EXOMETABOLITES OF ENDOSPORE-FORMING BACTERIA OF *BACILLUS* GENUS IDENTIFIED BY GENOMIC-METABOLOMIC PROFILING

A. M. OSTAPCHUK, M. D. SHTENIKOV✉, V. O. IVANYTSIA

Odesa I. I. Mechnykov National University, Ukraine;
✉*e-mail: shtenikovnik@gmail.com*

A number of unique bioactive metabolites produced by marine bacilli is already known but the metabolomic of these bacteria is underinvestigated. The aim of this work was to carry out the comparative analysis of metabolomic and genomic traits of Bacillus velezensis ONU 553, Bacillus pumilus ONU 554, Bacillus subtilis ONU 559 strains isolated from Black Sea bottom sediments. Organic extracts of each strain were analyzed using high-resolution liquid chromatography-mass spectrometry. General annotation of genomes was performed using PATRIC, search for secretory signals in the primary structure of selected proteins with using Signal IP analysis. The search of biosynthetic gene clusters was performed using antiSMASH, PRISM 3 and BiG-SCAPEs, the reconstruction of metabolites- with PRISM 3 and TransATor analyzes. The study allowed to find and identify 90, 33 and 43 metabolites in the strains Bacillus velezensis ONU 553, Bacillus pumilus ONU 554 and Bacillus subtilis ONU 559, respectively. The compounds found in metabolome were subdivided into two groups: those which are known members of the genus Bacillus and those new to both genus and prokaryotes in general. Among the secondary metabolites of studied strains, the variants of the nonribosomal peptide class surfactins (anhteisoC16-surfactin, surfactin B2-me ester), gageostatins, fengycins and amicoumacins, and the secreted protease inhibiting pentapeptide GPFPI were identified. The biosynthetic clusters of lipopeptides of the pumilacidin subgroup and amicoumacin antibiotic AI-77A were identified for the first time with the use of bioinformatic tools. The data obtained deepen the understanding of the marine bacilli biosynthetic potential.

K e y w o r d s: Bacillus, metabolomics, genomics, bioinformatic analysis, LC-MS.

In our century the search for previously unknown secondary metabolites is one of the main problems of the microorganisms biochemistry, especially bacteria. This allows us to solve some both applied and fundamental issues in prokaryotic biology. Particularly interesting is the study of microorganisms isolated from inaccessible marine biotopes with extreme conditions, which were understudied on earlier stages of metabolite screening researches. And the representatives of the genus *Bacillus* are promising and relatively under-researched sources of new metabolites [1, 2].

The study of bacteria of the genus *Bacillus* includes the search for new biologically active, including antimicrobial, substances using both metabolic and genomic methods. Recently, an integrated genomic-metabolic approach has been actively developed, which provides information on the structure of these compounds and their biosynthesis [3].

The following major groups of antibiotics are already known for members of the *Bacillus* genus: nonribosomal peptides, including lipopeptides, polyketides, amicoumacins, and their derivatives [4, 5]. These molecules are synthesized on polyketide synthetases (PKS) and nonribosomal polypeptides synthetases (NRPS), large molecular complexes where they are assembled from low molecular weight monomers and modified according to the sequence of corresponding modules - groups of protein domains that have corresponding biosynthetic func-

tions. In many cases, the synthesized precursor molecule undergoes significant modifications also after leaving the synthetase (for example, conjugates with fatty acids residues). The diversity of synthetase modules leads to the fact that nonribosomal peptides and polyketides form groups that are similar in structure but differ in several monomers or in residues conjugates such as fatty acids [6]. The systematics of these compounds is still inconsistent, and quite often certain groups of these compounds with moderate "homology" to one class are considered separately. Besides, even organisms that have only one cluster of a particular polypeptide in the genome are capable to synthesize a whole spectrum of variants, thereof [7]. This is due to a certain indeterminacy of the process itself, such as the effect of thermal motion on the work of massive synthetase. In addition to the complex compounds in *Bacillus*, there are also numerous low molecular weight secondary metabolites, with high biological activity usual, such as butirosin, bacilysin, 3,3′-neotrehalosadiamine, and the like [2, 4, 8].

Representatives of this genus are traditionally considered to be inhabitants of the soil, but they are also regularly isolated from aquatic biotopes, including the sea. The metabolomics of marine bacilli is understudied, but many metabolites have already been known to be isolated from marine bacilli [9].

Aim of this work was in the comparative analysis of metabolomic and genomic traits of antagonistic strains of the genus *Bacillus* isolated from bottom sediments of the Black Sea.

Material and Methods

Strains. Antagonistically active strains of *Bacillus velezensis* ONU 553, *Bacillus pumilus* ONU 554, *Bacillus subtilis* ONU 559, which were selected from more than 100 strains of endospore-forming facultative anaerobic bacteria isolated from deep-sea bottom sediments of the Black Sea were studied [10].

Metabolites extraction. The exometabolite profiling was done by using the method from [11]. For this, the bacteria were grown on tryptic soy broth (Sigma-Aldrich) and M medium of the next composition: soy meal – 20 g/l, potato starch – 3 g/l, yeast extract – 1 g/l, $(NH_4)_2SO_4$ – 1 g/l, $MgSO_4$ – 1 g/l, KH_2PO_4 – 0,6 g/l, pH-7) for 7 day on a rotary shaker for temperature 28°C and rotation speed 180 rpm. The metabolites from the culture liquid of the studied strains were extracted with ethyl acetate

and butanol, evaporated to dryness and dissolved in methanol for further manipulations. The total obtained extracts were fractionated by column size-exclusion chromatography using a 1 m column filled with LH 20 Sephadex (Sigma-Aldrich). The elution was carried out with methanol as the mobile phase at a flow rate of 1 ml/min. The fractions were collected every 15 min term, concentrated by evaporation and dissolved in 0.5 ml of methanol.

LC-MS. The component composition of the total extracts and their active fractions were analyzed by high-resolution liquid chromatography-mass spectrometry on an amaZon (Brucker, Daltonics) liquid chromatographic mass spectrometer, BEH C18 50 × 2.1 mm column, 1.7 μm (Waters, Germany), mobile phase A - water with 0.1% formic acid and B - acetonitrile with 0.1% formic acid. A linear gradient from 5 to 95% was used for 9 min [11].

Genome sequencing. DNA was isolated using a modified method [12]. The daily bacterial culture was grown on the LB broth medium at 30°C on an orbital shaker at 160 rpm. Sequencing was done using two libraries with short and long inserts on HiSeq 1500 equipment (Illumina). The *de novo* read assembly is done using the Newbler 2.8 (Roche) assembler.

Genome analysis. The TYGS tool was used to determine the species based on the 16S RNA gene sequence and full-genome comparisons [13]. General annotation of genomes was performed using PATRIC [14], search for secretory signals in the primary structure of selected proteins - Signal IP [15]. The search and analysis of biosynthetic gene clusters were performed using antiSMASH [16], PRISM 3 [17] and BiG-SCAPE [18]. Reconstruction of metabolites according to the specific cluster was performed for all clusters except trans-PKC using PRISM 3, and for trans-PKC clusters using TransA-Tor [8]. Experimentally verified reference genetic clusters were taken from the MIBiG database [19]. Homology visualization for biosynthetic clusters was performed using Easyfig [20].

Results and Discussion

By using liquid chromatography-mass spectrometry in the culture *Bacillus velezensis* ONU 553 were identified 90 metabolites, *Bacillus pumilus* ONU 554 - 33 and *Bacillus subtilis* ONU 559 – 43, respectively. These compounds can be divided in two groups: those which are known as a products of members of the genus *Bacillus* and new to both this

genus and prokaryotes in general. The compounds of the second group are subject to further study and identification.

Integrated data, obtained by the LC-MS investigation of samples from different methods of growing and extraction of cultures if studied strains are presented on tables 1, 2, and 3.

Among the secondary metabolites of the studied strains, were identified variants of the best-studied class of nonribosomal peptides - surfactins (Tables 1, 2, 3). The metabolites of this class are already known in the representatives of *B. velezensis*, *B. pumilus*, *B. subtilis* [2, 4, 21], and this was confirmed in our results. Surfactins are acylated cyclic nonribosomal heptapeptides with strong surface activity as well as hemolytic and antimicrobial (bacteriostatic and fungicidal) properties [6]. They are characterized by the structural pattern GluLeuLeu^d (Val/Ala/Iso)AspLeu^dLeu and contains β-hydroxy fatty acid with a chain length in the range of 11-14 atoms. The lichenisins subgroup is characterized by the substitution of glutamic acid for glutamine; besides, they are much more likely than isoleucine to be substituted for leucine at positions 2 and 7 rather than in classical surfactins. It is known that the surface activity of lichenisin is higher than that of surfactin. It should be noted that the biological activity of numerous variants of surfactin has not been thoroughly studied to date.

The structure of the surfactin genetic cluster, is very well known, in contrast. It includes three nonribosomal polypeptide synthetase (*srfAA*, *sfrAB*, *srfAC*) subunit genes and several relatively small ancillary open reading frames (ORFs), including the obligatory polyketide and NRPS phosphopantethenyltransferase [6].

In the studied strains the compounds such as anteiso-C16-surfactin, surfactin B2 1-Me ester, etc. were found. All of these compounds are known in members of the *Bacillus subtilis* group (*B. subtilis*, *B. amyloliquefaciens*, *B. pumilus*, etc.) [4, 6]. There is some species-specificity - so as described in the past a "pumilacidins" variants of surfactins with terminal valine (pumilacidins B, D, F, and G) and isoleucine (pumilacidins A, C, and E) were found in bacteria of strain *B. pumilus* ONU 554 (Table 2). Analysis of the genome of this strain using PRISM 3 revealed the presence of a gene cluster encoding the proteins of NRPS, whose modular composition corresponds to the sequence of components of pumilacidin A, C and E. Thus, it can conclude that the pumilacidin cluster identification was the first time performed (Fig. 1).

The strains of *Bacillus velezensis* ONU 553 and *Bacillus subtilis* 559 are most similar in their metabolite profiles. In particular, they are associated by the presence of fengycin class compounds: plipastatins, fengycins *sensu stricto*, and agrastatins (Table 1, 2). Fengycins are decapeptides, unlike surfactins. It is important to note that the surfactant and bactericidal properties of fengycin-like substances are less pronounced than in surfactins. Differences of fengycins sensu stricto from plipastatins lay in the chiral confirmation of two tyrosine residues (at positions 4 and 10) in the cycle. Agrastatins from plipastatins, in turn, differ in the presence of β-hydroxy- rather than amino fatty acid residue in the structure [22].

In strains *B. velezensis* ONU 553 and *B. subtilis* ONU 559 were found (Table 1, 3) already known for *Bacillus cereus* and *Bacillus subtilis* plipastatins A1, A2, B1, B2 [23], which are phospholipase inhibitors. Detected by systems fengycin IX is a derivative of plipastatin A1 and it differs from the latter by the presence of bound K⁺ ion [24]. These strains also possess agrastatin A (which source is *Bacillus subtilis* [22]), an antifungal agent active against many phytopathogens. Lipodepsipeptides SNA 60-367-17 and SNA 60-367-23 are found only in *B. subtilis* strain ONU 559. They are poorly characterized lipopeptides known for *Bacillus sp.* and differ only in the type of fatty acid. They are known for their ability to inhibit aromatase [25].

Genomic studies confirm the presence of biosynthetic clusters similar to the surfactin and fengycin cluster in all strains tested.

Interesting is the issue of the ability to determine the synthesis of different compounds by the same genetic cluster. The most likely solution to this problem is the known phenomenon of domain skipping, which leads to the formation of molecules "processed" by different sets of domains on the same synthetase [7].

In the culture of the strain *Bacillus velezensis* ONU 553 also revealed the presence of three types of gageostatins - A, B, C (Table 1). These compounds have the same peptide skeleton and differ in fatty acid residues at the N-terminus, the recruitment of which in nonribosomal peptides is usually determined by the specificity of the thioesterase domain. The very presence of gageostatins in the strain under study is problematic to explain since a specific genetic cluster from polyketide synthetase domains, that could synthesize them in the genome of this species, has not been identified. Gageostatins

T a b l e 1. Compounds identified in Bacillus velezensis ONU 553 extracts

M[1]	Name[2]	Calculated Mass[3]	Biological Source[4]	Use / Importance[4]
1490,827914	Plipastatin B1 Antibiotic Y10	1490,827238	*B. cereus*	Phospholipase inhibitor
1476,815714	Plipastatin A2	1476,811588	*B. cereus*	Phospholipase inhibitor
1504,848514	Plipastatin B2	1504,842888	*B. cereus*	Phospholipase inhibitor
1448,783354	Agrastatin A	1448,780288	*B. subtilis*	Antifungal agent, active against many plant pathogens
1476,817244	Unknown	1476,811588	*B. cereus*	Phospholipase inhibitor
1462,803854	Plipastatin A1 Fengycin IX	1462,795938	*B. cereus* *B. subtilis*	Phospholipase inhibitor
1476,814834	Unknown	1476,811588	*B. cereus*	Phospholipase inhibitor
993,637664	4-L-Alaninesur-factin B2	993,636188	*B. subtilis*	Unknown
1021,668224	Variant: ai-C14-Surfactin	1021,667488	*B. licheniformis*	Unknown
1007,653654	Gageostatin C	1007,65183	*B. subtilis*	Displays good antifungal activity and cytotoxicity
1039,682324	Gageostatin A	1039,678053	*B. subtilis*	Displays good antifungal activity and cytotoxicit
1053,698034	Gageostatin B	1053,693703	*B. subtilis*	Displays good antifungal activity and cytotoxicit
993,637754	4-L-Alaninesur-factin B1	993,636188	*B. subtilis*	Surface-active agent
1035,687124	Surfactin 1	1035,683138	*B. amylolique-faciens*	Surface-active agent, fibrinolytic, antiinflammatory, anticoagulant
1021,669474	Surfactin	1021,667488	*B. licheniformis*	Surface-active agent, fibrinolytic, antiinflammatory, anticoagulant
1049,701824	ai-C16-Surfactin	1049,698788	*B. licheniformis*	Cyclic lipopeptide
1035,682834	Surfactin B2 1-Me ester	1035,683138	*B. pumilus*	Surface-active agent, fibrinolytic, antiinflammatory, anticoagulant
1049,700514	Surfactin C2 1-Me ester	1049,698788	*B. pumilus*	Surface-active agent, fibrinolytic, antiinflammatory, anticoagulant
1063,717504	Lipodepsipeptides	1063,714438	*B. pumilus*	Lipodepsipeptide antibiotic complex related to surfactin and bacircines

Note: 1 – Molecular mass of compound, identified in this study; 2 – Name of identified compound; 3 – Molecular mass of identified compound according to the LC-MS device database; 4 – source and characteristics of identified compound according to the LC-MS device database (Dictionary of Natural Products, CRC Press)

T a b l e   2. Compounds identified in Bacillus pumilus ONU 554 extracts

M[1]	Name[2]	Calculated Mass[3]	Biological Source[4]	Use/Importance[4]
1035,685954	Surfactin C	1035,683138	*B. subtilis* *B. natto*	Shows cytolytic activity. Has surface-active props. A potent clotting inhibitor in the fibrinogen system
1049,701784	Surfactin C2 1-Me ester	1049,698788	*B. pumilus*	Fibrinolytic, antiinflammatory, anticoagulant
1063,716334	Pumilacidin D	1063,714438	*Bacillus* sp. *B. pumilus*	Phospholipase A2 inhibitor, antiviral agent
1077,733354	Pumilacidin C	1077,730088	*Bacillus* sp. *B. pumilus*	Phospholipase A2 inhibitor, antiviral agent
423,201684	Antibiotic AI-77A	423,200552	*B. pumilus*	Active against Gram-positive and -negative bacteria and tumours, Protects against stress ulcers, Shows antiinflammatory and anticholesterolemic props, related to the amicoumacins
300,205144	Lipoamide A	300,204908	*B. pumilus*	Unknown
1021,671364	Surfactin 7-L-Valinesurfactin C1 Lipobacillin	1021,667488	*B. subtilis* *B. pumilus*	Unknown
1035,688144	Surfactin B2 1-Me ester	1035,683138	*B. pumilus*	Surface-active agent, fibrinolytic, antiinflammatory, anticoagulant

Note: 1 – Molecular mass of compound, identified in this study; 2 – Name of identified compound; 3 – Molecular mass of identified compound according to the LC-MS device database; 4 – source and characteristics of identified compound according to the LC-MS device database (Dictionary of Natural Products, CRC Press)
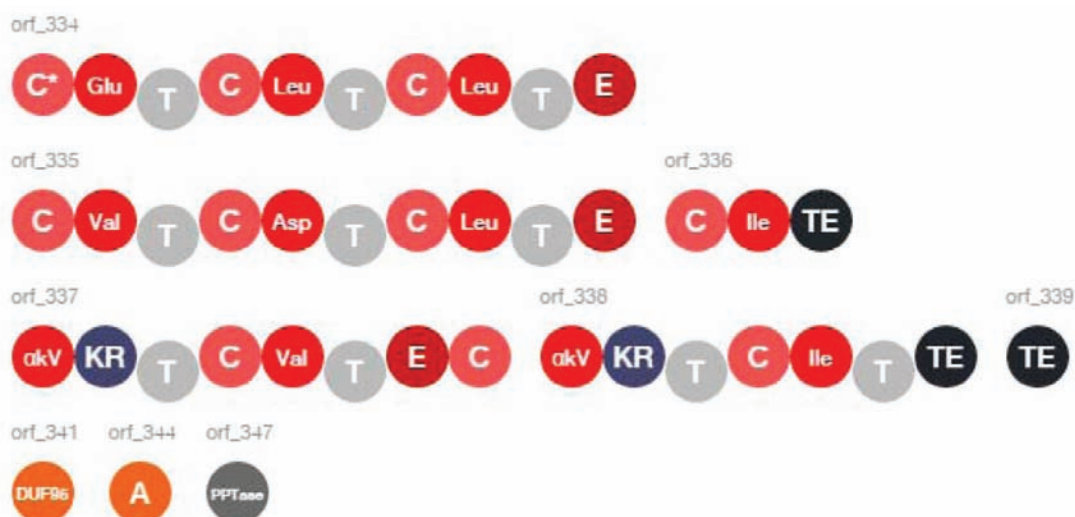

Fig. 1. The domain structure of the products of the pumilacidin cluster from the genome Bacillus pumilus 554, identified by PRISM 3. C – condensation domains; T – thiolation domains, TE – thioesterase domains, KR – ketoreductase domain. Letter combinations indicate amino acids for that activation are responsible the corresponding adenylation domains

T a b l e 3. Compounds identified in Bacillus subtilis ONU 559 extracts

M[1]	Name[2]	Calculated Mass[3]	Biological Source[4]	Use/Importance[4]
1488,853264	SNA 60-367-23	1488,847973	*Bacillus* sp.	Lipodepsipeptide complex
1474,837954	SNA 60-367-17	1474,832323	*Bacillus* sp.	Lipodepsipeptide complex
529,290334	Glycylprolylphe-nylalanylprolyliso-leucine	529,290035	*B. subtilis*	Dipeptidyl carboxypeptidase inhibitor
1476,816594	Plipastatin A2	1476,811588	*B. cereus*	Phospholipase inhibitor
1462,802674	Plipastatin A1 Fengycin IX	1462,795938	*B. cereus* *B. subtilis*	Phospholipase inhibitor
1490,832564	Plipastatin B1	1490,827238	*B. cereus*	Phospholipase inhibitor
1504,847384	Plipastatin B2	1504,842888	*B. cereus*	Phospholipase inhibitor
1448,783754	Agrastatin A	1448,780288	*B. subtilis*	Antifungal agent, active against many plant pathogens
1007,650294	Unknown	1007,651838	*B. subtilis*	Displays good antifungal activity and cytotoxicity
1021,660214	Bamylocin A	1021,649728	*B. amyloliquefaciens*	Depsipeptide, related to surfactin
1035,676354	Surfactin 1	1035,683138	*B. amyloliquefaciens*	Surface-active agent, fibrinolytic, antiinflammatory, anticoagulant
1049,703814	ai-C16-Surfactin	1049,698788	*B. licheniformis*	Cyclic lipopeptide

Note: 1 – Molecular mass of compound, identified in this study; 2 – Name of identified compound; 3 – Molecular mass of identified compound according to the LC-MS device database; 4 – source and characteristics of identified compound according to the LC-MS device database (Dictionary of Natural Products, CRC Press)

have previously been detected in only one strain *Bacillus subtilis* 109GGC020 isolated from the South China Sea [26]. These are linear lipopeptides whose mechanism of biosynthesis has not yet been established, but from the outset, it seems plausible that these are classic nonribosomal peptides; however, a cluster that could be responsible for their biosynthesis has not been identified. We emphasize the fact that these compounds were previously detected in strains of marine origin of similar species. For *Bacillus subtilis* 109GGC020 strain also were described two other types of antimicrobial peptides, gageobactins, and gageotetrins, but they were not detected in this case.

Based on impressive structural similarity, a cautious assumption can be made regarding the synthesis of gageostatin on surfactin synthetase with the skipping of epimerization steps for the corresponding leucine residues and the final addition of fatty acid without final product cyclization.

We conducted a comparative analysis of the domain composition of the surfactin-like cluster of strain *B. velezensis* ONU 553 (Fig. 2) and showed that the size and location of the domains in this cluster and the reference cluster of surfactin biosynthesis are not different.

Another mysterious finding was the pentapeptide GPFPI (dipeptidyl carboxypeptidase inhibitor known for *Bacillus subtilis*) (Table 3). This relatively poorly known compound was discovered in 1992 [27]. It can be a by-product of metabolism, for example, a residue of the signal peptide, or to be a pheromone. A direct search of this oligopeptide in the genome revealed that such a pentapeptide is part of a hypothetical YvkN peptide, the length of which is 76 amino acid residues. This peptide is included in the cluster with three hypothetical membrane peptides and one hypothetical non-membrane peptide (Table 4). It can be assumed that among these peptides are those that provide propeptide processing
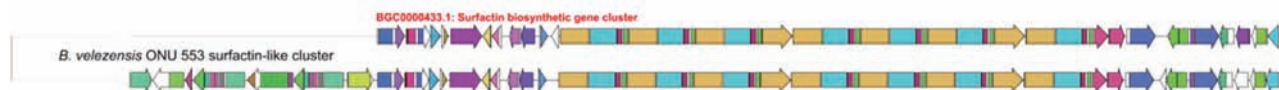
Fig. 2. Comparison cluster from genome Bacillus velezensis ONU 553 (below) with experimentally verified surfactin cluster (above) [19]. Different colors designate domains of the corresponding type: light-brown – condensation domains; blue and red – N and C end domains of adenylation complex domains; green – thiolation domains, dark-pink – thioesterase domain

T a b l e 4. Genomic context of the YvkN peptide gene supposedly containing secreted peptide GPFPI

Start	End	Chain	ORF identification
3387246	3383377	-	Hypothetical protein
3387805	3387446	-	Hypothetical membrane protein YvlD
3388004	3387807	-	Hypothetical membrane protein YvlC
3389106	3388009	-	Hypothetical protein YvlB
3389457	3389131	-	Hypothetical membrane protein YvlA
3389675	3389905	+	YvkN
3390943	3390104	-	Flagellar protein FlaA

with the formation of pentapeptide and its transport. Search with Signal IP 5.0 does not detect any secretory signals in this peptide.

Much more specific, in biochemical terms, is the strain *Bacillus pumilus* ONU 554 (Table 2). The presence of pumilacidins C and D in its exometabolome has already been mentioned. Genomic mining made it possible to identify the gene cluster that encodes the NPPS of this lipopeptide (Fig. 1).

A much less studied compound, lipoamide A, was also found in this strain. Previously, it was already reported for the marine strain *Bacillus pumilus* SP21 [28]. It is characterized by relatively weak (minimal inhibitory concentration is about 100μg/ ml) antimicrobial activity against *S. aureus* and *P. aeruginosa*. Lipoamides are very simple compounds and are an asparagine conjugate with anteiso-fatty acids with chain lengths in the range

11-13. [28]. It is important not to confuse these compounds with lipoic acid amides. The capabilities of modern methods of genomics at this time do not allow to localize in the genome the determinants of lipoamide A biosynthesis.

Another interesting and specific for *Bacillus pumilus* is the compound AI-77A, active against gram-positive and gram-negative bacteria, which is cytotoxic, also having anticholesterol and anti-inflammatory properties [29]. AI-77A belongs to the class of amicoumacins, antibiotics that are very characteristic of bacilli. Only three such compounds are known for actinobacteria and almost thirty for bacilli, which is not characteristic of any other class of antibiotics [5]. The substances of this group consist of a conservative dihydroisocoumarin core and the type-specific moiety coupled with the core by an amide bond.

Compound AI-77A has antibacterial activity against the causative agent of gastric ulcer *Helicobacter pylori*. Its mechanism of action is to bind the molecule simultaneously with the mRNA and 16S RNAs of the small subunit of the ribosome, causing the termination of protein synthesis [5]. Biosynthetic clusters for amicoumacin-like antibiotics have been poorly studied. By comparison, the homology of the hybrid (NRPS /PKC) synthetase cluster from the *Bacillus pumilus* ONU 554 genome and the cluster of the well-studied antibiotic of this group - xenocoumacin (Fig. 3) was demonstrated by Easyfig software.

The analysis of the remaining metabolites (for *Bacillus velezensis* ONU 553 - 58, for *Bacillus pumilus* ONU 554 - 22, *Bacillus subtilis* ONU 559 - 28) is challenging. As was established by the TransATor analyzer, the structure of the trans-AT polyketide synthetase products, which were detected in the genomes studied and identified by antiSMASH as macrolactin, dificidin, and bacillaene does not correspond to the known structure for these compounds [4]. Therefore, the actual products of these genetic complexes need further identification.
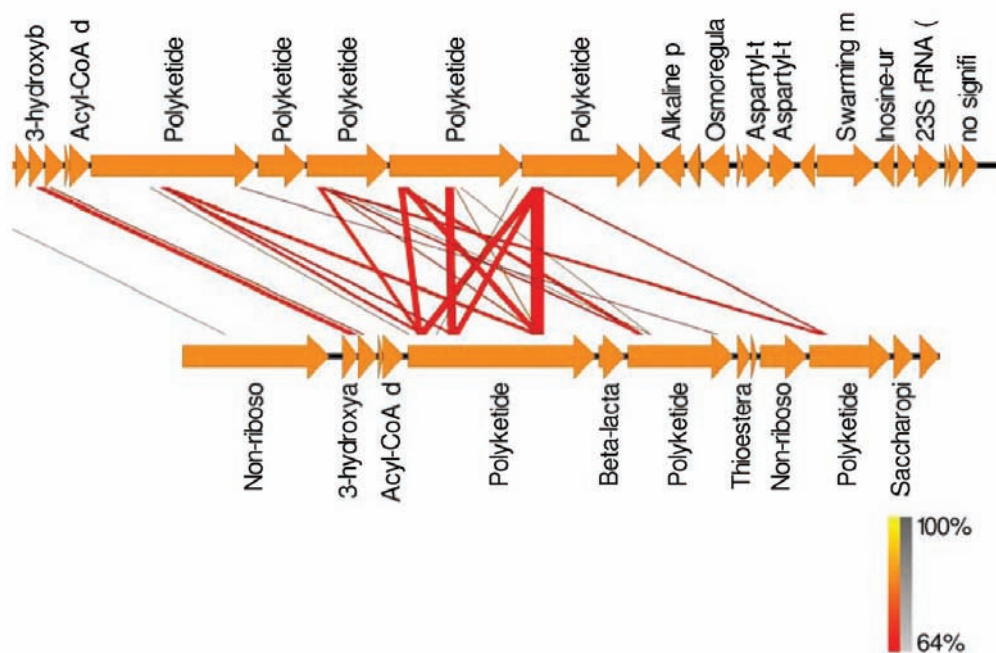
Fig. 3. Comparison of AI-77A cluster form genome Bacillus pumilus ONU 554 (above) and xenocoumacin cluster from genome Xenomonas sp. (below) [19]. Red color corresponds loci with homology >90%

Conclusions. As a result of the conducted researches it is established that all studied strains of Black Sea bacilli synthesize compounds, the structure of which fits into a small number of structural categories, and the diversity of individual compounds strongly depends on the species of the producer. Even strains of relatively phylogenetically related species *B. velezensis* and *B. subtilis* differ significantly in the set of metabolites. It should also be noted that the detected proportion of identified compounds is small, and not yet identified compounds are the subject for further study and identification.

In the studied strains of bacilli isolated from bottom sediments of the Black Sea, biosynthetic clusters have been identified and described that are interesting from a fundamentally biological point of view and may be useful for biotechnological research. Of particular interest may be the peptidase inhibitor pentapeptide *B. subtilis* ONU 554 and the *B. velezensis* ONU 553 surfactin cluster, the study of which can shed light on the little-studied issue of biosynthesis of various compounds on a single polypeptide synthetase complex.

ЕКЗОМЕТАБОЛІТИ ЧОРНОМОРСЬКИХ БАКТЕРІЙ РОДУ *BACILLUS*, ІДЕНТИФІКОВАНІ ШЛЯХОМ ГЕНОМНО-МЕТАБОЛОМНОГО ПРОФІЛЮВАННЯ

А. М. Остапчук, М. Д. Штеніков✉, В. О. Іваниця

Одеський національний університет
ім. І. І. Мечникова, Україна;
✉e-mail: shtenikovnik@gmail.com

Метаболоміку морських бацил мало вивчено, але вже відома низка унікальних біологічно активних метаболітів, які вони продукують. Мета роботи полягала у порівняльному аналізі метаболомних та геномних характеристик антагоністичноактивних штамів, ізольованих із донних відкладень Чорного моря. У роботі методом рідинної хроматомас-спектрометрії досліджували органічні екстракти штамів *Bacillus velezensis* ONU 553, *Bacillus pumilus* ONU 554, *Bacillus subtilis* ONU 559. Загальне анотування геномів виконували за допомогою PATRIC, пошук секреторних сигналів у первинній структурі обраних протеїнів – SignalIP відповідно. Пошук та аналіз кластерів біосинтетичних генів проводили за допомогою antiSMASH, PRISM 3 та BiG-SCAPE, а реконструкцію метаболітів – PRISM 3 та TransATor. Дослідження дозволили виявити та ідентифікувати 90, 33 та 43 метаболіти штамів *Bacillus velezensis* ONU 553, *Bacillus pumilus* ONU 554, *Bacillus subtilis* ONU 559 відповідно. Сполуки виявлені в метаболомі було розділено на дві групи: такі, що вже відомі для представників роду *Bacillus*, та нові як для цього роду, так і для прокаріот взагалі. Серед вторинних метаболітів досліджуваних штамів виявлено варіанти найвивченішого класу нерибосомних пептидів – сурфактинів (антеізо-С16-сурфактин, сурфактин В2 1-Me ester), гагеостатинів, фенгіцинів та амікумацинів, а також секретований пентапептид зі здатністю до інгібування протеаз GPFPI. Із використанням біоінформативних інструментів вперше ідентифіковано біосинтетичні кластери ліпопептидів підгрупи пумілацидинів, амікумацинового антибіотика AI-77A та зроблено припущення про виявлення біосинтетичного кластера пентапептиду GPFPI. Одержані ре-зультати поповнюють наші уявлення про біосинтетичний потенціал морських бацил.

К л ю ч о в і с л о ва: *Bacillus*, метаболоміка, геноміка, біоінформативний аналіз, LC-MS.

References

1. Mandic-Mulec I, Stefanic P, van Elsas JD. Ecology of *Bacillaceae*. *Microbiol Spectr.* 2015; 3(2): TBS-0017-2013.

2. Zhao X, Kuipers OP. Identification and classification of known and putative antimicrobial compounds produced by a wide variety of *Bacillales species*. *BMC Genomics.* 2016; 17(1): 882.

3. Tiam SK, Gugger M, Demay J, Le Manach S, Duval C, Bernard C, Marie B. Insights into the diversity of secondary metabolites of planktothrix using a biphasic approach Combining global genomics and metabolomics. *Toxins (Basel).* 2019; 11(9): 498.

4. Kaspar F, Neubauer P, Gimpel M. Bioactive secondary metabolites from *Bacillus subtilis*: a comprehensive review. *J Nat Prod.* 2019; 82(7): 2038-2053.

5. Tyurin AP, Efimenko TA, Prokhorenko IA, Rogozhin EA, Malanicheva IA, Zenkova VA, Efremenkova OV, Korshun VA. Chapter 12 - Amicoumacins and related compounds: chemistry and biology. *Stud Nat Prod Chem.* 2018; 55: 385-441.

6. Cochrane SA, Vederas JC. Lipopeptides from *Bacillus* and *Paenibacillus* spp.: a gold mine of antibiotic candidates. *Med Res Rev.* 2016; 36(1): 4-31.

7. Wenzel SC, Meiser P, Binz TM, Mahmud T, Müller R. Nonribosomal peptide biosynthesis: point mutations and module skipping lead to chemical diversity. *Angew Chem Int Ed Engl.* 2006; 45(14): 2296-2301.

8. Helfrich EJN, Ueoka R, Dolev A, Rust M, Meoded RA, Bhushan A, Gianmaria C, Costa R, Gugger M, Steinbeck C, Moreno P, Piel J. Automated structure prediction of trans-acyltransferase polyketide synthase products. *Nat Chem Biol.* 2019; 15(8): 813-821.

9. Ma Z, Hu J. Complete genome sequence of a marine-sediment-derived bacterial strain Bacillus velezensis SH-B74, a cyclic lipopeptides producer and a biopesticide. *3 Biotech.* 2019; 9(4): 162.

10. Ivanytsia VO, Shtenikov MD, Ostapchuk AM. Facultatively anaerobic sporeforming bacteria of deep sea sediments of the Black sea. *Microbiol Biotechnol.* 2017; (4(40)): 94-103. (In Ukrainian).

11. Paulus C, Rebets Y, Tokovenko B, Nadmid S, Terekhova LP, Myronovskyi M, Zotchev SB, Rückert C, Braig S, Zahler S, Kalinowski J, Luzhetskyy A. New natural products identified by combined genomics-metabolomics profiling of marine *Streptomyces* sp. MP131-18. *Sci Rep.* 2017; 7: 42382.

12. De S, Kaur G, Roy A, Dogra G, Kaushik R, Yadav P, Singh R, Datta TK, Goswami SL. A simple method for the efficient isolation of genomic DNA from Lactobacilli isolated from traditional Indian fermented milk (*dahi*). *Indian J Microbiol.* 2010; 50(4): 412-418.

13. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun.* 2019; 10(1): 2182.

14. Wattam AR, Brettin T, Davis JJ, Gerdes S, Kenyon R, Machi D, Mao C, Olson R, Overbeek R, Pusch GD, Shukla MP, Stevens R, Vonstein V, Warren A, Xia F, Yoo H. Assembly, annotation, and comparative genomics in PATRIC, the all bacterial bioinformatics resource center. *Methods Mol Biol.* 2018; 1704: 79-101.

15. Armenteros JJA, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol.* 2019; 37(4): 420-423.

16. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 2019; 47(W1): W81-W87.

17. Skinnider MA, Merwin NJ, Johnston CW, Magarvey NA. PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res.* 2017; 45(W1): W49-W54.

18. Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, De Los Santos ELC, Yeong M, Cruz-Morales P, Abubucker S, Roeters A, Lokhorst W, Fernandez-Guerra A, Cappelini LTD, Goering AW, Thomson RJ, Metcalf WW, Kelleher NL, Barona-Gomez F, Medema MH. A computational framework to explore large-scale biosynthetic diversity. *Nat Chem Biol.* 2020; 16(1): 60-68.

19. Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, van der Hooft JJJ, van Santen JA, Tracanna V, Suarez Duran HG, Andreu VP, Selem-Mojica N, Alanjary M, Robinson SL, Lund G, Epstein SC, Sisto AC, Charkoudian LK, Collemare J, Linington RG, Weber T, Medema MH. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* 2020; 48(D1): D454-D458.

20. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics.* 2011; 27(7): 1009-1010.

21. Fan B, Wang C, Song X, Ding X, Wu L, Wu H, Gao X, Borriss R. *Bacillus velezensis* FZB42 in 2018: the gram-positive model strain for plant growth promotion and biocontrol. *Front Microbiol.* 2018; 9: 2491.

22. Patel H, Tscheka C, Edwards K, Karlsson G, Heerklotz H. All-or-none membrane permeabilization by fengycin-type lipopeptides from *Bacillus subtilis* QST713. *Biochim Biophys Acta.* 2011; 1808(8): 2000-2008.

23. Nishikiori T, Naganawa H, Muraoka Y, Aoyagi T, Umezawa H. Plipastatins: new inhibitors of phospholipase A2, produced by *Bacillus cereus* BMG302-fF67. III. Structural elucidation of plipastatins. *J Antibiot (Tokyo).* 1986; 39(6): 755-761.

24. Honma M, Tanaka K, Konno K, Tsuge K, Okuno T, Hashimoto M. Termination of the structural confusion between plipastatin A1 and fengycin IX. *Bioorg Med Chem.* 2012; 20(12): 3793-3798.

25. Balunas M. Natural products as aromatase inhibitors: Identification and structure-activity aspects J. University of Illinois at Chicago, ProQuest Dissertations Publishing, 2007. 3345547.

26. Tareq FS, Lee MA, Lee HS, Lee JS, Lee YJ, Shin HJ. Gageostatins A-C, antimicrobial linear lipopeptides from a marine *Bacillus subtilis.* *Mar Drug.* 2014; 12(2): 871-885.

27. Nagamori Y, Kusaka K, Nishimura T, Okada S. Isolation and Characterization of a Bacterial Dipeptidyl Carboxypeptidase Inhibitor from Bacillus subtilis 3-16-20. *J Fermentat Bioeng.* 1992; 73(4): 277-279.

28. Berrue F, Ibrahim A, Boland P, Kerr RG. Newly isolated marine *Bacillus pumilus* (SP21): A source of novel lipoamides and other antimicrobial agents. *Pure Appl Chem.* 2009; 81(6): 1027-1031.

29. Hayashi H, Shimojima Y, Shirai T, Ishida T, Shibukawa M. AI-77 compounds and pharmaceutically acceptable salts thereof. U.S. Patent No. 4,393,225. 1983.