.

25. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547–9.

26. Czech L, Hermann L, Stöveken N, Richter AA, Höppner A, Smits SHJ, et al. Role of the extremolytes ectoine and hydroxyectoine as stress protectants and nutrients: Genetics, phylogenomics, biochemistry, and structural analysis. Genes (Basel). 2018;9(4):1–58.

Re: Submission 80829s1 - manuscript word files for copyediting and fee

Journal of Genomics <publish@jgenomics.com>

Pn, 13.02.2023 00:24

Do: Jakub Lach <jakub.lach@biol.uni.lodz.pl>

DW: publish@jgenomics.com <publish@jgenomics.com>;j.roch.lach@gmail.com <j.roch.lach@gmail.com>

Dear Mr. Jakub Lach,

The revision of your manuscript 80829s1 is acceptable in Journal of Genomics and we are currently scheduling it for publication. Please check the manuscript carefully again and then submit a last, final version (in the original Microsoft Word file) to us for copyediting and publication. Since changes after copyediting and generation of an xml version are very difficult, it is important that any changes you would like to make are included before submitting the final files. Additional changes are generally not permitted after copyediting.

Please submit the final manuscript Word file after signing in https://www.jgenomics.com/ms/status

(Manuscript ID: 80829s1, Login code: RPDBF)

References: please ensure your references follow the format in
https://www.jgenomics.com/ms/author
Make sure publication years and other information are not missing for journal papers.

Figures:
* Please make sure images are of the highest resolution, and any letters inside are clear and legible. Images should have resolution >600 dpi and line art >1200dpi.
* If any figures are adapted or taken from other existing publications, please obtain permissions from the publishers or copyright owners as soon as possible.

Publication fee: the invoice is available in the same manuscript login page after signing in. The publication fee is required in our journal to cover the cost of publication, and should be paid before copyediting and publication. Please proceed to make the payment. Thank you for your support.

Kind regards,
Publishing Team
Journal of Genomics
https://www.jgenomics.com

Manuscript information
Title: Draft Genomes of Halophilic Chromohalobacter and Halomonas Strains Isolated from Brines of The Carpathian Foreland, Poland
Author(s): Jakub Lach, Dominik Strapagiel, Agnieszka Matera-Witkiewicz, Paweł Stączek

Draft genomes of halophilic Archaea strains isolated from brines of the Carpathian Foreland, Poland

Jakub Lach[a,b,#], Dominik Strapagiel[b], Agnieszka Matera-Witkiewicz[c], Paweł Stączek[a,#]

a - Department of Molecular Microbiology, Faculty of Biology and Environmental Protection, University of Lodz, Lodz, Poland

b - Biobank Lab, Department of Oncobiology and Epigenetics, Faculty of Biology and Environmental Protection, University of Lodz. Lodz, Poland

c - Screening of Biological Activity Assays and Collection of Biological Material Laboratory, Wroclaw Medical University Biobank, Faculty of Pharmacy, Wroclaw Medical University, Wroclaw, Poland

#Address correspondence to Jakub Lach and Paweł Stączek: jakub.lach@biol.uni.lodz.pl, pawel.staczek@biol.uni.lodz.pl

Abstract

Halophilic Archaea are a unique group of microorganisms living in saline environments. They constitute a complex group whose biodiversity has not been thoroughly studied. Here, we report three draft genomes of halophilic Archaea isolated from brines, representing the genera of *Halorubrum*, *Halopenitus,* and *Haloarcula*. Two of these strains, Boch-26 and POP-27, were identified as members of the genera *Halorubrum* and *Halopenitus*, respectively. However, they could not be assigned to any known species because of the excessive difference in genome sequences between these strains and any other described genomes. In contrast, the third strain, Boch-26, was identified as *Haloarcula hispanica*. Genome lengths of these isolates ranged from 2.7 Mbp to 3.0 Mbp, and GC content was in the 63.77%-68.77% range. Moreover, functional analysis revealed biosynthetic gene clusters (BGCs) related to terpenes production in all analysed genomes and one BGC for RRE (RiPP recognition element)-dependent RiPP (post-translationally modified peptides) biosynthesis. Moreover, the obtained results enhanced the knowledge about the salt mines microbiota biodiversity as a poorly explored environment so far.

Introduction

Archaea as representatives of an extremely diverse group of organisms are present in almost all types of environments worldwide. They can be found in the guts of humans and animals, in harsh habitats such as salt mines or hot springs and in food products [1–3]. Over the years, research on Archaea has been very limited due to the limited availability of methods for studying these microorganisms. However, in the last decade, significant progress has been made in understanding the taxonomic and metabolic diversity of this group of organisms [4]. It is associated with the development of research methods related to high-throughput DNA sequencing and metagenome analyses, which provided greater insight into the microbial "dark matter" [5]. Metagenomic analysis of various environments has allowed the identification of

new taxonomic groups among Archaea, such as Asgard archaea or DPANN [6,7]. Research is currently underway to develop Archaea cultivation methods, which remains a significant challenge, and for many taxa, pure cultures have still not been isolated [8].

Archaea constitute a significant part of extremophilic microorganisms that are adapted to survive in conditions inaccessible to the most known microbes [4]. This group also includes halophilic microorganisms living in high-salinity environments. Most of them can be found in the classes *Halobacteria* and *Methanomicrobia* belonging to the phylum *Euryarcheota* [9]. Due to their ability to live in high salinity conditions, these microorganisms can be used for industrial purposes in bioprocesses [10,11]. Bioproducts such as ectoine, polyhydroxyalkanoate (PHA) or extremozymes produced by halophiles found applications in many areas of biotechnology including the production of polyunsaturated fatty acids, biopolymers and osmoprotectants [11,12]. Research was also conducted to identify and characterize compounds that may find application in medicine, as was the case with Actinomycin C2, Streptomonomicin or halocins [13].

The microbiota of salt mines and brines is still not very well explored. Only a few strains of microorganisms have been isolated from this kind of habitat, and only a few metagenomes have been sequenced. Such species, like *Halorhabdus rudnickae*, *Halorubrum trueperi* and *Halorubrum amylolyticum*, were identified for the first time in samples collected from such environments [14–16]. Metagenomic analysis for salt mines has been performed so far for the Karak Salt Mine, Pakistan [2]. Despite the research carried out, there are still significant gaps in the knowledge about halophiles inhabiting salt mines and brines what determines the need for a thorough study of such environments for a better understanding their biodiversity and to discover possible applications of extremophiles.

The salt mine in Bochnia (southern Poland) is an example of an environment that, due to its uniqueness, can be a habitat for many previously unknown microbial strains and which

has not yet been thoroughly investigated. It was established in the 13th century and was actively exploited until 1990. It means that it was one of the oldest and longest-exploited salt mines in Europe. Due to its long history and unique character, it was placed on the UNESCO World Heritage List [17]. It was first established in a fragment of the marine sediments of the Miocene salt-bearing formation. Currently, the historic mine consists of nine post-mining galleries reaching 350 m below the surface, and a significant part of the excavations is open to the public [18]. Since the end of salt mining, the activity of the Bochnia salt mine has changed its character and focused on tourism, recreation, and health protection.

In this paper, we report three genomes of Archaea isolated from brines collected from the Bochnia Salt Mine, located in southern Poland, near the city of Kraków. The analysed strains were characterized in terms of taxonomy and functionality, thus enriching the knowledge of the microbiota of the Bochnia Salt Mine.

Materials and methods

Three strains of halophilic Archaea were isolated from brines collected in the Bochnia Salt Mine (49°58′09″N 20°25′03″E). Isolation of strains was carried out in accordance with the methodology used in earlier studies by Albuquerque et al. [15]. All strains were cultured on plates containing halobacteria medium (DSMZ 372) with 25% (w/v) NaCl concentration. The culture temperature was 37°C for the strains *Halorubrum* Boch-26 and *Halopenitus* POP-27, and 28°C for the strain *Haloarcula hispanica* Boch-4. Genomic DNAs were extracted using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany).

For each strain, 1 ng of high-quality genomic DNA was used for DNA library preparation. Nextera XT DNA sample preparation kit (Illumina Inc., San Diego, USA) was used for the preparation of paired-ends libraries. The libraries were sequenced in a 2 × 150 bp configuration using a NextSeq 500 instrument (Illumina, San Diego, USA). The quality of reads

was checked with FastQC [19]. Adaptors and low-quality sequences were trimmed out from the reads with Trim galore! v. 0.6.4 set on default parameters [20]. For *de novo* genomes assembly, SPAdes v3.15.0 was used [21]. Contigs shorter than 500 bp or with sequencing depths lower than 2 were removed from the assemblies. For quality control of assemblies, CheckM [22] and Quast [23] were used. Genome annotation was performed using Prokka v.1.14.0 [24] and the eggNOG-mapper website v.5.0.0 [25]. Identification of biosynthetic gene clusters (BGCs) was performed with antiSMASH v.6.0.1 [26]. Taxonomic annotation of genomes was performed with gtdb-tk version 1.5.1 [27]. Type (Strain) Genome Server (TYGS) was used for phylogenomic analysis [28], and the resulting whole-genome sequence-based phylogenetic trees were visualized by iTOL [29]. The d4 formula from TYGS was used to calculate dDDH. This metric was calculated as a sum of all identities found in high-scoring segment pairs (HSPs) divided by overall HSP length. For TYGS analysis, genomes of 29 reference strains were used. They are available in NCBI database under accessions: FOPZ00000000, VCNM00000000, JANHDN000000000, AOLR00000000, FNBO00000000, AOLS00000000, VCNL00000000, LOAJ00000000, SDJP00000000, FXTD00000000, LKIR00000000, NHPJ00000000, FNPC00000000, BMON00000000, BMPD00000000, AOLY00000000, AOJG00000000, AOLW00000000, AOLX00000000, AOJH00000000, BBJP00000000, AOJE00000000, RBWW00000000, CP002921-CP002923, AY596290-AY596298, CP001365-CP001367, HF582854, NSKC00000000, FNWU00000000. A phylogenetic tree for *gyrB* sequences was prepared in MEGA X software [30]. The evolutionary distances between sequences were computed using the Kimura 2-parameter method.

All sequencing data are publicly available from the National Institutes of Health under WGS accessions JAQYWM000000000, JARANU000000000, JARANT000000000.

Results and discussion

The three draft genomes sequences of the strains analysed contained between 33 and 177 contigs. Their sizes ranged from 2.7-3.0 Mbps. Completeness of the genomes was assessed using CheckM and was between 95.97%-97.93%. Contamination of genomes has been estimated at a level of 0.0%-0.95%, and the GC content was in the 63.77%-68.77% range. General genome feature statistics have been summarized in Table 1.

Based on the genomic sequences, a preliminary taxonomic annotation was performed using the gtdb-tk tool. During the analysis, strains Boch-4 and POP-27 were assigned to *Haloarcula hispanica* and *Halopenitus malekzadehii,* respectively. For strain Boch-26, assignment to species was impossible due to the too high difference between the genome sequence of this strain and those of other known genomes collected in the GTDB (the Genome Taxonomy Database). In order to better describe the evolutionary relationships between the strains analysed and other closely related species, a phylogenetic tree based on sequences of the *gyrB* gene was prepared. The reconstructed phylogenetic tree partially confirmed the taxonomic classification performed with gtdb-tk (Figure 1). Strain Boch-26 was placed close to other *Halorubrum* species but did not form a single cluster with any of them. The same was true for strain Boch-4, which was placed in a single cluster with *Haloarcula hispanica.* In the case of strain POP-27, it was clustered together with *Halopenitus persicus,* but the sequence of the *gyrB* gene for *Halopenitus malekzadehii* was unavailable. To verify the results obtained from gtdb-tk and *gyrB* analysis, a phylogenetic tree based on the whole-genome sequences was prepared using the Genome BLAST Distance Phylogeny approach. The location of strains Boch-4 and Boch-26 on the reconstructed phylogenetic tree confirmed their taxonomic annotation. Strain Boch-4 clustered with *Haloarcula hispanica*, and the digital DNA-DNA hybridization (dDDH) value between these two genomes was 82.6%. Therefore, all analyses performed clearly indicated that strain Boch-4 belongs to the species *Haloarcula hispanica*. In the case of strain

Boch-26, its taxonomic classification is also unequivocal. On a whole-genome sequence-based phylogenetic tree, this strain was placed among other *Halorubrum* species, but its distance from its other close relative, which was *Halorubrum depositum*, was too great to identify both strains as the same species. It was also confirmed by low dDDH between these two strains, which was 35.45%. The results for the POP-27 strain remain inconclusive. On the phylogenetic tree, this strain was located close to *Halopenitus malekzadehii* but dDDH between these two genomes was 59.2%, which proves that the genomes do not belong to the same species of microorganisms. The difference in the classification indications based on gtdb-tk and dDDH means that it is impossible to conclude unequivocally from the available data whether the POP-27 strain belongs to the species *Halopenitus malekzadehii*. Accordingly, the strain will be referred to as *Halopenitus* POP-27.

In order to characterise the functional profiles of the strains studied, a functional annotation of the genomes analysed was performed. The profiles obtained, based on eggNOG categories, are shown in Table 2. The *Halopenitus* strain POP-27 was characterised by a higher number of genes related to defence mechanisms, which amounted to 31 [1.27%] in this genome, while for the strains *Halorubrum* Boch-26 and *Haloarcula hispanica* Boch-4 these numbers were 15 [0.59%] and 27 [0.90%], respectively. The reverse was true for genes related to cell motility. In the *Halopenitus* strain POP-27, there was a significantly lower number of such genes (17 [0.70%]) than in the other two strains, *Halorubrum* Boch-26 (29 [1.15%]) and *Haloarcula hispanica* Boch-4 (41 [1.37%]). It is also significant that genes related to the biosynthesis of secondary metabolites were detected in all genomes. The secondary metabolites produced may be related to the adaptation of the studied strains to a harsh environment and may be of interest for industrial application. The content of these genes in particular strains was similar and amounted to 42 [1.66%], 37 [1.51%], and 43 [1.44%] for strains *Halorubrum* Boch-26, *Halopenitus* POP-27, and *Haloarcula hispanica* Boch-4, respectively. The results obtained

were the basis for an attempt to identify the BGCs present in the analysed genomes. AntiSMASH was used to identify BGCs. Two BGCs associated with terpene production were identified in all three strains. These clusters were not identical between the strains however, each one was organised around genes coding phytoene synthase. It can therefore be assumed that these strains may produce carotenoids. Three BGCs were identified in the *Haloarcula hispanica* strain. As previously mentioned two of them were associated with the terpene production, and the last one with RRE (RiPP recognition element)-dependent RiPP (ribosomally synthesized and post-translationally modified peptides) biosynthesis. Both of BGC related to the terpene production were organised around genes coding phytoene synthase. This enzyme is usually involved in carotenoid production. The first BGC contained 23 genes, and antiSMASH identified two strains (*Haloarcula* sp. CBA1115 and *Haloarcula* sp. K1K1) containing BGCs in which 100% of the genes showed similarity to genes in the BGC analysed. The second BGC consisted of 20 genes, and BGCs in which 95% of the genes showed similarity to genes in the analysed BGC were found in strains such as *Haloarcula* sp. K1K1, *Haloarcula hispanica* ATCC33960, and *Haloarcula vallismortis* DSM3756. The third BGC associated with RiPP biosynthesis contained 19 genes, and antiSMASH did not identify any highly similar BGCs. The most similar BGC with 18% of similar genes was associated with lasso peptide production and was identified in *Halopiger xanaduensis* SH-6 plasmid. The BGC analysed contained genes related to the synthesis of B1 and B2 proteins involved in lasso peptides production. A gene encoding a kinase that plays an important role in lasso peptide synthesis of was also identified in this BGC.

Concluding, the draft genomes of three Archaea strains have been reported and characterised in this paper. These genomes represent a significant value due to the uniqueness and isolation of the environment from which they were extracted from other saline environments. Significant differences in genome sequences compared to other known

187 halophiles provide a unique insight into the diversity of microorganisms that can inhabit salt
188 mines. The analyses conducted indicate that two of the strains analysed, *Halorubrum* Boch-26
189 and *Halopenitus* POP-27, may belong to hitherto unknown species. However, confirmation of
190 this observation requires further analyses in aspects related to the physiology and morphology
191 of the isolates studied. The obtained results allow us to expand our knowledge about the
192 biodiversity of the halophilic Archaea living in salt mine environments.

193 ## Acknowledgments

194 This research was financed from "InterDOC-STARt" project (POWR.03.02.00-00-
195 I033/16-00), the MINI InterDOC-STARt grant and by the subsidy for scientific activity
196 provided to the Department of Molecular Microbiology by the University of Lodz
197 (B2111000000038.01). The authors are grateful to the management of the Bochnia Salt Mine
198 for granting permission for sampling, Mr Tomasz Migdas, the Technical Director of the
199 Bochnia Salt Mine, for his invaluable help in organizing sampling and for inspiring
200 conversation, and to the mine geologists: Michał Flasza and Aleksandra Puławska for guiding
201 the sampling tours. We would also like to thank Dr Magdalena Kowalewicz-Kulbat and Dr
202 Dominika Drzewiecka for preparing pure strain cultures and Professor Łukasz Dziewit for a
203 fruitful discussion and helpful advice regarding this publication.

204

205 **Table 1**. Genome features

	Archaea		
	***Halorubrum* Boch-26**	***Halopenitus* POP-27**	***Haloarcula hispanica* Boch-4**
Genome length (bp)	2 763 236	2 670 493	3 024 851
Number of contigs	177	33	34
Largest contig (bp)	66 542	332 978	285 480
GC content (%)	68.77	66.27	63.77
N50 (bp)	26 505	116 967	136 784
Number of CDSs	2 749	2 602	3 061
Number of rRNAs	1	1	1
Number of tRNAs	42	48	46
Number of repeat regions	0	0	0
Completeness (%)	95.97	97.84	97.93
Contamination (%)	0.95	0.38	0.0

206

207 **Table 2**. eggNOG categories of genes present in in the analysed genomes

Class	**Description**	**Archaea**		
		***Halorubrum* Boch-26**	***Halopenitus* POP-27**	***Haloarcula hispanica* Boch-4**
Information storage and processing				
J	Translation, ribosomal structure, and biogenesis	122 [4.82]	156 [6.38]	158 [5.29]
A	RNA processing and modification	0	0	0
K	Transcription	160 [6.32]	174 [7.12]	196 [6.56]

L	Replication, recombination, and repair	126 [4.98]	125 [5.11]	128 [4.28]
B	Chromatin structure and dynamics	1 [0.04]	2 [0.08]	1 [0.03]
Cellular processes and signalling				
D	Cell cycle control, cell division, chromosome partitioning	21 [0.83]	23 [0.94]	25 [0.84]
Y	Nuclear structure	0	0	0
V	Defence mechanisms	15 [0.59]	31 [1.27]	27 [0.90]
T	Signal transduction mechanisms	109 [4.31]	71 [2.91]	128 [4.28]
M	Cell wall/membrane/envelope biogenesis	55 [2.17]	58 [2.37]	80 [2.68]
N	Cell motility	29 [1.15]	17 [0.70]	41 [1.37]
Z	Cytoskeleton	0	0	0
W	Extracellular structures	0	0	0
U	Intracellular trafficking, secretion, and vesicular transport	17 [0.67]	17 [0.70]	16 [0.54]
O	Posttranslational modification, protein turnover, chaperones	82 [3.24]	90 [3.68]	100 [3.35]
Metabolism				
C	Energy production and conversion	172 [6.80]	148 [6.06]	169 [5.65]
G	Carbohydrate transport and metabolism	83 [3.28]	64 [2.62]	74 [2.48]

E	Amino acid transport and metabolism	247 [9.76]	258 [10.56]	248 [8.30]
F	Nucleotide transport and metabolism	71 [2.81]	64 [2.62]	69 [2.31]
H	Coenzyme transport and metabolism	106 [4.19]	132 [5.40]	128 [4.28]
I	Lipid transport and metabolism	64 [2.53]	62 [2.54]	78 [2.61]
P	Inorganic ion transport and metabolism	133 [5.26]	133 [5.44]	138 [4.62]
Q	Secondary metabolites biosynthesis, transport, and catabolism	42 [1.66]	37 [1.51]	43 [1.44]
Poorly characterized				
R	General function prediction only	0	0	0
S	Function unknown	460 [18.18]	405 [16.57]	513 [17.16]
All proteins		2530	2444	2989

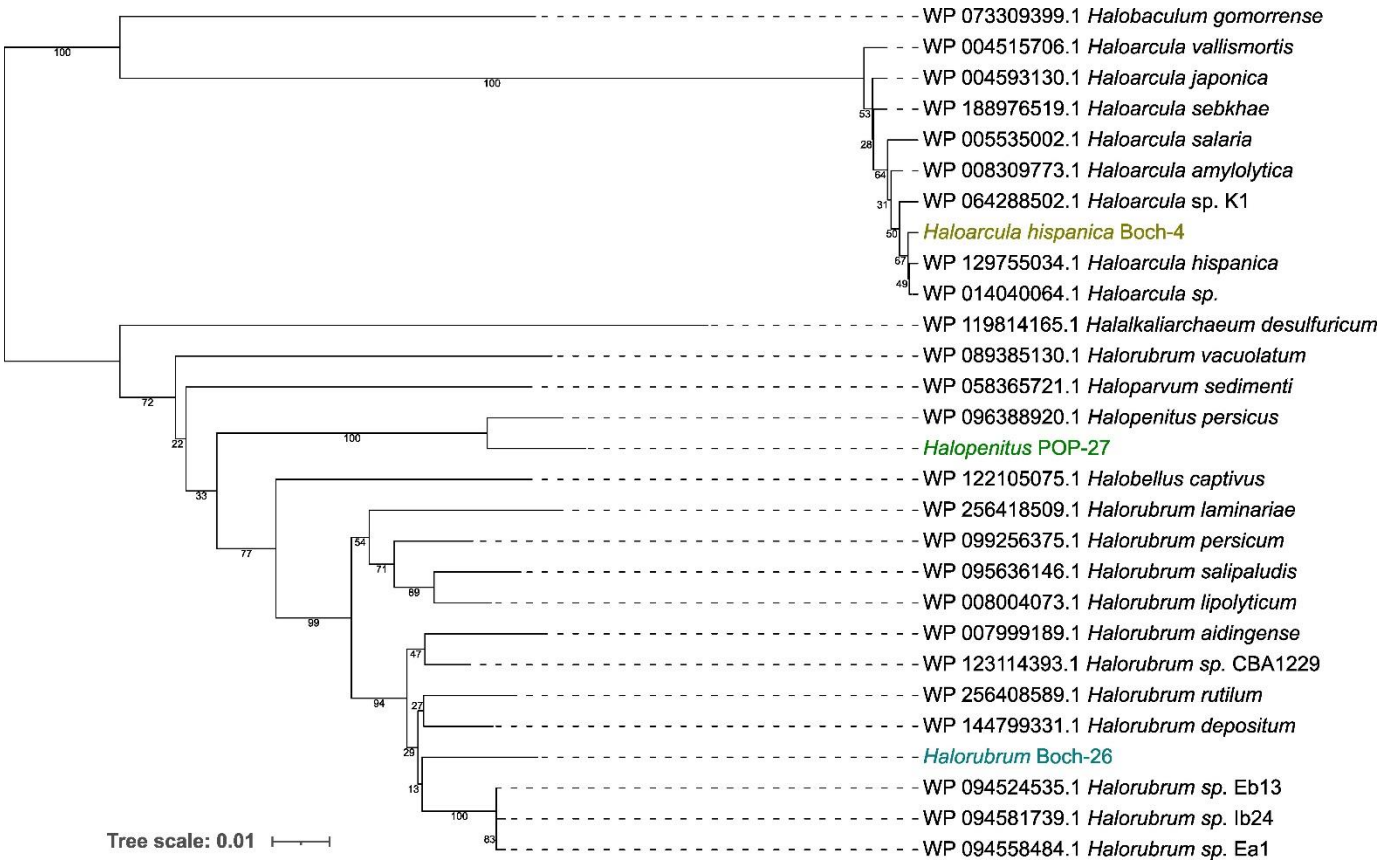


Figure 1. Phylogenetic tree based on the analysis of the *gyrB* sequences using the Neighbor-Joining method and showing the relationships between analysed strains and other *Chromohalobacter* and *Halomonas* strains.

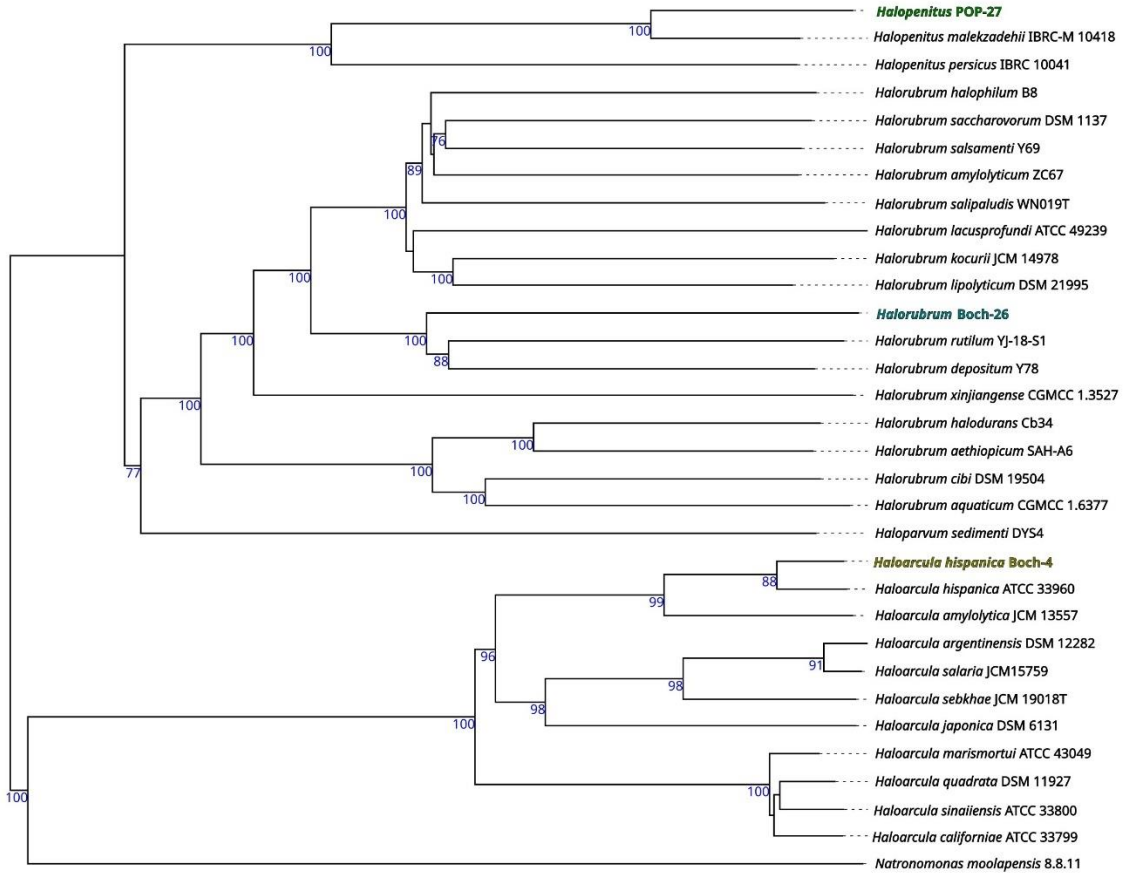


Figure 2. Whole-genome sequence-based phylogenetic tree build with Genome BLAST Distance Phylogeny approach (GBDP) by Type (Strain) Genome Server (TYGS).

References

1. Kim, J.Y.; Whon, T.W.; Lim, M.Y.; Kim, Y.B.; Kim, N.; Kwon, M.S.; Kim, J.; Lee, S.H.; Choi, H.J.; Nam, I.H.; et al. The human gut archaeome: Identification of diverse haloarchaea in Korean subjects. *Microbiome* **2020**, *8*, 1–17, doi:10.1186/s40168-020-00894-x.

2. Cycil, L.M.; DasSarma, S.; Pecher, W.; McDonald, R.; AbdulSalam, M.; Hasan, F. Metagenomic Insights Into the Diversity of Halophilic Microorganisms Indigenous to the Karak Salt Mine, Pakistan. *Front. Microbiol.* **2020**, *11*, 1–14, doi:10.3389/fmicb.2020.01567.

3. Kim, Y.B.; Kim, J.Y.; Song, H.S.; Lee, S.H.; Shin, N.R.; Bae, J.W.; Myoung, J.; Lee, K.E.; Cha, I.T.; Rhee, J.K.; et al. Haloplanus rubicundus sp. nov., an extremely halophilic archaeon isolated from solar salt. *Syst. Appl. Microbiol.* **2020**, *43*, 126085, doi:10.1016/j.syapm.2020.126085.

4. Baker, B.J.; De Anda, V.; Seitz, K.W.; Dombrowski, N.; Santoro, A.E.; Lloyd, K.G. Diversity, ecology and evolution of Archaea. *Nat. Microbiol.* **2020**, *5*, 887–900, doi:10.1038/s41564-020-0715-z.

5. Lok, C. Mining the microbial dark matter. *Nature* **2015**, *522*, 270.

6. Zaremba-Niedzwiedzka, K.; Caceres, E.F.; Saw, J.H.; Bäckström, Di.; Juzokaite, L.; Vancaester, E.; Seitz, K.W.; Anantharaman, K.; Starnawski, P.; Kjeldsen, K.U.; et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **2017**, *541*, 353–358, doi:10.1038/nature21031.

7. Rinke, C.; Schwientek, P.; Sczyrba, A.; Ivanova, N.N.; Anderson, I.J.; Cheng, J.F.; Darling, A.; Malfatti, S.; Swan, B.K.; Gies, E.A.; et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **2013**, *499*, 431–437, doi:10.1038/nature12352.

8. Nimonkar, Y.S.; Kajale, S.; Dake, M.; Ranade, D.R.; Yadav, K.K.; Kumar, R.; Prakash, O. A culture-based and culture-independent approach to the study of landfill leachate bacterial and archaeal communities. *Anaerobe* **2022**, *77*, 102626, doi:10.1016/j.anaerobe.2022.102626.

9. Edbeib, M.F.; Wahab, R.A.; Huyop, F. Halophiles: biology, adaptation, and their role in decontamination of hypersaline environments. *World J. Microbiol. Biotechnol.* **2016**, *32*, 1–23, doi:10.1007/s11274-016-2081-9.

10. Amoozegar, M.A.; Safarpour, A.; Noghabi, K.A.; Bakhtiary, T.; Ventosa, A. Halophiles and their vast potential in biofuel production. *Front. Microbiol.* **2019**, *10*, 1–17, doi:10.3389/fmicb.2019.01895.

11. Yin, J.; Chen, J.C.; Wu, Q.; Chen, G.Q. Halophiles, coming stars for industrial biotechnology. *Biotechnol. Adv.* **2015**, *33*, 1433–1442, doi:10.1016/j.biotechadv.2014.10.008.

12. Dumorné, K.; Córdova, D.C.; Astorga-Eló, M.; Renganathan, P. Extremozymes: A potential source for industrial applications. *J. Microbiol. Biotechnol.* **2017**, *27*, 649–659, doi:10.4014/jmb.1611.11006.

13. Corral, P.; Amoozegar, M.A.; Ventosa, A. Halophiles and their biomolecules: Recent advances and future applications in biomedicine. *Mar. Drugs* **2020**, *18*, doi:10.3390/md18010033.

14. Sun, S.; Chen, F.; Xu, Y.; Liu, J.; Chen, S. Halorubrum amylolyticum sp. nov., a novel halophilic archaeon isolated from a salt mine. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **2019**, *112*, 1849–1861, doi:10.1007/s10482-019-01313-4.

15. Albuquerque, L.; Kowalewicz-Kulbat, M.; Drzewiecka, D.; Staczek, P.; d'Auria, G.; Rosselló-Móra, R.; da Costa, M.S. Halorhabdus rudnickae sp. nov., a halophilic archaeon isolated from a salt mine borehole in Poland. *Syst. Appl. Microbiol.* **2016**, *39*, 100–105, doi:10.1016/j.syapm.2015.12.004.

16. Chen, S.; Xu, Y.; Ke, L.X. Halorubrum trueperi sp. nov., a halophilic archaeon isolated from a salt mine. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 1564–1570, doi:10.1099/ijsem.0.001762.

17. UNESCO World Heritage Centr Wieliczka and Bochnia Royal Salt Mines Available online: https://whc.unesco.org/en/list/32/.

18. Puławska, A.; Manecki, M.; Flasza, M.; Styszko, K. Origin, distribution, and perspective health

benefits of particulate matter in the air of underground salt mine: a case study from Bochnia, Poland. *Environ. Geochem. Health* **2021**, *43*, 3533–3556, doi:10.1007/s10653-021-00832-2.

19. Andrews, S. FastQC: a quality control tool for high throughput sequence data Available online: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

20. Krueger, F. Trim Galore Available online: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

21. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477, doi:10.1089/cmb.2012.0021.

22. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015**, *25*, 1043–1055, doi:10.1101/gr.186072.114.

23. Mikheenko, A.; Prjibelski, A.; Saveliev, V.; Antipov, D.; Gurevich, A. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* **2018**, *34*, i142–i150, doi:10.1093/bioinformatics/bty266.

24. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069, doi:10.1093/bioinformatics/btu153.

25. Huerta-Cepas, J.; Szklarczyk, D.; Heller, D.; Hernández-Plaza, A.; Forslund, S.K.; Cook, H.; Mende, D.R.; Letunic, I.; Rattei, T.; Jensen, L.J.; et al. EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **2019**, *47*, D309–D314, doi:10.1093/nar/gky1085.

26. Blin, K.; Shaw, S.; Kloosterman, A.M.; Charlop-Powers, Z.; Van Wezel, G.P.; Medema, M.H.; Weber, T. AntiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* **2021**, *49*, W29–W35, doi:10.1093/nar/gkab335.

27. Chaumeil, P.A.; Mussig, A.J.; Hugenholtz, P.; Parks, D.H. GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* **2020**, *36*, 1925–1927, doi:10.1093/bioinformatics/btz848.

28. Meier-Kolthoff, J.P.; Göker, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* **2019**, *10*, doi:10.1038/s41467-019-10210-3.

29. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296, doi:10.1093/nar/gkab301.

30. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549, doi:10.1093/molbev/msy096.

Re: Submission 82493k1 - manuscript word files for copyediting and fee

Journal of Genomics <publish@jgenomics.com>

Śr, 22.02.2023 06:38

Do: Jakub Lach <jakub.lach@biol.uni.lodz.pl>

DW: publish@jgenomics.com <publish@jgenomics.com>;j.roch.lach@gmail.com <j.roch.lach@gmail.com>

Dear Mr. Jakub Lach,

The revision of your manuscript 82493k1 is acceptable in Journal of Genomics and we are currently scheduling it for publication. Please check the manuscript carefully again and then submit a last, final version (in the original Microsoft Word file) to us for copyediting and publication. Since changes after copyediting and generation of an xml version are very difficult, it is important that any changes you would like to make are included before submitting the final files. Additional changes are generally not permitted after copyediting.

Please submit the final manuscript Word file after signing in https://www.jgenomics.com/ms/status

(Manuscript ID: 82493k1, Login code: RAADP)

References: The current references are not in our journal's required format. Please follow the format, e.g.
1. Eknoyan G, Beck GJ, Cheung AK, et al. Effect of dialysis dose and membrane flux in maintenance hemodialysis. N Engl J Med. 2002; 347: 2010-9.

See the format at https://www.jgenomics.com/ms/author

Figures:
* Please make sure images are of the highest resolution, and any letters inside are clear and legible. Images should have resolution >600 dpi and line art >1200dpi.
* If any figures are adapted or taken from other existing publications, please obtain permissions from the publishers or copyright owners as soon as possible.

Publication fee: the invoice is available in the same manuscript login page after signing in. The publication fee is required in our journal to cover the cost of publication, and should be paid before copyediting and publication. Please proceed to make the payment. Thank you for your support.

Kind regards,
Publishing Team
Journal of Genomics
https://www.jgenomics.com

Manuscript information
Title: Draft genomes of halophilic Archaea strains isolated from brines of the Carpathian Foreland, Poland
Author(s): Jakub Lach, Dominik Strapagiel, Agnieszka Matera-Witkiewicz, Paweł Stączek

European Journal of Pharmaceutics and Biopharmaceutics

Novel AMPs from saline environments - promising glimmer for inhibition of multidrug resistant

--Manuscript Draft--

Manuscript Number:	EJPB-D-23-00244
Article Type:	Research Paper
Keywords:	antimicrobial peptides; halophiles; metagenomics; molecular docking; sequencing
Corresponding Author:	Jakub Lach University of Lodz Lodz, POLAND
First Author:	Jakub Lach
Order of Authors:	Jakub Lach
	Magdalena Krupińska
	Aleksandra Mikołajczyk
	Dominik Strapagiel
	Paweł Stączek
	Agnieszka Matera-Witkiewicz
Abstract:	Microorganisms inhabiting saline environments have been known for many years as producers of many valuable bioproducts. These substances also include antimicrobial peptides (AMPs), the most recognizable of which are halocins produced by halophilic Archaea. As agents with a different mode of action than most conventionally used antibiotics, AMPs are currently the main subject of various research. The aim of this study was to investigate the antimicrobial activity against multi drug resistant human pathogens of three peptides, which have been synthetized based on sequences identified in metagenomes from saline environments. The investigations have been performed against Enterococcus faecalis, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Escherichia coli and Candida albicans. Subsequently, the cytotoxicity and haemolytic properties of the tested peptides were verified. An in silico analysis of the interaction of the tested peptides with molecular targets for reference antibiotics was also carried out in order to verify whether they can act on a similar way. The P1 peptide manifested the growth inhibition of E. faecalis at a MIC50 of 32 µg/ml and the P3 peptide at a MIC50 of 32 µg/ml was shown to inhibit the growth of both E. faecalis and S. aureus. Furthermore, the P1 and P3 peptides have been shown to have no cytotoxic or haemolytic activity against human cells.
Suggested Reviewers:	Zülal Kesmen Erciyes University zkesmen@erciyes.edu.tr
	Chunye Zhang University of Missouri zhangcherryuniversity@gmail.com
	Ming Yang University of Missouri yangmin@health.missouri.edu
	Yihua Chen Institute of Microbiology Chinese Academy of Sciences chenyihua@im.ac.cn
	Yipeng Wang Soochow University yipengwang@suda.edu.cn
Opposed Reviewers:	

Lodz, 16th March 2023

Editorial Office
European Journal of Pharmaceutics and Biopharmaceutics

Dear Editor,

We wish to submit an original research paper entitled “Novel AMPs from saline environments - promising glimmer for inhibition of multidrug resistant E. faecalis and S. aureus infections?" prepared by Jakub Lach, Magdalena Krupińska, Aleksandra Mikołajczyk, Dominik Strapagiel, Paweł Stączek and Agnieszka Matera-Witkiewicz for consideration by the European Journal of Pharmaceutics and Biopharmaceutics.

We hereby state that this work has not been published or accepted for publication, and is not under consideration for publication in another journal or book. We have not submitted our manuscript to a preprint server before submitting it to the European Journal of Pharmaceutics and Biopharmaceutics. The manuscript in its submitted form has been read and approved by all authors. The authors declare no conflict of interest.

In this paper, we investigate the antimicrobial activity against multi drug resistant human pathogens of three peptides, which have been synthetized based on sequences identified in metagenomes from saline environments. The P1 peptide manifested the growth inhibition of *E. faecalis* at a MIC50 of 32 μg/ml and the P3 peptide at a MIC50 of 32 μg/ml was shown to inhibit the growth of both *E. faecalis* and *S. aureus*. Furthermore, the P1 and P3 peptides have been shown to have no cytotoxic or haemolytic activity against human cells.

We hope that the manuscript meets the high standards of your journal. We are looking forward to receiving a favorable response from you regarding the acceptance of our manuscript.

Yours sincerely
Jakub Lach

Jakub Lach
Biobank Lab / Department of Molecular Microbiology
Faculty of Biology and Environmental Protection,
University of Lodz
Pomorska 139, Lodz 90-235, Poland,
Tel.: +48 513 006 750
E-mail: Jakub.lach@biol.uni.lodz.pl

Novel AMPs from saline environments - promising glimmer for inhibition of multidrug resistant E. faecalis and S. aureus infections?

Jakub Lach[a,b,#], Magdalena Krupińska[c], Aleksandra Mikołajczyk[c], Dominik Strapagiel[b], Paweł Stączek[a], Agnieszka Matera-Witkiewicz[c]

a - Department of Molecular Microbiology, Faculty of Biology and Environmental Protection, University of Lodz, Lodz, Poland

b - Biobank Lab, Department of Oncobiology and Epigenetics, Faculty of Biology and Environmental Protection, University of Lodz, Lodz, Poland

c - Screening of Biological Activity Assays and Collection of Biological Material Laboratory, Wroclaw Medical University Biobank, Faculty of Pharmacy, Wroclaw Medical University, Wroclaw, Poland

#Corresponding author: Jakub Lach, jakub.lach@biol.uni.lodz.pl

Keywords: antimicrobial peptides, halophiles, metagenomics, molecular docking, sequencing

Abstract

Microorganisms inhabiting saline environments have been known for many years as producers of many valuable bioproducts. These substances also include antimicrobial peptides (AMPs), the most recognizable of which are halocins produced by halophilic Archaea. As agents with a different mode of action than most conventionally used antibiotics, AMPs are currently the main subject of various research. The aim of this study was to investigate the antimicrobial activity against multi drug resistant human pathogens of three peptides, which have been synthetized based on sequences identified in metagenomes from saline environments.

The investigations have been performed against *Enterococcus faecalis*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Escherichia coli* and *Candida albicans*. Subsequently, the cytotoxicity and haemolytic properties of the tested peptides were verified. An *in silico* analysis of the interaction of the tested peptides with molecular targets for reference antibiotics was also carried out in order to verify whether they can act on a similar way.

The P1 peptide manifested the growth inhibition of *E. faecalis* at a MIC50 of 32 µg/ml and the P3 peptide at a MIC50 of 32 µg/ml was shown to inhibit the growth of both *E. faecalis* and *S. aureus*. Furthermore, the P1 and P3 peptides have been shown to have no cytotoxic or haemolytic activity against human cells.

Introduction

Microorganisms inhabiting high salinity environments are known as halophiles. Representatives of these class of extremophiles can be found in all three domains of life [1]. Halophiles are known as a source of active biomolecules for biotechnological and pharmaceutical applications and due to the fact of high variability and unique adaptation to extreme environments they are still interesting subjects to search of novel bioproducts useful in biomedicine and industry. In the era of multidrug-resistant bacteria, increasing morbidity rate of cancer, and extreme environmental pollution, research on the development of new active compounds is urgently obligated [2,3].

Antibiotic resistance is one of the most significant public health problems in recent years. Due to the continuous overuse of antibiotics, both in human medicine and in other applications, the number of cases of infections with pathogens resistant to standard antibiotics is increasing year by year [4]. In order to combat this threat, the World Health Organization introduced the "Global Action Plan on Antimicrobial Resistance" in 2015, where the guidelines for global actions to prevent the further development of antibiotic resistance was presented. One of the important activities under this plan is an exploration of the sources and mechanisms of antibiotic resistance are created in the environment in accordance with the One Health approach [5]. In addition, the key activity is also refer to seeking of new active antimicrobial compounds, in particular antibiotics with new mechanisms of action. However, antimicrobial peptides (AMPs) can be included in this group of potential biomolecules to cope with microbials resistance, respectively. Comparable activities have been undertaken by the European Union since 2003 through Public Health Programs. Here, supporting actions for research activities and programs in order to search for new antimicrobial drugs, innovative products and alternative methods of therapy, as well as the prevention of infections and infectious diseases have been underlined. Moreover, the European Union Joint Action on Antimicrobial Resistance and Healthcare Associated Infections has been established to support European Union members in developing and implementing One Health solutions to combat the growing threat of antibiotic resistance.

Antimicrobial peptides (AMPs), a diverse group of bioactive small proteins, are a part of the body's first line of defence for pathogen inactivation in eucaryotic organisms. They work by disrupting bacterial cell membranes, modulating the immune response, and regulating inflammation [6]. AMPs serving as important tools of microorganism competition in complex microbial communities, can inhibit growth of other species of microorganisms as well [7,8]. Valuable and important for biotechnology and medicine AMPs were identified in a number of eukaryotic organisms and numerous of bacteria and archaea. But there is also a plenty of unconventional sources of AMPs, including unculturable soil and marine bacteria, extremophilic microorganisms and methods available to produce vast libraries of derivatives [9–11]. Currently only few AMPs are approved to clinical usage as an alternative to antibiotics in terms of their antimicrobial potency. This behaviour is related to the challenges toward clinical application of AMPs which include cytotoxic effects, costs production and obstacles related to peptide bioavailability and efficacy [12]. However, in the age of antibiotic resistance, AMPs have the potential to be a valuable tool to combat multidrug resistant bacteria and reduce usage of antibiotics in clinical application due to the synergistic activity of AMPs with antibiotics [13,14]. AMPs can be used not only in microbial infections treating, but also can be applied in agriculture to control plant diseases, as therapeutics in aquaculture or food additives

for livestock [15–17]. Finally, AMPs usage as food preservatives is also proved. Nisin is an AMP representative as antimicrobial active packaging of food compound [18,19].

Microorganisms inhabiting saline and hypersaline environments are also producers of wide range of AMPs. Halophilic archaea produce group of AMPs known as halocins [10]. First discovered halocin-Halocin H4 was described by Francisco Rodriguez-Valera in 1982 [20]. Only for Halocin H4, C8 and S8 the amino acid and nucleotide sequences of their genes are known. Presented halocins are synthesized in the form of precursor proteins which undergo posttranslational modification to release the mature halocins as secretory proteins [10]. Also, halophilic bacteria are AMP producers, but they are less explored than archaea in that case [2,21]. Despite the ecological and environmental role of several halocins, their action against human pathogens has been less studied. The clinical significance of AMPs produced by halophilic microorganisms is minor reported and their antimicrobial activity against the most important human pathogens still remains an area of preliminary research [2]. However, halocins stability in hypersaline environments made them useful in industrial applications especially in preservations of salted food [10,21].

The aim of this study was to investigate the antibacterial activity against multidrug resistant human pathogens: *Enterococcus faecalis*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii, Pseudomonas aeruginosa, Escherichia coli* and *Candida albicans* of three peptides sequences identified on the bases of the metagenomes of the Bochnia Salt Mine and brine graduation towers. Activity of the selected peptides in combination with the reference antibiotics was also checked. To verify the cytotoxicity and potential haemolytic effect, also specific assays were performed. Potential mechanism of action for the peptides with confirmed antibacterial activity was studied with *in silico* analysis.

Materials and Methods

2.1 Sample collection and sequencing

Four brine samples were collected from brine wells situated in the Bochnia Salt Mine located in southern part of Carpathian Foreland in Poland near Kraków city (49°58′09″N 20°25′03″E), three samples of blackthorn were collected from three brine graduation towers located in Lodz city in the Podolski Park (51°74′14″N 19°49′17″E), the Botanik Residential (51°75′02″N 19°40′30″E), the Mikolaj Rej Residential (51°78′83″N 19°41′73″E) and one sample of brine was collected from the Tadeusz brine source in Zablocie (49°90′77″N 18°77′06″E). All samples were collected in 2019. To prepare the brines for DNA isolation, 2 ml of saline was transferred into a new sterile Eppendorf tube and centrifuged for 10 minutes at 14 000×G. After centrifugation, 1.8 ml of brine were withdrawn, and 1.8 ml of new brine was added. The centrifugation was then repeated. This procedure was repeated 5 times for each sample. In the case of blackthorn about 1 cm fragments of twigs were rinsed and shaken in sterile nuclease-free water and then the water was used for DNA isolation. DNA was extracted from the samples using PowerSoil DNA Isolation Kit (Mo Bio Laboratories Inc., Carlsbad, CA, USA). Samples were eluted with nuclease-free water in a volume of 50 µl. DNA concentration was determined using the Qubit high-sensitivity (HS) assay kit (ThermoFisher, Waltham, USA).

Shotgun sequencing libraries were prepared using the Vazyme TruePrep DNA Library Prep Kit V2 for Illumina (Nanjing, China). Libraries were sequenced on Illumina NextSeq 500 with 2x150bp paired end reads.

2.2 In-silico analysis of selected organisms

For shotgun sequencing data *de novo* assembly was performed using MEGAHIT v.1.2.9 with the "--min-contig-len 1000" parameter [22]. The quality of the assemblies was checked using metaQUAST [23].

Identification of AMP sequences were performed using marcel v. 0.3.1 on default parameters [24]. AMP sequences with non-haemolytic properties were verified and classified into functional types by using iAMP-2L platform [25].

In assembled contigs potentially AMPs sequences were identified. Sequences marked as a potentially haemolytic were excluded from further analysis. The probability of non-haemolytic AMPs sequence classification was ranged from 50.05% to 76.2%. Furthermore, sequences where the AMP probability exceeded 70% were selected for synthesis and *in vitro* evaluation of their antimicrobial and cytotoxic properties. Peptides labelled as P1 and P3 were supposed to affect Bacteria and P2 Fungi. Selected peptides are presented in Table 1.

Table 1. Peptides selected for antimicrobial activity assay

ID	Source	Sequence	AMP probability	Target organism
P1	Brine graduation tower - Botanik Residential	LAAIDALARACLKVKPDTTKIQNTARYPSVTSGT	76.2%	Bacteria
P2	Brine well – The Bochnia Salt Mine	AAALCVRAAVFKRGESNGYDPKPGDLRVGKVKRAERRVEAC	72.3%	Fungi
P3	Brine graduation tower – Podolski Park	NHFKNIGRVNYLGQPMLQRVSHCFGYPRPVIGSKSKPA	71.3%	Bacteria

The possible mode of action and target organisms were assumed using CAMP database [26]. CAMPR4 contains information on the AMP sequence, protein definition, accession numbers, activity, source organism, target organisms, protein family descriptions, N and C terminal modifications and links to databases like UniProt, PubMed and other antimicrobial peptide databases.
Peptide parameters such total net charge, molecular weight, hydrophobicity were calculated using APD3 database [27].

2.3 Antimicrobial peptides synthesis and preparation

Three peptides with the highest AMPs probability were synthesised by Pepmic Co., Ltd (Suzhou, China). At least 98% of purity of each peptide was required. Peptides were synthesised and ligated at pH~7 with thiophenol derivative. Quality analysis of peptides were done using MS spectrum. The QC certificate was delivered by the provider. Additional QC checking was done using ESI-MS mass spectrometer by Bruker Daltonik (model Compact, Billerica, Massachusetts, USA).

2.4 Antimicrobial activity assay

Seven reference strains from ATCC collection (*A. baumanii* 19606, *K. pneumoni*ae 700603, *S. aureus* 43300, *E.coli* 25922, *E. faecalis* 29212, *P. aeruginosa* 27853 and *C. albicans* 10231) were used for antimicrobial activity assay. The antimicrobial activity was performed according to the standard protocol using microdilution method with spectrophotometric measurement (λ = 580 nm at starting point and after 24 h) [28] convergent to the ISO standard 20776-1:2019 [29], ISO standard 16256:2012 [30] and modified Richard's method [31–33].

Stock peptide solutions were prepared in water. Serial dilutions were made on 96-well microplates in the range between 0.5 µg/ml and 256 µg/ml. Tryptone Soy Agar (TSA) plates were inoculated with microbial strains from performed stocks. After 24 h/37ºC incubation (for bacteria) or 24 h/25ºC (for fungus) a proper density of bacterial and fungal suspension was prepared using a densitometer (final inoculum ($5x10^5$ CFU/ml) was prepared in Tryptic Soy Broth (TSB). A positive (TSB+ strain) and negative control (TSB) were also included in the test. Spectrophotometric solubility control of each peptide was also performed. Microplates were incubated at 37°C or 25°C for 24 hours on the shaker (500 rpm). After this, the spectrophotometric measurement was performed at 580 nm and then 50 µL aliquots of 1% (m/v) 2,3,5-triphenyltetrazolium chloride (TTC) solution were added into each well. TTC is a chemical indicator which is converted into red formazan crystals in living microbial cells. Possible killing effect can be observed as the lowest concentration determined by visual analysis after 24 h incubation with TTC (did not change the color to pink) Thus simultaneous usage of microdilution method and TTC examination let us to determine potential MIC (Minimal Inhibitory Concentration) and MBC (Minimal Bactericidal Concentration) or MFC (Minimal Fungicidal Concentration) values.

For each strain, the validation process was performed using the following antibacterial/antifungal agents: levofloxacin, gentamicin, amphotericin B, according to the EUCAST examination. Minimal Inhibitory Concentration (MIC) was determined for each strain referred to antibacterial/antifungal agents. Moreover, for peptides investigation the 50% Minimal Inhibitory Concentration (MIC50) was determined. The original aim of the study was to determine the MIC for each of the tested compounds, including peptides, but due to the inability to determine the MIC of peptides in the tested concentration range, it was decided to determine the MIC50. As MIC50 the lowest concentration of an antimicrobial agent that inhibits the measured microbial growth to 50 % as referred to positive control was obtained.

After MIC50 determination for P1 (*E. faecalis*) and P3 (*E. faecalis, S. aureus*) synergy checkerboard assay was performed. Columns 1 to 11 contain 2-fold serial dilutions of P1 or P3 which starts from earlier obtained MIC50, and rows A to G contain 2-fold serial dilutions of levofloxacin. Levofloxacin was selected for this study as the reference antibiotic for tested strains. Column 12 contains a serial dilution of levofloxacin alone, while row H contains a serial dilution of Compound A alone. A detail scheme of the plate preparation for checkerboard assay is presented at Figure 1.

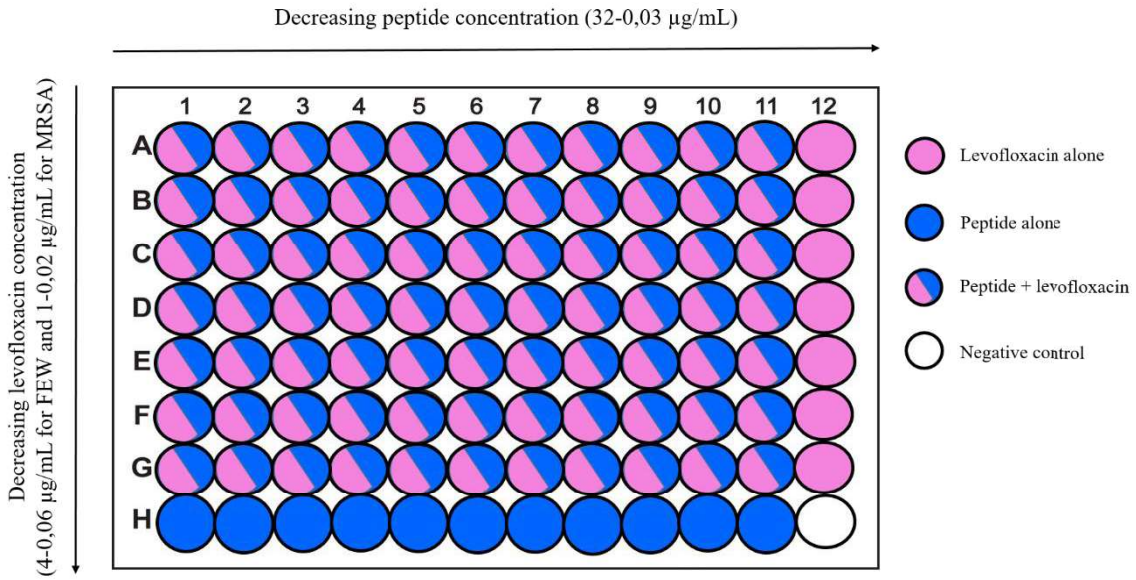


Figure 1. Synergy checkerboard assay plate scheme.

2.5 Neutral Red cytotoxicity of selected AMPs

For each peptide, where the antimicrobial activity was determined against *E.faecalis* and *MRSA* strains, a Neutral Red (NR) cytotoxicity assay was performed using human primary renal proximal tubule epithelial cells (RPTEC) from ECACC collection. The experiment was performed according to ISO:10993 guidelines (Biological evaluation of medical devices; Part 5: Tests for in vitro cytotoxicity; Part 12: Biological evaluation of medical devices, sample preparation and reference materials: ISO 10993-5:2009 and ISO/IEC 17025:2005). A standard protocol for the NR assay was used from Nature Protocol [34]. MEMα supplemented with 10% FBS, 2 mM L-glutamine and suitable amount of antibiotics (amphotericin B, gentamycin) was used for the experiment. Stock peptide solutions were prepared in water and then 100x diluted in the medium. P1 and P3 tested concentration was in range from 64-8µg/ml. After adding proper mixtures of testing compounds and cells ($1x10^5$ cells/ml) into each well, plates were incubated for 24, 48 and 72 h in 5% CO_2 at 37ºC. Next, medium was removed and 100 µL of NR solution (40 µg/ml) was added to each well followed by incubation for 2 h at 37ºC. After removing the dye, wells were rinsed with PBS and left to dry. Then, NR detain solution (1% of glacial acetic acid, 50% of 96% ethanol and 49% of deionized water; v/v) was added to each well. The plates were shaken (30 min, 500 rpm) until NR was extracted from the cells and formed a homogenous solution. The absorbance was measured using microplate reader at λ =540 nm. As a negative control untreated cells were considered as 100% of potential cellular growth. Furthermore, cells incubated with 1 µM staurosporine were used as a positive control.

2.6. Haemolytic activity of AMPs

Haemolytic activity was determined by incubating a 5% (v/v) suspension of human erythrocytes with selected peptides. Red blood cells were rinsed three times in PBS, by centrifugation for 15 min at 1500×g. Then were incubated at 37°C for 3h with saponin (positive control), with additional PBS (negative control) and with specific peptides in MIC50. Afterwards, the samples were centrifuged at 1500×g for 15 min, the supernatant was separated from the pellet and absorbance of supernatant was measured at λ =540 nm. Pure PBS was used as blank. The relative haemolysis percentage (H) was computed using the equation:

$$H = \frac{absorbance\ of\ sample - absorbance\ of\ blank}{absorbance\ of\ positive\ control - absorbance\ of\ blank} \times 100$$

2.7 Molecular docking

Structures of selected peptide sequences were predicted by Psipred database [35] and modelled using RoseTTAFold provided by computing resources by the Baker lab. Peptides P1 and P3 were found to have the antimicrobial activity against 3 pathogens out of 7 used in laboratory tests (*E.faecalis* 29212*, S.aureus* 43300 and *E.coli* 25922). Molecular docking was performed to model the interaction between molecules and a receptor at the atomic level, in order to characterize the behavior of molecules in the binding site of a target proteins as well as to elucidate fundamental biochemical processes.

Molecular docking was performed using AutodockVina [36] in order to dock levofloxacin and gentamicin, and ClusPro [37] server for docking of peptides P1, P2, P3. Input sequences were prepared by AutoDockTools software.

The structures used in docking were optimized by the energy minimization in MMFF94 force field [38]. Water molecules were removed from receptors and polar hydrogens were added, as well as missing atoms were repaired. ADT 1.5.6 software [39] is used to investigate the activity in terms of binding affinity (Kcal/mol). The docking outcomes e.g. bonds between ligand and receptor, binding affinity score for best-docked conformation are compared for reference antibiotics and analysed peptides and presented in *Results* section.

Results

3.1 AMPs sequences identification

The physico-chemical properties such as net charge, molecular weight, hydrophobicity of the peptides selected for *in vitro* analysis have been determined in order to evaluate their essential properties. All peptides are characterised by positive net charge and contain significant number of hydrophobic amino acids which may suggest their amphipathic properties. Described properties has been presented in Table 2.

Table 2. Physicochemical properties of peptides selected for antimicrobial activity testing

Sequence	**Number of residues / Total net charge**	**Hydrophobic aminoacids**	**Molecular weight [g]**
LAAIDALARACLKVKPDTTKIQNTARYPSVTSGT	34 (Total net charge +3)	13/34 (38%)	3575.142
AAALCVRAAVFKRGESNGYDPKPGDLRVGKVKRAERRVEAC	41 (Total net charge +5)	17/41 (41%)	4445.154
NHFKNIGRVNYLGQPMLQRVSHCFGYPRPVIGSKSKPA	38 (Total net charge +6.5)	12/38 (32%)	4298.023

In assembled contigs from all eight samples, 271 potentially AMPs sequences were identified. 191 sequences were marked as possible haemolytic activity, which resulted in their exclusion from further analysis. From 80 remaining sequences, peptides with the probability of being AMP above 70% were selected. The results predict only 3 sequences, thus they have been chosen for synthesis and *in vitro* evaluation of their antimicrobial properties.
In Table 3 the pathogens and the reference antibiotic which are using in treatment of bacterial infection were presented.

Table 3. Pathogens used in molecular docking with reference antibiotics and their molecular targets

Strain	**Reference antibiotic**	**Receptor**
E.faecalis 29212	Levofloxacin	Topoisomerase IV
S.aureus 43300	Levofloxacin	DNA gyrase
E.coli 25922	Gentamicin	16S rRNA

The receptors listed in Table 2 were also used in docking for peptides P1, P2, P3. *E. faecalis* Topoisomerase IV, *S. aureus* DNA gyrase and *E.coli* rRNA 16S, were selected as receptors for molecular docking due to the levofloxacin and gentamicin mode of action. Gentamicin is supposed to incorporate and destabilise the bacteria outer membrane and reach the 16S rRNA where bind in helix 44 near the A site of the 30S ribosomal subunit [40]. Levofloxacin inhibits two bacterial enzymes DNA gyrase and Topoisomerase IV. Both targets are type II topoisomerases but have unique functions within the bacterial cell [41].

3.2 Antimicrobial activity of selected peptides.

Antimicrobial activity assay was performed using multidrug resistant pathogens. First, the MICs for antibiotics used as reference agents were verified. Obtained MIC values were for *S. aureus* 43300: levofloxacin 0.25 µg/ml, *E.coli* 25922: gentamicin 4 µg/ml, *C. albicans* 10231: amphotericin B 1 µg/ml, *P. aeruginosa* 27853: levofloxacin 1 µg/ml, *E. faecalis* 29212: levofloxacin 1 µg/ml, *A. baumannii* 19606: levofloxacin 0,5 µg/ml and *K. pneumoniae* 700603: gentamicin 4 µg/ml.

The results for P1, P2 and P3 did not present a full growth inhibition for any of the strains. Nevertheless, MIC50 has been achieved for P1 and P3 against selected MDR strains. The assay showed MIC50 at 32 µg/ml for P1 and P3 against *E. faecalis*. Moreover, P3 inhibited the growth of MRSA with MIC50 at 32 µg/ml and *E.coli* with MIC50 at 256 µg/ml. P2 did not inhibit the growth of any of the tested microorganisms (Table 4).

Table 4. MIC50 (µg/ml) of peptides against selected microbials

Strain	P1	P2	P3
A. baumannii 19606	>256	>256	>256
C. albicans 10231	>256	>256	>256
E. coli 25922	>256	>256	**256**
E. faecalis 29212	**32**	>256	**32**
K. pneumoniae 700603	>256	>256	>256
S. aureus 43300 (MRSA)	>256	>256	**32**
P. aeruginosa 27853	>256	>256	>256

3.3 Antimicrobial activity of peptides in combination with levofloxacin against *E. faecalis* and *MRSA*

After MIC50 detection, peptides P1 and P3 were tested in combination with levofloxacin against *E. faecalis* and *MRSA* using microdilution method and checkerboard assay. The checkerboard assay was performed for potential synergy effect identification. Levofloxacin was selected for this study as the reference antibiotic for all strains, where MIC50 was defined. For microdilution method, MIC value against *E. faecalis* for P1 and P3 was obtained in combination of 0.03 µg/ml of levofloxacin and 32 µg/ml of each peptide, respectively. In comparison, levofloxacin alone showed an inhibitory effect (MIC) at concentration 1 µg/ml. MBC for levofloxacin alone and for combination with peptides was the same and it was 1 µg/ml.

In the case of *S. aureus* 43300, the MIC for the combination of levofloxacin and peptide was determined at 0.03 µg/ml of levofloxacin and 32 µg/ml of peptide P3. Single use of levofloxacin presented MIC at 0.25 µg/ml. Furthermore, 0.25 µg/ml of levofloxacin in combination with 32 µg/ml of peptide P3 was essential for MBC determination.

For checkerboard assay results FIC index could not be calculated, because MIC value was not determined in a single treatment of P1 and P3. Nevertheless, the checkerboard test confirmed the experimental values of microdilution method. Here, also MIC values for *E. faecalis* and

MRSA were obtained (levofloxacin for *E. faecalis* 4/2/1 µg/ml and 1/0.5/0,25 µg/ml for *MRSA*). Moreover, for P1 and P3 in the concentration 32 µg/ml only MIC 50 was determined and confirmed. For all other concentrations combinations, no antimicrobial activity has been determined. It can suggest that no synergy effect can be obtained for analysed systems.

3.4 Cytotoxicity and haemolytic properties

For P1 and P3 peptides cytotoxic properties were also investigated using Neutral Red assay. The experiment was performed for P1, P1-levofloxacin (1 µg/ml), P3, and P3-levofloxacin (1 µg/ml) in 3-point time frame (24/48/72 hours). Viability calculation for all systems in each time estimation presents positive results. The viability range was: 100%-107% after 24 hours, 100%-115% after 48 hours and 99%-107% after 72 hours, respectively. Figure 2 presents the results for all systems after 72 hours of incubation.

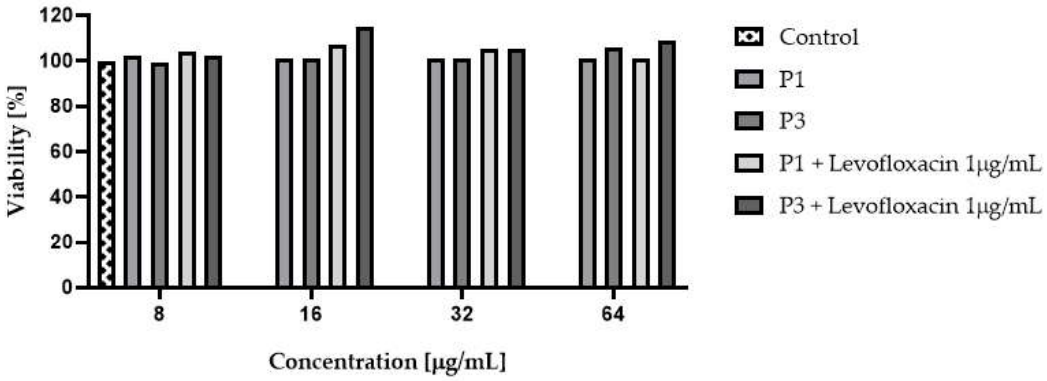


Figure 2. RPTEC viability after 72 hours incubation with P1/P3 peptides alone and in combination with levofloxacin

In the second assay, where potential hemolytic activity was checked, the analyzed peptides did not show cytotoxicity towards human erythrocytes. The level of hemolysis for the peptide-treated red blood cells was similar to the level of hemolysis in the untreated samples (Table 5).

Table 5. Hemolytic activity of tested peptides

Treatment group	**Mean percentage hemolysis (± SEM)**
Untreated cells	7.53 ± 0.61
Saponin	100 ± 5.74
P1	6.81 ± 0.45
P2	7.65 ± 0.10
P3	7.29 ± 0.43

3.5 Molecular docking analysis.

3.5.1 Molecular docking results between E. *faecalis* 29212 Topoisomerase IV and levofloxacin as well as *E. faecalis* 29212 Topoisomerase IV and peptides P1 and P3

Docking studies were performed in order to estimate the binding interactions of antibiotics and compare the results with studied peptides P1, P2, P3. Docking studies revealed that levofloxacin showed better binding affinity (-7.7 kcal/mol) than peptide P1 (-7.0 kcal/mol) and P3 (-6.6 kcal/mol). In-depth analysis revealed that levofloxacin made hydrogen bond in its best ranked conformation with Asn650 of *E. faecalis* Topoisomerase IV (Figure 3), while peptide P1 made hydrogen bonds with Asn650 and Val596, Ile594, Leu540, Thr649, Glu647, His511, Glu807, Asn702, Ile589, Glu512, Tyr514, Lys534, Glu578, Glu575, Ile576, Arg572 (Figure 4). Peptide P3 interacted by forming hydrogen bonds with Asn650 and Lys613, Asp590, Gln645, Lys613, Glu805, Asn702, Glu808, Leu540, Tyr597, Val651, Glu815, Lys606, Leu653, Asp656 and Asp658 (Figure 5).

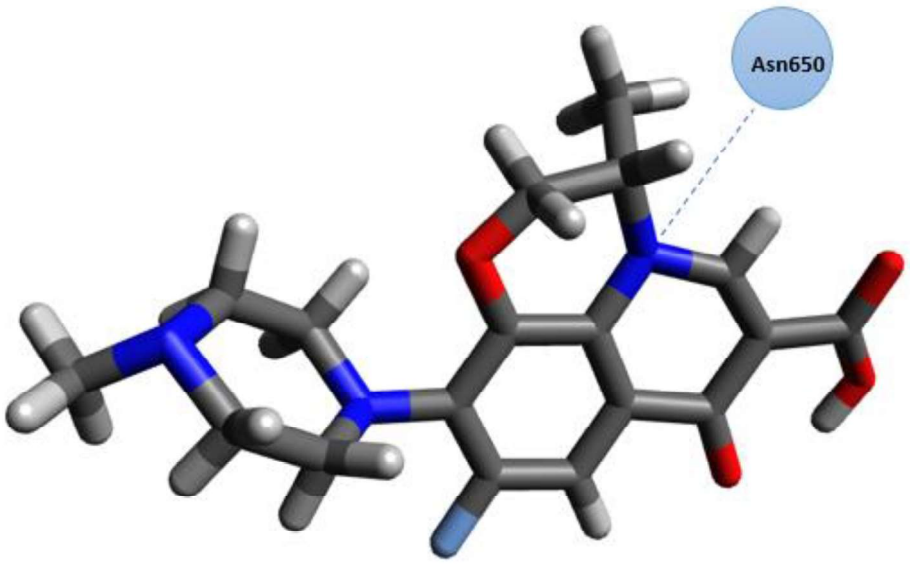


Figure 3. Topoisomerase IV residue involved in hydrogen bond formation with levofloxacin.

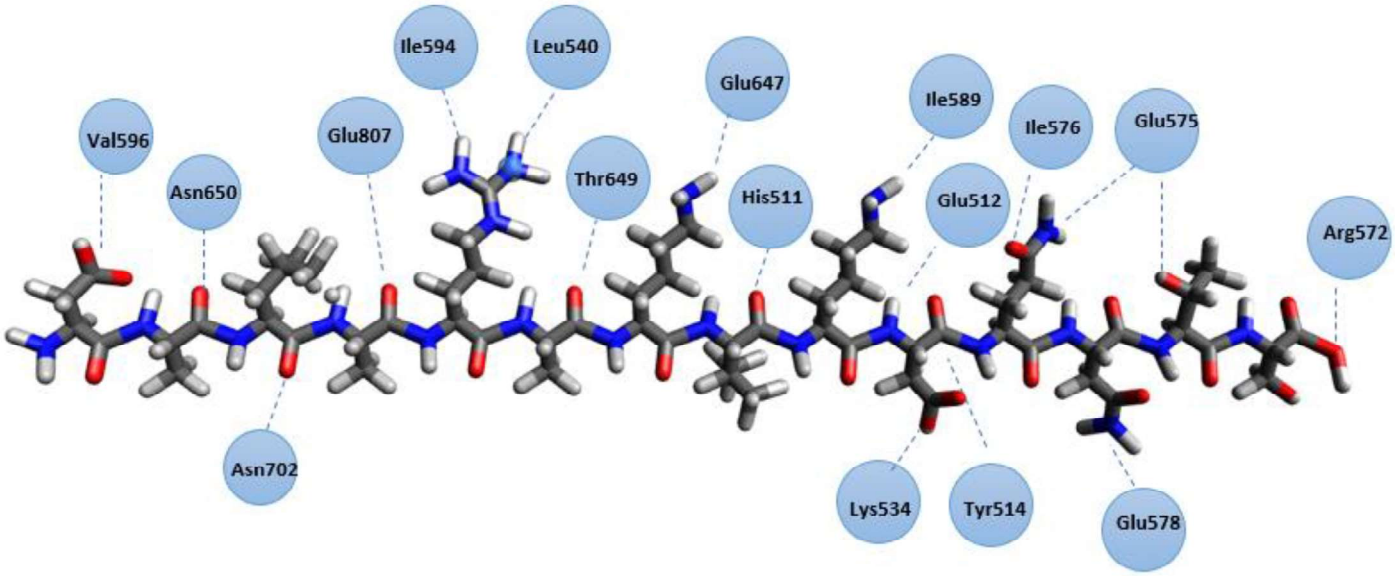


Figure 4. Topoisomerase IV residues involved in hydrogen bonds formation with peptide P1.

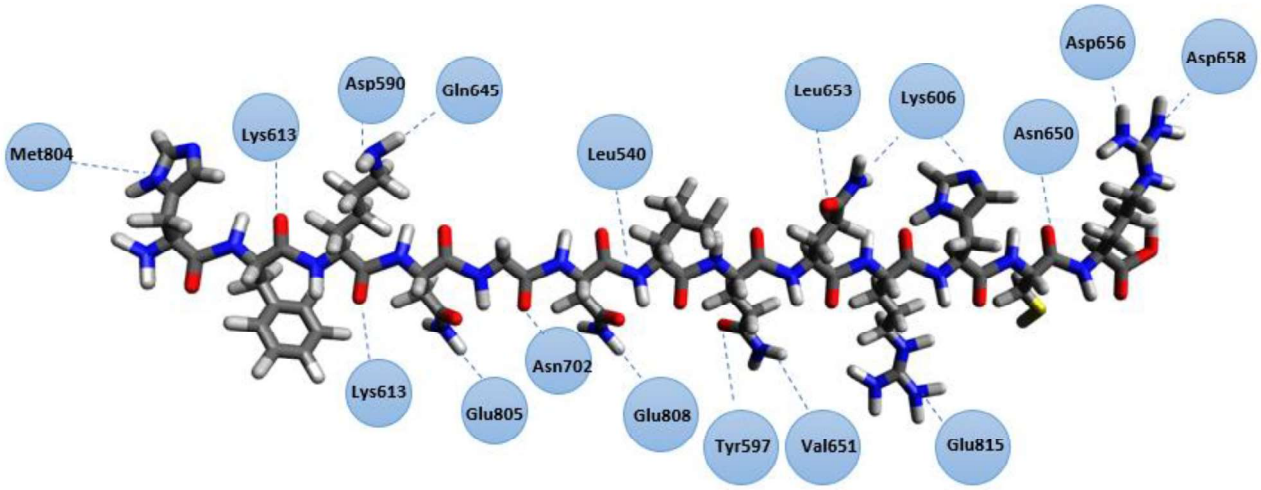


Figure 5. Topoisomerase IV residues involved in hydrogen bonds formation with peptide P3.

3.5.2 Molecular docking results between *S. aureus* 43300 DNA gyrase and levofloxacin as well as *S. aureus* 43300 DNA gyrase and peptide P3.

Levofloxacin was found to form hydrogen bond with Asn383 with *S. aureus* 43300 DNA gyrase (Figure 6). Binding affinity –10.2 kcal/mol of a best ranked pose. Peptide P3 made hydrogen bonds with Glu436, Arg433, Ser31, Asn535, Asp182, Lys59 and Arg61 (Figure 7). Binding affinity of the interaction was –8.8 kcal/mol.

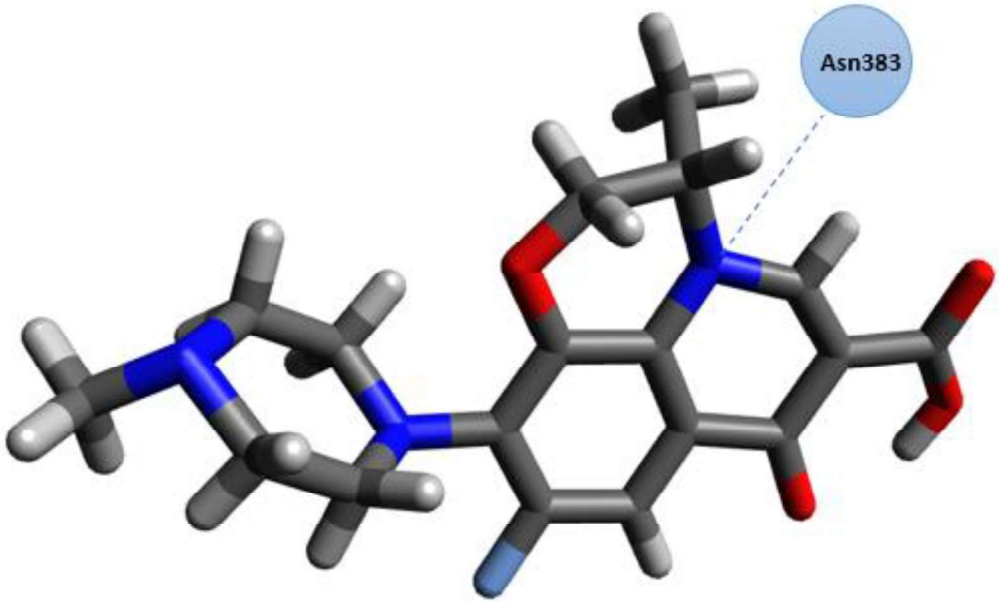


Figure 6. DNA gyrase residue involved in hydrogen bond formation with levofloxacin.

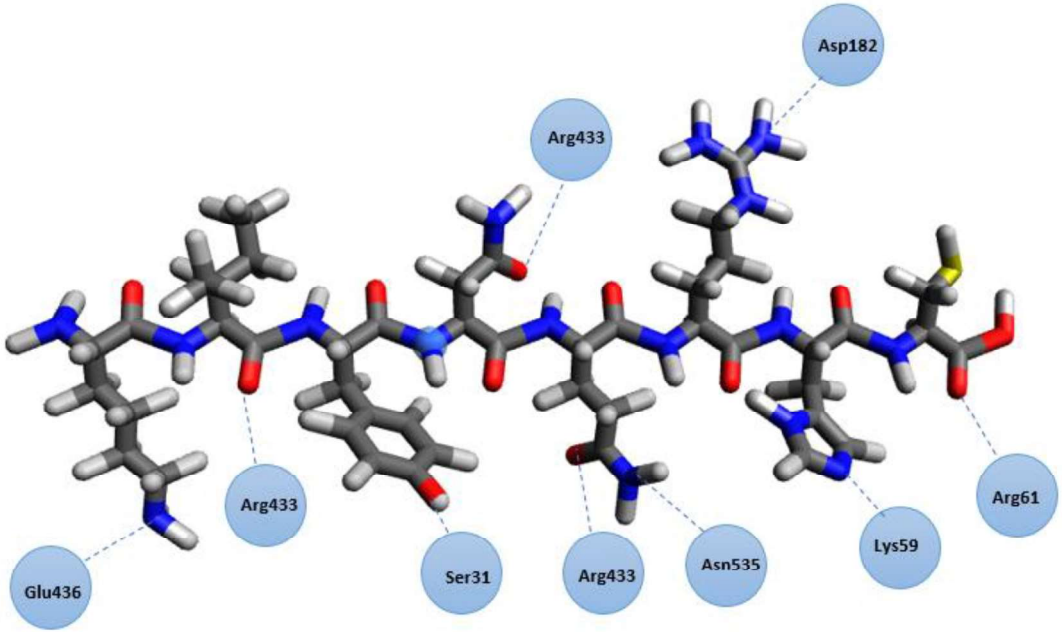


Figure 7. DNA gyrase residues involved in hydrogen bond formation with peptide P3.

3.5.3 Molecular docking results between *E. coli* 25922 16S rRNA and gentamicin as well as *E. coli* 25922 16S rRNA and peptide P3.

Gentamicin C1a binds in the A-site of 16S rRNA and form hydrogen bonds to four nucleotides G1405, U1495, A1408 and A1492. Binding affinity –7.8 kcal/mol of a best ranked pose. Peptide P3 made hydrogen bonds in the A-site rRNA (binding affinity –3.3 kcal/mol) with G1405, U1495, G1494, A1493, A1492, G1491 (Figure 8 and 9).

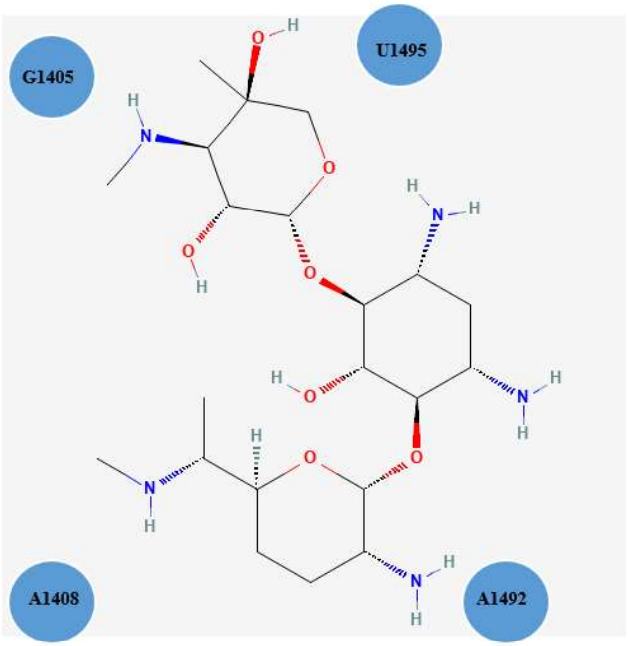


Figure 8. Gentamicin in A-site of 16S rRNA

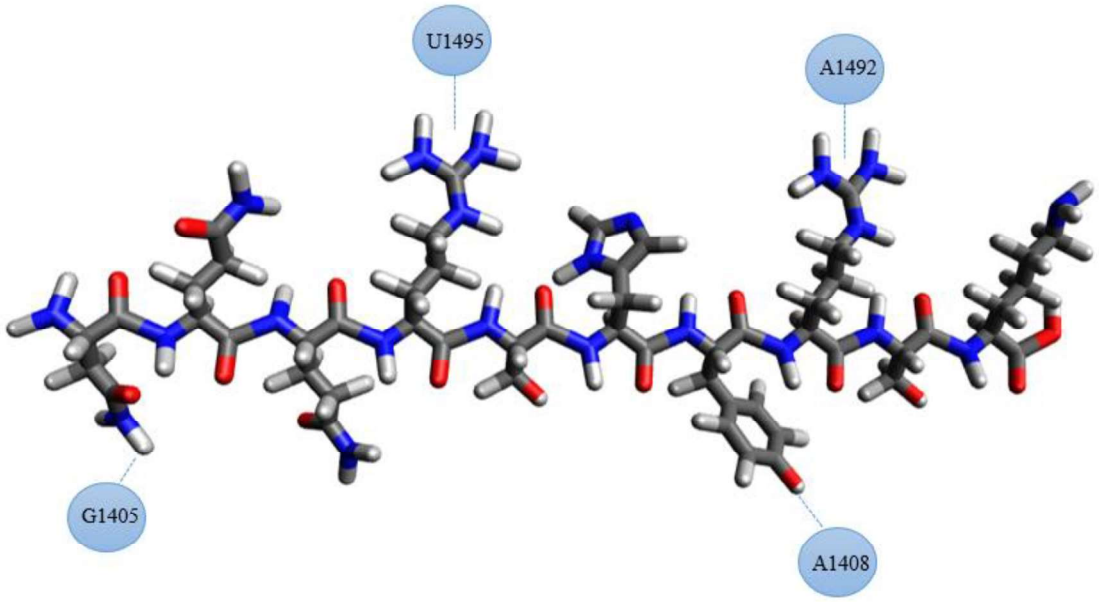


Figure 9. 16S rRNA nucleotides involved in the hydrogen bonds formation with peptide P3.

Discussion

Due to the challenging fact of increased antibiotic resistance, the crucial task is to seek for alternatives in pathogens defence. Multi-drug resistance is extremely hard to neglect and may be untreatable with conventional antibiotics [42]. Antimicrobial peptides - AMPs can be an alternative to overcome bacterial resistance. In our study, the sequences of 271 AMPs were identified in metagenomic data from saline environments and three of them with the highest probability of being AMP were selected for further analysis. Their antimicrobial properties, cytotoxic and potential haemolytic effects against eukaryotic cells and potential mechanism of action were characterized. Due to the characteristics of the environmental niche of microorganisms in the metagenome of which genes encoding the analysed AMPs were detected, it can be assumed that they may belong to halocins. Unlike to the majority of studies that assess AMPs produced by halophilic microorganisms, we decided to focus on assessing their activity against pathogens characterized by increasing antibiotic resistance [10]. Most of the known AMPs isolated from halophilic microorganisms were tested or showed antimicrobial activity only against closely related species. The production of AMPs active against closely related strains indicates their important role in competition between strains inhabiting the same niches [20,43,44]. However, when it comes to the application of AMPs produced by halophiles against pathogenic strains, only few examples can be found. Halocin KPS1 showed a broad spectrum of antimicrobial activity, including pathogenic strains such as: *Streptococcus mutans* MTCC896, *Bacillus subtilis* MTCC1134, *Escherichia coli* MTCC1671, *Staphylococcus aureus* MTCC916, *Pseudomonas aeruginosa* MTCC6538 [45]. The cell-free supernatant of the halocins producing strain *Halobacterium salinarum* ETD5 has also been shown to inhibit the growth of *P. aeruginosa*. However, it has not yet been possible to isolate the halocin responsible for this effect [46,47].

Due to their activity against *E. faecalis* and *S. aureus* and the lack of cytotoxicity and haemolytic activity, AMPs P1 and P3 may be interesting candidates for clinical use. They can also be applied as a starting point for further modifications and the creation of semi-synthetic AMPs with preferred properties. Particularly promising seems to be P3, which induced growth inhibition of both previously indicated strains at MIC50 of 32 µg/ml. P1 caused growth inhibition only for *E. faecalis.* Comparing the MIC50 obtained for our AMPs to AMPs tested in clinical and preclinical trials, peptides such as HB1275, HB1345 or LTX-109 show inhibitory activity at concentrations in the range of 1-4 µg/ml [48]. On the other hand, there are also examples of AMPs that are active in a similar or higher range of concentrations. An example here is the MSI-78 peptide, which has a MIC50 of 31.3 µg/ml for *E. faecalis* and >125 µg/ml for MRSA. It is worth noting that the MSI-78 peptide has reached the third stage of clinical trials and is treated as a promising basis for further modifications to improve its therapeutic properties [49]. Another example may be Peptidolipin B-F for which the MIC50 for MRSA is 64 µg/ml [50]. Combinations of the test peptides with reference antibiotics were also tested to investigate their possible interactions. However, unlike some AMPs, the P1 and P3 peptides did not show any beneficial interactions with the reference antibiotics [51].

In terms of length, molecular weight and net charge, our AMPs fall within the range of other AMPs tested in clinical and preclinical trials. Most of the AMPs in this group are rather small and their length and molecular weight are less than 25 amino acids and 3100 Da. With a molecular weight of 3575 Da and 4298 Da, our peptides can be classified as quite large.

However, there are also larger peptides, such as LL-37 (4490 Da), AP114 (4417 Da) and AP138 (4460 Da). In addition, the net charge of our peptides is almost in the middle of the range for clinically tested AMPs, which extends from -3 to 14 [48].

Molecular modelling techniques are being used in pharmaceutical industry as the effective tool in the drug discovery processes. Most of the drugs produce their effect by interacting with a biological macromolecule such as an enzyme, DNA or receptor. Discovering and developing any new medicine is a long and expensive process. A new compound must have the ability to replace the existing cure with desired response with minimal side effects [52]. Molecular docking is used to model the interaction between compound and target protein receptors.

An estimation of the binding pose demonstrates possible interactions [53]. Docking of substances with known action mode as levofloxacin and gentamicin was used as starting point to compare interactions of studied proteins. Levofloxacin is bactericidal and exerts its antimicrobial effects *via* inhibition of bacterial DNA replication. Levofloxacin exerts its antimicrobial activity via the inhibition of two key bacterial enzymes: DNA gyrase and topoisomerase IV [54]. Gentamicin is effective against both gram-positive and gram-negative organisms but is particularly useful for the treatment of severe gram-negative infections. Structural and cell biological studies suggest that it binds to the 16S rRNA [55]. Due to the published mechanism of action for levofloxacin and gentamicin Topoisomerase IV of *E. faecalis*, DNA gyrase of S. *aureus,* and 16S rRNA of *E. coli* were selected as target receptor for in-silico analyses. Docking of levofloxacin and gentamicin were performed to compare the possible mechanisms of action and potential hydrogen bonds formed in order to stabilize the docked peptide. Kiranpreet Kaur et. al. in their analysis also compared the docking of novel modified AMPs based on AMPs with known function and mechanism of action to find out the best conformation of a compound [56]. In the study conducted by Abdulmalik Aliyu et al., AMPs were designed, and molecular docking were used to compare the antimicrobial activity of novel AMPs with the reference [57]. Roy A. et. al. investigated the inhibition of Topoisomerases by synthetic peptides. Synthetic peptides have been prophesied to be the ideal inhibitors of enzyme activity either alone or in combination with small-molecule drugs [58]. Chandrashekar S. et al. conducted a study, where the protein NAP1 inhibited a *S. aureus* gyrase–AM8191 complex [59]. Molecular modelling study revealed that peptide P1 and P3 interacted with Topoisomerase IV of *E. faecalis* and formed the hydrogen bonds with Ans650 as a reference antibiotic used in treatment – levofloxacin. Peptides P1 and P3 require additional in – depth analyses. Molecular docking of peptides P1 and P3 revealed that peptides formed the hydrogen bonds to different residues than reference antibiotics in case of *E. coli* and *S. aureus.*

According to Zhang et al., AMPs may act as the novel therapeutic option of treating antibiotic-resistant bacteria either alone or applied in a synergistic with antibiotics manner. Penicillin or chloramphenicol combined with nisin improved antibacterial effect in *E. faecalis* where single antibiotic alone had no significant activity [60]. In addition, due to the fact that the tested peptides come from strains living in saline environments, it is quite likely that they are characterized by high tolerance to high salinity. This creates the opportunity not only to use the tested peptides in medicine, but also to use them in industry, for example for food preservation [18,19].

In conclusion, as a result of the analysis of metagenomic sequences from environmental samples from saline environments, three peptide sequences with a high probability of having antimicrobial activity were selected. On the basis of these sequences, peptides with physico-chemical parameters typical for AMP were synthesized. Laboratory evaluation has shown that the P1 and P3 peptides have antimicrobial activity. The P1 peptide inhibited the growth of the *E. faecalis* and the P3 peptide of the *E. faecalis* and *S. aureus* strains. It has also been shown that these peptides have no cytotoxic or haemolytic effect. As a part of the *in silico* analyses, the possible mode of action of the peptides was demonstrated, through pathways analogous to the reference antibiotics. However, it should be borne in mind that the presented variant of the mechanism of action is one of many options, and determining the proper mode of action for each of the peptides requires in-depth research in this area. The results obtained from all tests performed indicate that the P1 and P3 peptides may be valuable candidates for the role of antimicrobial agents.

Acknowledgements:

This research was financed from the following funds: "InterDOC-STARt" project (POWR.03.02.00-00-I033/16-00), the MINI InterDOC-STARt grant, Wroclaw Medical University SUBZ.D250.23.035 project, and WBIOŚ B2111000000038.01 project.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

1. Ma, Y.; Galinski, E.A.; Grant, W.D.; Oren, A.; Ventosa, A. Halophiles 2010: Life in saline environments. *Appl. Environ. Microbiol.* **2010**, *76*, 6971–6981, doi:10.1128/AEM.01868-10.

2. Corral, P.; Amoozegar, M.A.; Ventosa, A. Halophiles and their biomolecules: Recent advances and future applications in biomedicine. *Mar. Drugs* **2020**, *18*, doi:10.3390/md18010033.

3. Lach, J.; Jęcz, P.; Strapagiel, D.; Matera-Witkiewicz, A.; Stączek, P. The methods of digging for "gold" within the salt: Characterization of halophilic prokaryotes and identification of their valuable biological products using sequencing and genome mining tools. *Genes (Basel).* **2021**, *12*, doi:10.3390/genes12111756.

4. Larsson, D.G.J.; Flach, C.F. Antibiotic resistance in the environment. *Nat. Rev. Microbiol.* **2022**, *20*, 257–269, doi:10.1038/s41579-021-00649-x.

5. Hernando-Amado, S.; Coque, T.M.; Baquero, F.; Martínez, J.L. Defining and combating antibiotic resistance from One Health and Global Health perspectives. *Nat. Microbiol.* **2019**, *4*, 1432–1442, doi:10.1038/s41564-019-0503-9.

6. Mookherjee, N.; Hancock, R.E.W. Cationic host defence peptides: Innate immune regulatory peptides as a novel approach for treating infections. *Cell. Mol. Life Sci.* **2007**, *64*, 922–933, doi:10.1007/s00018-007-6475-6.

7. Ouertani, A.; Chaabouni, I.; Mosbah, A.; Long, J.; Barakat, M.; Mansuelle, P.; Mghirbi, O.; Najjari, A.; Ouzari, H.I.; Masmoudi, A.S.; et al. Two new secreted proteases generate a casein-derived antimicrobial peptide in Bacillus cereus food born isolate leading to bacterial competition in milk. *Front. Microbiol.* **2018**, *9*, 1–12, doi:10.3389/fmicb.2018.01148.

8. Simons, A.; Alhanout, K.; Duval, R.E. Bacteriocins, antimicrobial peptides from bacterial origin: Overview of their biology and their impact against multidrug-resistant bacteria. *Microorganisms* **2020**, *8*, doi:10.3390/microorganisms8050639.

9. Haney, E.F.; Straus, S.K.; Hancock, R.E.W. Reassessing the host defense peptide landscape. *Front. Chem.* **2019**, *7*, 1–22, doi:10.3389/fchem.2019.00043.

10. Kumar, V.; Singh, B.; van Belkum, M.J.; Diep, D.B.; Chikindas, M.L.; Ermakov, A.M.; Tiwari, S.K. Halocins, natural antimicrobials of Archaea: Exotic or special or both? *Biotechnol. Adv.* **2021**, *53*, 107834, doi:10.1016/j.biotechadv.2021.107834.

11. Souza, P.F.N.; Marques, L.S.M.; Oliveira, J.T.A.; Lima, P.G.; Dias, L.P.; Neto, N.A.S.; Lopes, F.E.S.; Sousa, J.S.; Silva, A.F.B.; Caneiro, R.F.; et al. Synthetic antimicrobial peptides: From choice of the best sequences to action mechanisms. *Biochimie* **2020**, *175*, 132–145, doi:10.1016/j.biochi.2020.05.016.

12. Dijksteel, G.S.; Ulrich, M.M.W.; Middelkoop, E.; Boekema, B.K.H.L. Review: Lessons Learned From Clinical Trials Using Antimicrobial Peptides (AMPs). *Front. Microbiol.* **2021**, *12*, doi:10.3389/fmicb.2021.616979.

13. Magana, M.; Pushpanathan, M.; Santos, A.L.; Leanse, L.; Fernandez, M.; Ioannidis, A.; Giulianotti, M.A.; Apidianakis, Y.; Bradfute, S.; Ferguson, A.L.; et al. The value of

antimicrobial peptides in the age of resistance. *Lancet Infect. Dis.* **2020**, *20*, e216–e230, doi:10.1016/S1473-3099(20)30327-3.

14. Wu, C.L.; Hsueh, J.Y.; Yip, B.S.; Chih, Y.H.; Peng, K.L.; Cheng, J.W. Antimicrobial peptides display strong synergy with vancomycin against vancomycin-resistant E. Faecium, S. Aureus, and wild-type E. coli. *Int. J. Mol. Sci.* **2020**, *21*, 1–12, doi:10.3390/ijms21134578.

15. Liu, S.; Wang, S.; Liu, X.; Wen, L.; Zou, J. Effects of dietary antimicrobial peptides on intestinal morphology, antioxidant status, immune responses, microbiota and pathogen disease resistance in grass carp Ctenopharyngodon idellus. *Microb. Pathog.* **2022**, *165*, 105386, doi:10.1016/j.micpath.2021.105386.

16. Valdez-Miramontes, C.E.; De Haro-Acosta, J.; Aréchiga-Flores, C.F.; Verdiguel-Fernández, L.; Rivas-Santiago, B. Antimicrobial peptides in domestic animals and their applications in veterinary medicine. *Peptides* **2021**, *142*, doi:10.1016/j.peptides.2021.170576.

17. Colombo, M.; Masiero, S.; Rosa, S.; Caporali, E.; Toffolatti, S.L.; Mizzotti, C.; Tadini, L.; Rossi, F.; Pellegrino, S.; Musetti, R.; et al. NoPv1: a synthetic antimicrobial peptide aptamer targeting the causal agents of grapevine downy mildew and potato late blight. *Sci. Rep.* **2020**, *10*, 1–18, doi:10.1038/s41598-020-73027-x.

18. Santos, J.C.P.; Sousa, R.C.S.; Otoni, C.G.; Moraes, A.R.F.; Souza, V.G.L.; Medeiros, E.A.A.; Espitia, P.J.P.; Pires, A.C.S.; Coimbra, J.S.R.; Soares, N.F.F. Nisin and other antimicrobial peptides: Production, mechanisms of action, and application in active food packaging. *Innov. Food Sci. Emerg. Technol.* **2018**, *48*, 179–194, doi:10.1016/j.ifset.2018.06.008.

19. Abdulhussain Kareem, R.; Razavi, S.H. Plantaricin bacteriocins: As safe alternative antimicrobial peptides in food preservation—A review. *J. Food Saf.* **2020**, *40*, doi:10.1111/jfs.12735.

20. Rodriguez-Valera, F.; Juez, G.; Kushner, D.J. Halocins: salt-dependent bacteriocins produced by extremely halophilic rods. *Can. J. Microbiol.* **1982**, 3–6.

21. Chhetri, V.; Prakitchaiwattana, C.; Settachaimongkon, S. A potential protective culture; halophilic Bacillus isolates with bacteriocin encoding gene against Staphylococcus aureus in salt added foods. *Food Control* **2019**, *104*, 292–299, doi:10.1016/j.foodcont.2019.04.043.

22. Li, D.; Liu, C.M.; Luo, R.; Sadakane, K.; Lam, T.W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **2015**, *31*, 1674–1676, doi:10.1093/bioinformatics/btv033.

23. Mikheenko, A.; Saveliev, V.; Gurevich, A. MetaQUAST: Evaluation of metagenome assemblies. *Bioinformatics* **2016**, *32*, 1088–1090, doi:10.1093/bioinformatics/btv697.

24. Santos-Júnior, C.D.; Pan, S.; Zhao, X.M.; Coelho, L.P. Macrel: Antimicrobial peptide screening in genomes and metagenomes. *PeerJ* **2020**, *8*, 1–20, doi:10.7717/peerj.10555.

25. Xiao, X.; Wang, P.; Lin, W.Z.; Jia, J.H.; Chou, K.C. IAMP-2L: A two-level multi-label classifier for identifying antimicrobial peptides and their functional types. *Anal. Biochem.* **2013**, *436*, 168–177, doi:10.1016/j.ab.2013.01.019.

26. Waghu, F.H.; Barai, R.S.; Gurung, P.; Idicula-Thomas, S. CAMPR3: A database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Res.* **2016**, *44*, D1094–D1097, doi:10.1093/nar/gkv1051.

27. Wang, G.; Li, X.; Wang, Z. APD3: The antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res.* **2016**, *44*, D1087–D1093, doi:10.1093/nar/gkv1278.

28. Weinstein, M.P. Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically. *Natl. Comm. Clin. Lab. Stand.* **2018**.

29. Susceptibility testing of infectious agents and evaluation of performance of antimicrobial susceptibility test devices – Part 1: Broth micro-dilution reference method for testing the in vitro activity of antimicrobial agents against rapidly growing aerobi. *Int. Organ. Standarization* **2019**.

30. Clinical laboratory testing and in vitro diagnostic test systems – Reference method for testing the in vitro activity of antimicrobial agents against yeast fungi involved in infectious diseases. *Int. Organ. Standarization* **2012**.

31. Gabrielson, J.; Hart, M.; Jarelöv, A.; Kühn, I.; McKenzie, D.; Möllby, R. Evaluation of redox indicators and the use of digital scanners and spectrophotometer for quantification of microbial growth in microplates. *J. Microbiol. Methods* **2002**, *50*, 63–73, doi:10.1016/S0167-7012(02)00011-8.

32. Francisco, F.L.; Saviano, A.M.; Pinto, T. de J.A.; Lourenço, F.R. Development, optimization and validation of a rapid colorimetric microplate bioassay for neomycin sulfate in pharmaceutical drug products. *J. Microbiol. Methods* **2014**, *103*, 104–111, doi:10.1016/j.mimet.2014.05.023.

33. Sabaeifard, P.; Abdi-Ali, A.; Soudi, M.R.; Dinarvand, R. Optimization of tetrazolium salt assay for Pseudomonas aeruginosa biofilm using microtiter plate method. *J. Microbiol. Methods* **2014**, *105*, 134–140, doi:10.1016/j.mimet.2014.07.024.

34. Repetto, G.; del Peso, A.; Zurita, J.L. Neutral red uptake assay for the estimation of cell viability/ cytotoxicity. *Nat. Protoc.* **2008**, *3*, 1125–1131, doi:10.1038/nprot.2008.75.

35. Jones, D.T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **1999**, *292*, 195–202, doi:10.1006/jmbi.1999.3091.

36. Allouche, A. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2012**, *32*, 174–182, doi:10.1002/jcc.

37. Kozakov, D.; Hall, D.R.; Xia, B.; Porter, K.A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S. The ClusPro web server for protein-protein docking. *Nat. Protoc.* **2017**, *12*, 255–278, doi:10.1038/nprot.2016.169.

38. Halgren, T.A. MMFF VII. Characterization of MMFF94, MMFF94s, and Other Widely Available Force Fields for Conformational Energies and for Intermolecular Interaction Energies and Geometries. **2000**, *20*, 730–748.

39. Eberhardt, J.; Santos-Martins, D.; Tillack, A.F.; Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model.* **2021**, *61*, 3891–3898, doi:10.1021/acs.jcim.1c00203.

40. Ying, L.; Zhu, H.; Shoji, S.; Fredrick, K. Roles of specific aminoglycoside–ribosome interactions in the inhibition of translation. *Rna* **2019**, *25*, 247–254, doi:10.1261/rna.068460.118.

41. Fàbrega, A.; Madurga, S.; Giralt, E.; Vila, J. Mechanism of action of and resistance to quinolones. *Microb. Biotechnol.* **2009**, *2*, 40–61, doi:10.1111/j.1751-7915.2008.00063.x.

42. Frieri, M.; Kumar, K.; Boutin, A. Antibiotic resistance. *J. Infect. Public Health* **2017**, *10*, 369–378, doi:10.1016/j.jiph.2016.08.007.

43. de Castro, I.; Mendo, S.; Caetano, T. Antibiotics from Haloarchaea: What Can We Learn from Comparative Genomics? *Mar. Biotechnol.* **2020**, *22*, 308–316, doi:10.1007/s10126-020-09952-9.

44. Kis-Papo, T.; Oren, A. Halocins: Are they involved in the competition between halobacteria in saltern ponds? *Extremophiles* **2000**, *4*, 35–41, doi:10.1007/s007920050005.

45. Kavitha, P.; Lipton, A.P.; Sarik, A.R.; Aishwarya, M.. Growth characteristics and halocin production by a new isolate, Haloferax volcanii KPS1 from Kovalam Solar Saltern (India). *Res. J. Biol. Sci.* **2011**, *6*, 257–262.

46. Ghanmi, F.; Carré-Mlouka, A.; Vandervennet, M.; Boujelben, I.; Frikha, D.; Ayadi, H.; Peduzzi, J.; Rebuffat, S.; Maalej, S. Antagonistic interactions and production of halocin antimicrobial peptides among extremely halophilic prokaryotes isolated from the solar saltern of Sfax, Tunisia. *Extremophiles* **2016**, *20*, 363–374, doi:10.1007/s00792-016-0827-9.

47. Ghanmi, F.; Carré-Mlouka, A.; Zarai, Z.; Mejdoub, H.; Peduzzi, J.; Maalej, S.; Rebuffat, S. The extremely halophilic archaeon Halobacterium salinarum ETD5 from the solar saltern of Sfax (Tunisia) produces multiple halocins. *Res. Microbiol.* **2020**, *171*, 80–90, doi:10.1016/j.resmic.2019.09.003.

48. Koo, H.B.; Seo, J. Antimicrobial peptides under clinical investigation. *Pept. Sci.* **2019**, *111*, doi:10.1002/pep2.24122.

49. Lin, B.; Hung, A.; Li, R.; Barlow, A.; Singleton, W.; Matthyssen, T.; Sani, M.A.; Hossain, M.A.; Wade, J.D.; O'Brien-Simpson, N.M.; et al. Systematic comparison of activity and mechanism of antimicrobial peptides against nosocomial pathogens. *Eur. J. Med. Chem.* **2022**, *231*, 114135, doi:10.1016/J.EJMECH.2022.114135.

50. Barbosa, F.; Pinto, E.; Kijjoa, A.; Pinto, M.; Sousa, E. Targeting antimicrobial drug resistance with marine natural products. *Int. J. Antimicrob. Agents* **2020**, *56*, 106005, doi:10.1016/J.IJANTIMICAG.2020.106005.

51. Zhu, Y.; Hao, W.; Wang, X.; Ouyang, J.; Deng, X.; Yu, H.; Wang, Y. Antimicrobial peptides, conventional antibiotics, and their synergistic utility for the treatment of drug-resistant infections. *Med. Res. Rev.* **2022**, *42*, 1377–1422, doi:10.1002/med.21879.

52. Leach, A. *Molecular Modelling: Principles and Applications*; 2001;

53. Maden, S.F.; Sezer, S.; Acuner, S.E. Fundamentals of Molecular Docking and Comparative Analysis of Protein–Small-Molecule Docking Approaches. In *Molecular Docking - Recent Advances*; 2022.

54. Bearden, D.; Danziger, L. Mechanism of Action of and Resistance to Quinolones. *Pharmacotherapy* **2001**, *21*, doi:10.1007/BF02855708.

55. Borovinskaya, M.A.; Pai, R.D.; Zhang, W.; Schuwirth, B.S.; Holton, J.M.; Hirokawa, G.; Kaji, H.; Kaji, A.; Cate, J.H.D. Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nat. Struct. Mol. Biol.* **2007**, *14*, 727–732, doi:10.1038/nsmb1271.

56. Kaur, K.; Kaur, P.; Mittal, A.; Nayak, S.K.; Khatik, G.L. Design and molecular docking studies of novel antimicrobial peptides using autodock molecular docking software. *Asian J. Pharm. Clin. Res.* **2017**, *10*, 28–31, doi:10.22159/ajpcr.2017.v10s4.21332.

57. Aliyu, A.; Ibrahim, Y.; Tytler, B.; Olowo-Okere, A. Antimicrobial Peptide Design, Molecular Docking and ADMET Studies Against the Methicillin-Resistant Staphylococcus aureus and Carbapenem- resistant and Carbapenemase-producing Pseudomonas aeruginosa. *Trends Pept. Protein Sci.* **2022**, *7*.

58. Roy, A.; D'Annessa, I.; Nielsen, C.J.F.; Tordrup, D.; Laursen, R.R.; Knudsen, B.R.; Desideri, A.; Andersen, F.F. Peptide Inhibition of Topoisomerase IB from Plasmodium falciparum . *Mol. Biol. Int.* **2011**, *2011*, 1–10, doi:10.4061/2011/854626.

59. Napin, A.; Protein, B. In Vitro and In Silico Screening and Characterization of. **2021**, 1–19.

60. Zhang, Q.Y.; Yan, Z. Bin; Meng, Y.M.; Hong, X.Y.; Shao, G.; Ma, J.J.; Cheng, X.R.; Liu, J.; Kang, J.; Fu, C.Y. Antimicrobial peptides: mechanism of action, activity and clinical potential. *Mil. Med. Res.* **2021**, *8*, 1–25, doi:10.1186/s40779-021-00343-2.

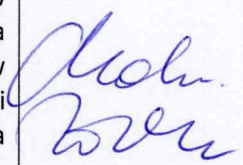
Jakub Lach, Paulina Jęcz, Dominik Strapagiel, Agnieszka Matera-Witkiewicz, Paweł Stączek (2021) "The Methods of Digging for "Gold" within the Salt: Characterization of Halophilic Prokaryotes and Identification of Their Valuable Biological Products Using Sequencing and Genome Mining Tools" *Genes* DOI: https://doi.org/10.3390/genes12111756 - artykuł wchodzący w skład rozprawy doktorskiej (IF: 4.414, pkt. MEiN: 100)

Imię i nazwisko	Rola autora	Podpis
Mgr Jakub Lach	Wiodąca rola w opracowaniu koncepcji pracy. Współudział w przygotowaniu wstępnej wersji manuskryptu, w szczególności sekcji: „Experiment Strategies Used in Halophiles Research", „In Silico Methods for Identification of Novel Halophiles Bioproducts", „Conclusions". Wiodąca rola w przygotowaniu finalnej wersji manuskryptu po uwzględnieniu uwag współautorów i recenzentów. Wiodąca rola w przygotowaniu odpowiedzi dla recenzentów. Autor korespondencyjny. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 50%	
Dr Paulina Jęcz	Współudział w opracowaniu koncepcji pracy. Współudział w przygotowaniu wstępnej wersji manuskryptu, w szczególności sekcji: „Introduction", „Global Distribution of Hypersaline Environments", „Biodiversity of Hypersaline Environments". Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 20%	
Dr Dominik Strapagiel, prof. UŁ	Współudział w opracowaniu koncepcji pracy. Krytyczna ocena wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 10%	
Dr inż. Agnieszka Matera-Witkiewicz	Współudział w opracowaniu koncepcji pracy. Recenzja i korekty wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 10%	

Dr hab. Paweł Stączek, prof. UŁ	Współudział w opracowaniu koncepcji pracy. Krytyczna ocena wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekta językowa. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 10%	

Jakub Lach, Klaudyna Królikowska, Monika Baranowska, Magdalena Krupińska, Dominik Strapagiel, Agnieszka Matera-Witkiewicz, Paweł Stączek (2023) „A First Insight into the Polish Bochnia Salt Mine Metagenome" *Environ Sci Pollut Res* DOI: https://doi.org/10.1007/s11356-023-25770-7 - artykuł wchodzący w skład rozprawy doktorskiej (IF: 5.190, pkt. MEiN: 100)

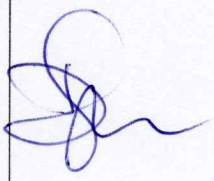
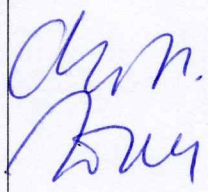
Imię i nazwisko	**Rola autora**	**Podpis**
Mgr Jakub Lach	Wiodąca rola w opracowaniu koncepcji pracy. Przygotowanie bibliotek DNA do sekwencjonowania metagenomów. Przeprowadzenie większości analiz bioinformatycznych i statystycznych (za wyjątkiem identyfikacji i charakterystyki AMP). Przygotowanie wszystkich rycin i tabel. Przygotowanie wstępnej wersji manuskryptu. Wiodąca rola w przygotowaniu finalnej wersji manuskryptu po uwzględnieniu uwag współautorów i recenzentów. Wiodąca rola w przygotowani odpowiedzi dla recenzentów. Autor korespondencyjny. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 50%	Jakub Lach
Mgr Klaudyna Królikowska	Współudział w zbieraniu prób środowiskowych. Współudział w izolacji DNA środowiskowego, przygotowaniu bibliotek DNA dla amplikonów 16S rRNA do sekwencjonowania oraz przeprowadzeniu sekwencjonowania DNA. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 10%	Klaudyna Królikowska
Mgr Monika Baranowska	Współudział w izolacji DNA środowiskowego, przygotowaniu bibliotek DNA dla amplikonów 16S rRNA do sekwencjonowania oraz przeprowadzeniu sekwencjonowania DNA. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 5%	Monika Baranowska
Mgr inż. Magdalena Krupińska	Identyfikacja i charakterystyka AMP w badanych metagenomach. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 5%	Magdalena Krupińska

Dr Dominik Strapagiel, prof. UŁ	Współudział w opracowaniu koncepcji pracy. Współudział w zbieraniu prób środowiskowych. Krytyczna ocena wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 10%	
Dr inż. Agnieszka Matera-Witkiewicz	Współudział w opracowaniu koncepcji pracy. Recenzja i korekty wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 5%	
Dr hab. Paweł Stączek, prof. UŁ	Współudział w opracowaniu koncepcji pracy. Współudział w zbieraniu prób środowiskowych. Krytyczna ocena wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekta językowa. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 15%	

Jakub Lach, Dominik Strapagiel, Agnieszka Matera-Witkiewicz, Paweł Stączek (2023) "Draft Genomes of Halophilic Chromohalobacter and Halomonas Strains Isolated from Brines of The Carpathian Foreland, Poland" *Journal of Genomics* DOI: 10.7150/jgen.80829 - artykuł wchodzący w skład rozprawy doktorskiej (IF: brak, pkt. MEiN: 100)

Imię i nazwisko	**Rola autora**	**Podpis**
Mgr Jakub Lach	Wiodąca rola w opracowaniu koncepcji pracy. Przeprowadzenie wszystkich analiz bioinformatycznych. Przygotowanie wszystkich rycin i tabel. Przygotowanie wstępnej wersji manuskryptu. Wiodąca rola w przygotowaniu finalnej wersji manuskryptu po uwzględnieniu uwag współautorów i recenzentów. Wiodąca rola w przygotowani odpowiedzi dla recenzentów. Autor korespondencyjny. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 70%	
Dr Dominik Strapagiel, prof. UŁ	Współudział w opracowaniu koncepcji pracy. Przygotowanie bibliotek DNA oraz przeprowadzenie sekwencjonowania genomów analizowanych szczepów. Krytyczna ocena wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 12,5%	
Dr Agnieszka Matera-Witkiewicz	Współudział w opracowaniu koncepcji pracy. Recenzja i korekty wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej. Szacunkowy udział (%) – ok. 5%	
Dr hab. Paweł Stączek, prof. UŁ	Współudział w opracowaniu koncepcji pracy. Dostarczenie izolatów DNA analizowanych szczepów. Krytyczna ocena wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekta językowa. Korekty i akceptacja wersji końcowej. Szacunkowy udział (%) – ok. 12,5%	

Jakub Lach, Dominik Strapagiel, Agnieszka Matera-Witkiewicz, Paweł Stączek (2023) "Draft genomes of halophilic Archaea strains isolated from brines of the Carpathian Foreland, Poland" *Journal of Genomics* DOI: 10.7150/jgen.82493 - artykuł wchodzący w skład rozprawy doktorskiej (IF: brak, pkt. MEiN: 100)

Imię i nazwisko	**Rola autora**	**Podpis**
Mgr Jakub Lach	Wiodąca rola w opracowaniu koncepcji pracy. Przeprowadzenie wszystkich analiz bioinformatycznych. Przygotowanie wszystkich rycin i tabel. Przygotowanie wstępnej wersji manuskryptu. Wiodąca rola w przygotowaniu finalnej wersji manuskryptu po uwzględnieniu uwag współautorów i recenzentów. Wiodąca rola w przygotowanie odpowiedzi dla recenzentów. Autor korespondencyjny. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 70%	
Dr Dominik Strapagiel, prof. UŁ	Współudział w opracowaniu koncepcji pracy. Przygotowanie bibliotek DNA oraz przeprowadzenie sekwencjonowania genomów analizowanych szczepów. Krytyczna ocena wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 12,5%	
Dr inż. Agnieszka Matera-Witkiewicz	Współudział w opracowaniu koncepcji pracy. Recenzja i korekty wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 5%	
Dr hab. Paweł Stączek, prof. UŁ	Współudział w opracowaniu koncepcji pracy. Dostarczenie izolatów DNA analizowanych szczepów. Krytyczna ocena wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekta językowa. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 12,5%	

Jakub Lach, Magdalena Krupińska, Aleksandra Mikołajczyk, Dominik Strapagiel, Paweł Stączek, Agnieszka Matera-Witkiewicz (2023) „Novel AMP from saline environments - promising glimmer for inhibition of multidrug resistant E.faecalis and S.aureus infections?" *Eur J Pharm Biopharm* (IF: 5.589, pkt. MEiN: 100) – nieopublikowany artykuł wchodzący w skład rozprawy doktorskiej, obecnie w trakcie recenzji

Imię i nazwisko	**Rola autora**	**Podpis**
Mgr Jakub Lach	Współudział w opracowaniu koncepcji pracy. Współudział w identyfikacji i charakterystyce AMP w analizowanych metagenomach. Przeprowadzenie części doświadczeń dotyczących oceny właściwości przeciwdrobnoustrojowych analizowanych AMP. Współudział w przygotowaniu wstępnej wersji manuskryptu, w szczególności wstępu oraz części dotyczących laboratoryjnej oceny właściwości przeciwdrobnoustrojowych badanych AMP. Przygotowanie tabel: 1, 2, 3, ?. Wiodąca rola w przygotowaniu finalnej wersji manuskryptu po uwzględnieniu uwag współautorów i recenzentów. Wiodąca rola w przygotowanie odpowiedzi dla recenzentów. Autor korespondencyjny. Korekty i akceptacja wersji końcowej. Szacunkowy udział (%) – ok. 40%	Jakub Lach
Mgr inż. Magdalena Krupińska	Współudział w opracowaniu koncepcji pracy. Przeprowadzenie analiz *in silico*. Współudział w przygotowaniu wstępnej wersji manuskryptu, w szczególności części dotyczących analizy *in silico*. Przygotowanie tabel: 1, 2, 3, ? oraz rycina 1, 2, 3, ?. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 20%	Magdalena Krupińska
Mgr inż. Aleksandra Mikołajczyk	Współudział w opracowaniu koncepcji pracy. Przeprowadzenie większości doświadczeń dotyczących oceny właściwości przeciwdrobnoustrojowych analizowanych AMP. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 10%	Aleksandra Mikołajczyk

Dr Dominik Strapagiel, prof. UŁ	Współudział w opracowaniu koncepcji pracy. Krytyczna ocena wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej. Szacunkowy udział (%) – ok. 5%	
Dr hab. Paweł Stączek, prof. UŁ	Współudział w opracowaniu koncepcji pracy. Krytyczna ocena wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej. Szacunkowy udział (%) – ok. 10%	
Dr inż. Agnieszka Matera-Witkiewicz	Wiodąca rola w opracowaniu koncepcji pracy. Recenzja i korekty wstępnej wersji manuskryptu. Edycja pracy. Udział w przygotowywaniu odpowiedzi dla recenzentów. Współudział w przygotowaniu finalnej wersji manuskryptu. Korekty i akceptacja wersji końcowej Szacunkowy udział (%) – ok. 15%