С.Сейфуллин атындағы Қазақ агротехникалық университетінің **Ғылым жаршысы (пәнаралық)** = **Вестник науки** Казахского агротехнического университета им. С.Сейфуллина **(междисциплинарный)**. - 2022. - №3 (114). –Ч.2. - P. 147-157

COLLECTION OF MICROORGANISMS – BIORESOURCES AND DATABASE

Almagambetov Kairtai Khamitovich
Doctor of Medical Sciences, Professor
Republican Collection of Microorganisms
Nur-Sultan, Kazakhstan
E-mail: rcmkz@list.ru

Nagumanova Gulbakhyt Smanovna
Master of Biological Sciences
Republican Collection of Microorganisms
Nur-Sultan, Kazakhstan
E-mail: gulbakyt_n@mail.ru

Temirkhanov Aslan Zhanaevich
Candidate of biological sciences
Republican Collection of Microorganisms
Nur-Sultan, Kazakhstan
E-mail: rkm13@mail.ru

Sarmurzina Zinigul Serikovna
Candidate of Biological Sciences
Republican Collection of Microorganisms
Nur-Sultan, Kazakhstan
E-mail: sarmurzina@list.ru

Abstract

The Republican Collection of Microorganisms (RCM) is a depository of industrial microorganisms intended for research and analytical work, for the organization of biotechnological production.The Biobank of industrial microorganisms of the RCM stores more than 700 strains from various taxonomic groups (bacteria, actinomycetes, yeast, mycelial fungi). In Kazakhstan, there is no single annotated database (DB) of collectible strains of industrial microorganisms, including information about their- and genotypic characteristics.In this regard, the VRCM is working on the creation of a database, containing the genetic characteristics of deposited strains of industrial microorganisms.To compare genomes, the BLAST+ software package (Basic Local Alignment Search Tool) is used according to the instructions of the National Center for Biotechnology Information (NCBI).The creation of a database of genomic sequences of microorganisms based on the BLAST computer program (nucleotide group) and the storage of information in the Biobank of industrial microorganisms of the RCM will contribute to improving the quality of service function and safety of collection cultures.

Key words: collection of microorganisms; biobank; bioresources; database; BLAST.

Introduction

To preserve the resources of industrial microorganisms, the development of collections is important. It is the collections that make it possible to store deposited strains for a long time and use them in scientific research and biotechnological developments. The fundamental principles of the collections of microorganisms are to maintain the viability, purity and authenticity of the deposited strains [1,2,3].

The relevance of the preservation of microbial bioresources is evidenced by the steady growth in the number of international and national collections [4]. Currently, 803 culture collections in 78 countries contain 3,293,173 strains of bacteria, fungi, viruses and cell lines (http://www.wfcc.info/ccinfo) [5]. In the registry of the World Data Center for Microorganisms – World Data Center for Microorganisms (WDCM) The World Federation of Cultures – World Federation for Culture Collections (WFCC) has registered over 700 collections from more than 70 countries. Including 3 large collections in the USA: American Type Culture Collection (ATCC) contains over 4,500 cultures, Northern Regional Research Laboratory (NRRL) - about 20 thousand cultures, Quartermaster Quartermaster Culture Collection (QM) — about 9 thousand strains of cultures. There are 12 large collections in England, united by the common name British Common wealth Collections of Microorganisms, which support about 15 thousand cultures. The Japanese Federation of Microbial Culture Collections — Japanese Federation of Culture Collections of Microorganisms (JFCC) — has 7,600 cultures representing more than 2,100 species. The French collection at the National Museum of Natural History in Paris supports about 1,000 species. The open catalog fund of the All-Russian Collection of Microorganisms contains cultures of more than 750 genera and 3300 species [6-8].

The rapid growth of "omic" scientific directions, in particular metagenomics, has made it possible to preserve in biobanks not only pure cultures of microorganisms, but also the original biological material of human, animal or plant origin containing DNA samples of the community of microorganisms[9]. For example, the Network of National Environmental Observatories of the USA (NEON; https://www.neonscience.org/) has a microbiome archive with associated metadata, where raw biomaterials containing microbiomes are stored for future research using more modern technologies [10]. In the biobank of agriculturally significant microorganisms, created in 2021 at the National Research Tomsk State University of the Russian Federation for the purpose of collecting and storing biological material, accumulating and systematizing information associated with it, as well as for the purpose of conducting research, cryopreserved along with pure cultures of microorganisms, compost, feces and other biomaterial.

Databases are necessary for collecting and storing information about deposited strains, providing user-friendly access and search functions, as well as standardizing the presentation of information and their organization into knowledge [11]. Database of the global catalog of microorganisms (GCM; http://gcm.wfcc.info/) is an example of comprehensive information about collection microbial resources.

The world-famous National Center for Biotechnological Information of the USA (National Centre for Biotechnology Information - NCBI http://www.ncbi.nlm.nih.gov /) supports the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/index.html) for gene and genomic sequences.This web server is linked to other databases, including the National Library of Medicine http://www.nlm.nih.gov/ and provides a package of computer programs for GenBank search, including BLAST http://blast.ncbi.nlm.nih.gov/Blast.cgi. (Basic Local Alignment Search Tool - a means of finding the main local alignment).

BLAST, known since the 90s of the last century, is widely used to determine the functional and evolutionary relationships between the analyzed nucleotide sequences [12]. The basic local alignment search tool BLAST finds areas of similarity between biological sequences. The program compares the nucleotide sequence of the strain under study with sequence databases and calculates statistical significance.The convenience of the system is that the user does not need to contact various sources to obtain information, it is prepared and presented in an electronic database.

For example, thanks to the GenBank database (accession number MN908947), studies by Chinese scientists on the nucleotide sequence of a new strain of SARS-CoV-2 coronavirus have become publicly available [13,14]. A comparative analysis with the nucleotide sequences of viruses of this family available in the database showed that the new coronavirus clusters with the CoV clade related to SARS [15]. The information available in the database on other studied virus strains made it possible to find areas of matching/similar nucleotide sequences, protein profile, and, consequently, to predict and determine the characteristics of the virus faster, to develop vaccines. Impressive advances in translational biomedical research have demonstrated the need to integrate microbial collection databases both with high-level technological infrastructures in genomics, proteomics and bioinformatics and with information systems for patients and disease registries. This will make it possible to connect the information obtained on the biological characteristics of microorganisms with clinical information and the specific pathology of the patient and will increase the quality of biomedical research [16,17].

Material and methods

The strains of microorganisms deposited in the RCM were taken as the studied material. DNA isolation was performed using the Qiagen QIAamp DNA MiniKit kit (Qiagen, Germany) according to the manufacturer's instructions [18]. Qualitative assessment was carried out in 1% agarose gel using a 1x TAE buffer, quantitative - by the spectrophotometric method (NanoDrop).

DNA amplification was carried out on a ProFlex thermal cycler from AppliedBiosystems according to the following protocol: the total volume of the reaction mixture was 30 µl, of which water for PCR free of nucleases was 17 µl, 10x buffer with KCl was 2.5 µl, 15 mm MgCl2 was 2.5 µl, 2 mM dNTP Mix was 2.5 µl, 20 pmol/µl direct and reverse primers of 1.5 µl, 5 units /µl Taq polymerase – 0.5 µl, 20 ng/µl DNA matrix – 2 µl. Temperature regime: denaturation 95 ° C – 3 min, 30 cycles of denaturation, annealing and elongation 95 ° C - 30 sec, 60 ° C – 30 sec, 72 °C - 30 sec, the final stage of elongation 72 ° C – 5 min and storage at 4 ° C - ∞.

To compare genomes, the BLAST program (Basic Local Alignment Search Tool) was used according to the instructions of the NCBI website and the software Phylophlan 3.0.2 (https://pypi.org/project/PhyloPhlAn).

Results

The Republican Collection of Microorganisms (RCM) is a depository of collectible strains of industrial microorganisms intended for research and organization of biotechnological production (Decree of the Government of the Republic of Kazakhstan dated July 30, 2002 No. 850 "On the Republican collection of microorganisms").The Biobank of industrial microorganisms of the RCM contains more than 700 strains from various taxonomic groups (bacteria, actinomycetes, yeast, mycelial fungi), which are constantly replenished with new industrially valuable crops, selected from various regions of Kazakhstan and neighboring countries.

In the RCM, general information about the strains stored in the collection is presented on the official web portal of the organization (Figure1). However, the existing database requires further improvement, in particular, reflection of the genetic characteristics of collection cultures in it.

Therefore, within the framework of the scientific and technical program "Creation and replenishment of the collection of industrially valuable microorganisms, study and conservation of their biological diversity for the needs of biotechnology, medicine and agriculture" for 2021-22. KN of the Ministry of Education and Science of the Republic of Kazakhstan in the RCM, work is being carried out to create a database containing the genetic characteristics of deposited strains of microorganisms.

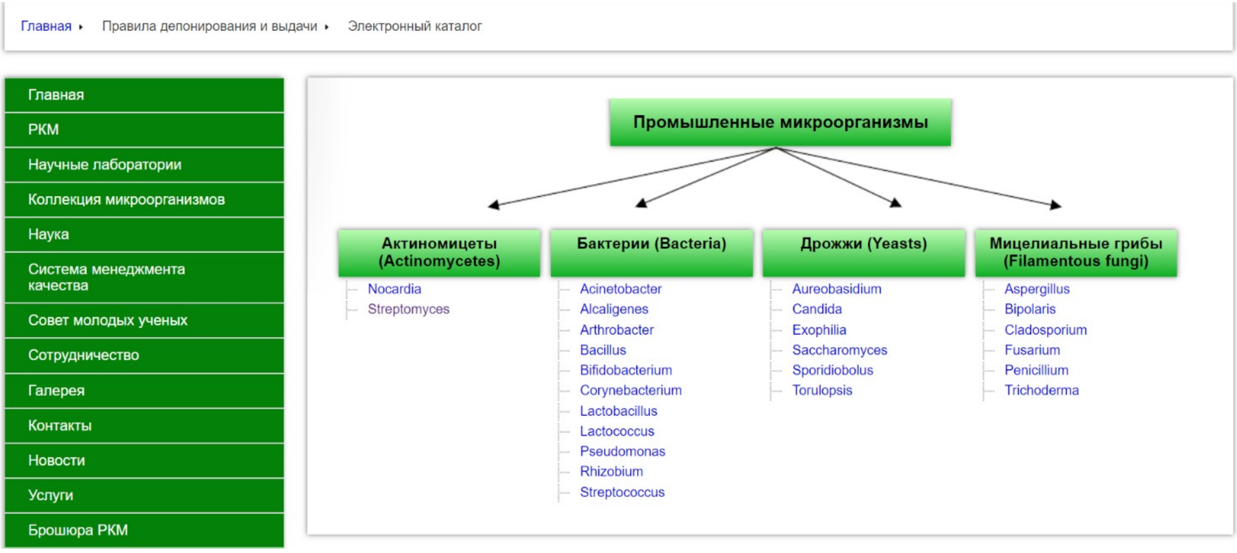


Figure 1. Structure of the electronic catalog of the RCM (2014)

To identify microorganisms, the necessary primers were selected from conservative sites of 16S rRNA and 18S rRNA. The sequence of selected primers is the most universal, used to identify the genomes of many prokaryotic and eukaryotic microorganisms (Table 1).

Table 1. Primers for identification of microorganisms 16S rRNA and 18S rRNA

Name	Oligonucleotide sequence	Concentration, [pmol/µl]
1F-16S	AGAGTTTGATCCTGGCTCAG	405,23

2R-16S	GGTTACCTTGTTACGACTT	344,19
3F-16S	TCCTACGGGAGGCAGCA	233,14
4R-16S	GACACGAGCTGACGACA	218,65
5R-16S	GGACTACCAGGGTATCTAAT	341,03
ITS1F	CTT GGT CAT TTA GAG GAA GTA A	235,82
ITS2R	GCT GCG TTC TTC ATC GAT GC	287,52
ITS1	TCC gTA ggT gAA CCT gCg g	377,36
ITS3	GCA TCG ATG AAG AAC GCA GC	263,18
ITS4	TCC TCC gCT TAT TgA TAT gC	293,11
ITS5	ggA AgT AAA AgT CgT AAC AAg g	361,70

Genotyping of collection strains is being carried out, the nucleotide sequence of some of them is shown in Table 2.

Table 2. Nucleotide sequences of collection strains

№	Name of the strain	Collection number	The nucleotide sequence
1	Proteus vulgaris 177	B-RKM 0038	CCATGCCGCGTGTATGAAGAAGGCCTTAGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGTGATAARGTTAATACCYTTRTCAATTGACGTTACCCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCAATTAAGTCAGATGTGAAAGCCCCGAGCTTAACTTGGGAATTGCATCTGAAACTGGTTGGCTAGAGTCTT

			GTAGAGGGGGGTAGAATTCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAATACCGGTGGCGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCTGTAAACGATGTCGATTTAGAGGTTGTGGTCTTGAACCGTGGCTTCTGGAGCTAACGCGTTAAATCGACCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGATGCAACGCGAAGAACCTTACCTACTCTTGACATCCAGCGAATCCTTTAGAGATAGAGGAGTGCCTTCGGGAACGCTGAGACAGGTGCTGCATGGCTGTCG
2	Staphylococcus aureus 6538	B-RKM 0039	ACGCCGCGTGAGTGATGAAGGTCTTCGGATCGTAAAACTCTGTTATTAGGGAAGAACATATGTGTAAGTAACTGTGCACATCTTGACGGTACCTAATCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATTATTGGGCGTAAAGCGCGCGTAGGCGGTTTTTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGAAAACTTGAGTGCAGAAGAGGAAAGTGGAATTCCATGTGTAGCGGTGAAATGCGCAGAGATATGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGTCTGTAACTGACGCTGATGTGCGAAAGCGTGGGGATCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGACCGCAAGGTTGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAAATCTTGACATCCTTTGACAACTCTAGAGATAGAGCYTTYCCCTTCGGGGGACAAAGTGACAGGTGGTGCATGGTTGTCGTC
3	Serratia marcescens 221 F	B-RKM 0059	RCCATGCCGCGTGTGTGAAGAAGGCCTTCGGGTTGTAAAGCACTTTCAGCGAGGAGGAAGGTGGTGARCTTAATACGYTCATCAATTGACGTTACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGTCAGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATTTGAAACTGGCAAGCTAGAGTCTC

			GTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCTGGACGAAGACTGACGCTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCTGTAAACGATGTCGATTTGGAGGTTGTGCCCTTGAGGCGTGGCTTCCGGAGCTAACGCGTTAAATCGACCGCCTGGGGAGTACGGCCGCAA GGTTAAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGATGCAACGCGAAGAACCTTACCTACTCTTGACATCCAGAGAACTTASCAGAGATGCTTTGGTGCCTTCGGGAACTCTGAGACAGGTGCTGCATGGCTGTCGTCAGCYTCGT
4	Bacillus thuringiensis kurstaki Z-52V3	B-RKM 0210	AGCCCTATCTCTAGGGTTTTCAGAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCCTTGCGGCCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAACTTCAGCACTAARGGGCGGAAACCCTCTAACACTTAGCACTCATCGTTTACGGCGTGGACTACCAGGGTATCTAAWCCTGTTTGCTCCCCACGCTTTCGCGCCTCAGTGTCAGTTACAGACCAGAAAGTCGCCTTCGCCACTGGTGTTCCTCCATATCTCTACSCATTTCACCGCTACACATGGAATTCCACTTTCCTCTTCTGCACTCAAGTCTCCCAGTTTCCAATGACCCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACTTAAGAAACCACCTGCGCGCGCTTTACGCCCAATAATTCCGGATAACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTTCTGGTTAGGTACCGTCAAGGTGCCAGCTTATTCAACTAGCACTTGTTCTTCCCTAACAACAGAGTTT
5	Bacillus cereus №4	B-RKM 0492	ATCTCTAGGGTTTTCAGAGGATGTSAAGACCTGGKAAGGKTTGGAGCGTGCWTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCCTTGCGGCCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAACTTCAGCACTAAAGGGCGGAAACCCTCTAACACTTAGCACTCATCGTTTACGGCGTGGACTACCAGGGTATCTAATCCTGTTTGCTCCCCACGCTTTCGCGCCTCAGTGTCAGTTACAGACMAGAAAGTCGCCTTCGCCACTGGKGTTCCTCCATATCTCTACGCATTTCACCGCTACACATG

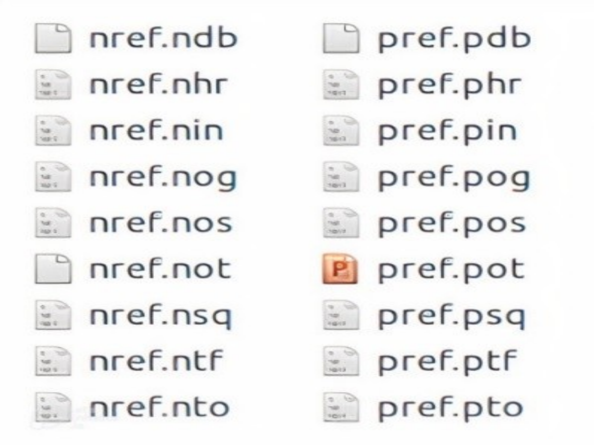
			KAATTCCACTTTCCTCTTCTGCACTCAAGTCTCCCAGTTTCCAATGACCCTYYATGGTTKAGCCGTGGGCTTTCACATCASACTTAAGAAACCACCTGCGCGCGCTTTACGCCCAATAATTCCGGATAACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTTCTGGTTAGGTACCGTCAAGGTGCCAGCTTATTCAACTAGCACTTGTTCTTCCCTAACAACAGAGTTTTACGACCCGAAAGCCTTCATCACTCACGCGGCGTTGCTCCSTCAGACTTTCGTCCATTGCGGAAGATTCCCTACTGCTGCCTCCCCGTAGG
6	Lactobacillus plantarum ATCC-9338	B-RKM 0019	GCCGCGTGAGTGAAGAAGGGTTTCGGCTCGTAAAACTCTGTTGTTAAAGAAGAACATATCTGAGAGTAACTGTTCAGGTATTGACGGTATTTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTTTTAAGTCTGATGTGAAAGCCTTCGGCTCAACCGAAGAAGTGCATCGGAAACTGGGAAACTTGAGTGCAGAAGAGGACAGTGGAACTCCATGTGTAGCGGTGAAATGCGTAGATATATGGAAGAACACCAGTGGCGAAGGCGGCTGTCTGGTCTGTAACTGACGCTGAGGCTCGAAAGTATGGGTAGCAAACAGGATTAGATACCCTGGTAGTCCATACCGTAAACGATGAATGCTAAGTGTTGGAGGGTTTCCGCCCTTCAGTGCTGCAGCTAACGCATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCTACGCGAAGAACCTTACCAGGTCTTGACATACTATGCAAATCTAAGAGATTAGACGTTCCCTTCGGGGACATGGATACAGGTGGTGCATGGTTG
7	Lactobacillus helveticus 5-2M	B-RKM 0807	CGCGTGAGTGAAGAAGGGTTTCGGCTCGTAAAACTCTGTTGTTAAAGAAGAACACCTTTGAGAGTAACTGTTCMAGGGTTGACGGTATTTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTTTTAAGTCTGATGTGAAAGCCTTCGGCTTAACCGGAGAAGTGCATCGGAAACTGGGAGACTTGAGTGCAGAAGAGGACAGTGGAACTCCATGTGTAGCGGTGGAATGCGTAGATATATGGAAGAACACCWGTGGCGAAGGCGGCTGTCTWGYCTGTAACTGACGCTGAGGCTCGAAAGCATGGGTAGCGAACAGGATTAGATACCCTG

			GTAGTCCATGCCGTAAACGATGAGWGCWAAGTGTTGGAGGGTTTCCGCCCTTCAGWGCTGCAGCTAACGCWTTAAGCACTCCGCCTGGGGAGTACGACCGCAAGGTTGA AACTCARAGKAATTGACGGGGGCCCGCACAAGCKGKGGAGCATGTGGTTTAATTCGWAGCT
8	Lactobacillus paracasei 8-2M	B-RKM 0810	GTCAAGACCTGGTAAGGTTCTTCGCGTWGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAACCTTGCGGTCGTACTCCCCAGGCGGARTGCTTAATGCGTTAGCTGCRGCACTGAAGGGSGGAAACCCTCCAACACYTAGCAYTCATCGTTTACGGCATGGACTACCAGGGTATCTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTACAGACYAGACAGCCGCCTTCGCCACTGGTGTTCTTCCATATATCTACGCATTYCACCGCTACACATGGAGTTCCACTGTCCTCTTCTGCACTCAAGTYTCCCAGTTTCCGATGCRCTTCYYCGGTTAAGCCGARGGCTTTCACATCAGACTTAAAAAACCGCCTGCGCTCGCTTTACGCCCAATAAATCCGGAYAACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTTCTGGTTRRATACCGTCAMSCCTASAACAGTTACTCTSMCGASYGTTCTTCTYTAACAACAGAGTTTTACGASCCGAAASCCTTCTTCACTCACGCGGCRTTGCTCCATCAGACTTKCGTCCATTGTGGAAGATTCCCTACTGCTGCCTCCCGTAGGAA

The BLAST program (Basic Local Alignment Search Tool) was used to compare genomes. The BLAST+ software package (Basic Local Alignment Search Tool) is installed according to the NCBI instructions (https://www.ncbi.nlm.nih.gov/books/NBK52640) [19].

The genome of the test microorganism was used to test the operation of the installed programs. Corynebacterium glutamicum was chosen for this purpose (https://www.ncbi.nlm.nih.gov/datasets/genomes/taxon=1718&utm_source=gquery&utm_medium=referral&utm_campaign=KnownItemSensor:taxname) [20].

Using the makeblastdb command, two databases were created called nref (for comparing nucleotide sequences) and pref (for comparing protein sequences) from the reference genome (GCF_000742715.1). All files necessary for the correct operation of the database were automatically created (Figure 2).



nref.ndb
nref.nhr
nref.nin
nref.nog
nref.nos
nref.not
nref.nsq
nref.ntf
nref.nto

pref.pdb
pref.phr
pref.pin
pref.pog
pref.pos
pref.pot
pref.psq
pref.ptf
pref.pto

Figure 2 - Reference database files

With the help of the blastn command, the test genome of KB (GCF_000742735.1) was compared with the nref database. blastn allows you to compare the sequence relative to the database in the nucleotide-nucleotide mode (Figure 3). As a result, a text file is created that contains a pairwise comparison of two nucleotide sequences.

```
BLASTN 2.12.0+

Reference: Zheng Zhang, Scott Schwartz, Lukas Wagner, and Webb Miller (2000),
"A greedy algorithm for aligning DNA sequences", J Comput Biol 2000; 7(1-
2):203-14.

Database:     data/GCF_000742715.1/GCF_000742715.1_ASM74271v1_genomic.
fna  2 sequences; 3,192,886 total letters


Query= NZ_CP007724.1 Corynebacterium glutamicum strain AR1 chromosome,
complete genome

Length=3145677

Sequences          producing          significant          alignments:
Score          E

                                                                   (Bits
)          Value

NZ_CP007722.1  Corynebacterium  glutamicum  strain  ATCC  21831  chromo...
2.027e+06    0.0

NZ_CP007723.1  Corynebacterium  glutamicum  strain  ATCC  21831  plasmi...
734            0.0
```

```
bNZ_CP007722.1 Corynebacterium glutamicum strain ATCC 21831 chromosome,
complete genome

Length=3176076

        Score = 2.027e+06 bits (1097615), Expect = 0.0

        Identities = 1097936/1098096 (99%), Gaps = 1/1098096 (0%)

        Strand=Plus/Plus

Query                                                              431875
CGAAGGTGTCCCAGGAGCCGCGGAGGTCATCCTGCCACCAGCCCCAGG
AGTGGGTGCCAG    431934

                | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | |

Sbjct                                                              462691
CGAAGGTGTCCCAGGAGCCGCGGAGGTCATCCTGCCACCAGCCCCAGG
AGTGGGTGCCAG    462750


Query                                                              431935
TGGGGCGCAGGTTCCAGTCAGCAGGGATGTCGGCACTGTCTAGCTTGG
CATTGAGATCGT      43199

               | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
        | | | | | | | | | | | | | | | | | | | |

Sbjct  462751 TGGGGCGCAGGTTCCAGTCAGCAGGGATGTCGGCACTGT
CTAGCTTGGCATTGAGATCGT      462810
```

Figure 3 - Results of nucleotide - nucleotide comparison

Preparatory work has been carried out on the construction of a phylogenetic tree. Laravel 8.0 was chosen as the software platform. This web framework is a free open source application designed for development using the MVC architectural model. Laravel products are manufactured under the MIT license (Massachusetts Institute of Technology, USA). The framework software platform is optimal for use at the corporate level. In addition, there is support for revision and modernization by developers. The search for information inside the site will be carried out by the name of the microorganism, genetic sequence and other parameters. Search filtering is in development.The database of deposited strains itself is formed on the basis of the MySQL program distributed as free software (users have the right to unlimited installation, launch and free use).

Along with genotyping of deposited strains, the RCM database will include information about their phylogeny. The deposit and issuance of collectible cultures will be carried out in digital format: requests are received electronically, the user can find the necessary strains on the website directly in the database of collectible strains, where all the necessary service information is available. Genotyping is mandatory for newly deposited strains, and the nucleotide sequences of the strain provided by the applicants will undergo mandatory verification in the depository.

Discussion

The research and methodological approaches to the introduction of the BLAST software package into the electronic resource of the RCM are presented. If the electronic catalog of collection cultures previously created on the RCM website covered general characteristics related to their taxonomic affiliation, target biological activity, then in the database being formed, along with phenotypic characteristics, the results of genotyping of the deposited strain are mandatory.

The created electronic database will subsequently expand information about the biological characteristics of collection cultures, introduce the results of the proteomic profile of deposited strains, including adhesiveness peptides or bacteriocins of lactic acid bacteria, specific proteins of other taxonomic groups of industrial microorganisms.

Conclusion

In the RCM, long-term storage of strains of industrial microorganisms is carried out, in compliance with the following basic principles: purity of the isolated culture of the microorganism; stability of the biological properties of the isolated strain, preservation of its authenticity in conditions of long-term storage; availability of detailed information about the phenotypic and genotypic characteristics of the strain.

The creation of a database of genomic sequences of microorganisms based on the BLAST computer program (nucleotide group) contributes to improving the quality of service function and safety of collection cultures deposited in the RCM.The database being developed will make it easier for the applicant to obtain the necessary information about the strain, simplify the procedures for depositing and issuing collection cultures.

Source of funding

This work was carried out within the framework of the project of program-targeted financing of the Ministry of Education and Science of the Republic of Kazakhstan (No. OR 11465530)".

References

1 Primak-R. Osnovy sokhraneniya [Text] / M: MGU, -2002. – P. 256.

2 Çaktü, K.; Türkoˇglu, E.A. Microbial culture collections: The essential resources for life [Text] / Gazi Univ. J. Sci. –2011. -№24. -P.175–180.

3 Stackebrandt, E. Deposit of microbial strains in public service collections as part of the publication process to underpin good practice in science [Text] / Smith, D., Casaregola, S., Varese, G.C., Verkleij, G., Lima, N., Bridge, P. / Springerplus, -2014. -Vol.3. –P. 1–4.

4 Luciana De Vero, Maria Beatrice Boniotti, Marilena Budroni Preservation, Characterization and Exploitation of Microbial Biodiversity: The Perspective of the Italian Network of Culture Collections [Text] / Microorganisms, – 2019. -№7. – P. 685. doi:10.3390/microorganisms7120685.

5 Bajerski, F., Nagel, M. & Overmann, J. Microbial occurrence in liquid nitrogen storage tanks: a challenge for cryobanking [Text] / Appl Microbiol Biotechnol 105, -2021. -P. 7635 – 7650. https://doi.org/10.1007/s00253-021-11531-4.

6 Sarmurzina Z. S. Kollekcii kultur mikroorganizmov dlya ustojchivogo sokhraneniya mikrobnogo bioraznoobraziya I geneticheskikh resursov. [Collections of cultures of microorganisms for the sustainable conservation of microbial biodiversity and genetic resources] [Text] / Proceedings of the International Scientific-Practical Conference "Actual Problems of Biodiversity and Biotechnology", dedicated to the Year of Youth in the Republic of Kazakhstan. -Nur-Sultan, -2019. – P.10-11.

7 Iskakova A. N., Nagumanova G. S., Eskaraevaa A. A., Kulametov Zh. A., Sarmurzina Z. S. Sozdanie otechestvennoj annotirovannoj bazy dannykh mikroorganizmov aktualnost I perspektivy razvitiya [Creation of the domestic annotated database of microorganisms: relevance and prospects for development] [Text] / Proceedings of the International Scientific and Practical Conference "Actual problems of microbiology, biotechnology and biodiversity", dedicated to the 30th anniversary of Independence of the Republic of Kazakhstan] / Nur-Sultan, -2021. – P. 183-188.

8 Obzornoe issledovanie mikroorganizmov, imeiushchikh znachenie dlia proizvodstva prodovolstviia i vedeniia selskogo khoziaistva / Prodovolstvennaia i

selskokhoziaistvennaia Organizatsiia Obieedinennykh Natsii, Dvenadtsataia ocherednaia sessiia. [An overview study of microorganisms of importance for food and agriculture [Text] / Food and Agriculture Organization of the United Nations, Twelfth Regular Session Rim, 19 – 23 oktiabria 2009 goda. –P.10.

9 Brovarnik V.V., Golovach T.M. Baza dannykh dlia depozitariia kultur mikroorganizmov [Text] / Mikrobiol. Zhurn. Database for microbial culture depository / Mikrobiol. Zhurnal, –2015. -V.77. -№5. -P.81-86.

10 Matthew J. Towards a unified data infrastructure to support European and global microbiome research: a call to action [Text] / Ryan, Michael Schloter, Gabriele Berg, Linda L. Kinkel, Kellye Eversole, James A. Macklin, Daria Rybakova and Angela Sessitsch / Environmental Microbiology, -2020. doi:10.1111/1462-2920.15323.

11 Buikin S.V., Bragina E.Iu., Koneva L.A., Puzyrev V.P. Bazy dannykh kollektsii biologicheskogo materiala: organizatsiia soprovoditelnoi informatsii [Text] / Biulleten sibirskoi meditsiny. Databases of biological material collections: organization of accompanying information / Bulletin of Siberian Medicine, –2012. - №1. –P.111-121.

12 Altschul,S. Basic local alignment search tool Affiliations expand [Text] / / J. Mol. Biol . –1990. –Vol.215. -№.3. – P. 403-10.

doi: 10.1016/S0022-2836(05)80360-2.

13 Holmes, E. Initial genome release of novel coronavirus 2020 [Text] / Available from: http://virological.org/t/initial-genome-release-of-novel-coronavirus/319.

14 GenBank. Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome 2020 [16 January 2020], [Text] / Available from: https://www.ncbi.nlm.nih.gov/nuccore/ MN908947.

15 Outbreak of acute respiratory syndrome associated with a novel coronavirus, China; First cases imported in the EU/EEA; second update [Text] / -European Centre for Disease Prevention and Control (ECDC). - 26 January, -2020.

16 De Paoli, P. Future of biobanking in microbiology for medical research [Text] / Futur microbiology, -2008. -Vol.3. -№1. –P.79-86. DOI: 10.2217/17460913.3. 1.79.

17 Hawkins, A.K., O'Doherty, K.C. "Who owns your poop?": insights regarding the intersection of human microbiome research and the ELSI aspects of biobanking and related studies [Text] / BMC medical genomics, -2011. -Vol.4. -№7(72). DOI: 10.1186/1755-8794-4-72.

18 Tolosa, J.M., Column-based method to simultaneously extract DNA, RNA, and proteins from the same sample [Text] / Schjenken, J.E., Civiti, T.D., Clifton, V.L., Smith, R. / Biotechniques, -2007. –Vol.43(6). –P. 799-804.

19 BLAST+ (Basic Local Alignment Search Tool) software package: https://www.ncbi.nlm.nih.gov/books/NBK52640 (accessed 18.09.2021).

20 Genome of the microorganism Corynebacterium glutamicum in bazedan NCBI[Text] https://www.ncbi.nlm.nih.gov/datasets/genomes/taxon=1718&utm_source=gquery&utm_medium=referral&utm_campaign=KnownItemSensor:taxname.

МИКРООРГАНИЗМДЕР ЖИНАҒЫ – БИОРЕСУРСТАР ЖӘНЕ МӘЛІМЕТТЕР БАЗАСЫ

Алмағамбетов Қайыртай Хамитұлы
Медицина ғылымдарының докторы, профессор
Микроорганизмдердің республикалық коллекциясы
Нұр-Сұлтан қ., Қазақстан
E-mail: rcmkz@list.ru
Нағыманова Гүлбахыт Сманқызы
Биология ғылымдарының магистрі
Микроорганизмдердің республикалық коллекциясы
Нұр -Сұлтан қ., Қазақстан
E-mail: gulbakyt_n@mail.ru

Темірханов Аслан Жаңаұлы
Биология ғылымдарының кандидаты
Микроорганизмдердің республикалық коллекциясы
Нұр -Сұлтан қ., Қазақстан
E-mail: rkm13@mail.ru

Сармурзина Зинигуль Серікқызы
Биология ғылымдарының кандидаты
Микроорганизмдердің республикалық коллекциясы
Нұр -Сұлтан қ., Қазақстан
E-mail: sarmurzina@list.ru

Түйін

Республикалық микроорганизмдер коллекциясы (РМК) – ғылыми-зерттеу және талдау жұмыстарын жүргізуге, биотехнологиялық өндірісті ұйымдастыруға арналған өндірістік микроорганизмдердің депозитарийі. РКМ Өндірістік микроорганизмдер биобанкінде әртүрлі таксономиялық топтардан (бактериялар, актиномицеттер, ашытқылар, жіп тәрізді саңырауқұлақтар) 700-ден астам штамдар сақталады. Қазақстанда өнеркәсіптік микроорганизмдердің коллекциялық штаммдарының, оның ішінде олардың фено- және генотиптік сипаттамалары туралы мәліметтердің бірыңғай аннотацияланған деректер базасы (БҚ) жоқ. Осыған байланысты РҚМ өнеркәсіптік микроорганизмдердің тұндырылған штаммдарының генетикалық сипаттамаларын қамтитын мәліметтер базасын құру бойынша жұмыс жүргізуде. Геномдарды салыстыру үшін Ұлттық биотехнологиялық ақпарат орталығының (NCBI) нұсқауларына сәйкес BLAST+ (Basic Local Alignment Search Tool) бағдарламалық пакеті пайдаланылды. BLAST компьютерлік бағдарламасы (нуклеотидтер тобы) негізінде микроорганизмдердің геномдық реттілігінің деректер қорын құру және ақпаратты ӨКМ Өнеркәсіптік микроорганизмдер биобанкінде сақтау қызмет көрсету функциясының сапасын және коллекциялық дақылдардың қауіпсіздігін арттыруға мүмкіндік береді.

Кілт сөздер: микроорганизмдер коллекциясы; биобанк; биоресурстар; мәліметтер базасы; BLAST.

НАБОР МИКРООРГАНИЗМОВ – БИОРЕСУРСЫ И БАЗЫ ДАННЫХ

Алмагамбетов Кайртай Хамитович
Доктор медицинских наук, профессор
Республиканская коллекция микроорганизмов
г. Нур-Султан, Казахстан
E-mail: rcmkz@list.ru

Нагуманова Гульбахыт Смановна
Магистр биологических наук
Республиканская коллекция микроорганизмов
г. Нур-Султан, Казахстан
E-mail: gulbakyt_n@mail.ru

Темирханов Аслан Жанаевич
Кандидат биологических наук
Республиканская коллекция микроорганизмов
г. Нур-Султан, Казахстан
E-mail: rkm13@mail.ru

Сармурзина Зинигуль Сериковна
Кандидат биологических наук
Республиканская коллекция микроорганизмов

г. Нур-Султан, Казахстан
E-mail: sarmurzina@list.ru

Аннотация

Республиканская коллекция микроорганизмов (РКМ) является депозитарием промышленных микроорганизмов, предназначенных для проведения научно-исследовательских и аналитических работ, для организации биотехнологического производства. В Биобанке промышленных микроорганизмов РКМ хранятся более 700 штаммов из различных таксономических групп (бактерии, актиномицеты, дрожжи, мицелиальные грибы). В Казахстане нет единой аннотированной базы данных (БД) коллекционных штаммов промышленных микроорганизмов, включающей сведения об их фено- и генотипических характеристиках. В этой связи в РКМ выполняются работы по созданию БД, содержащей генетические характеристики депонированных штаммов промышленных микроорганизмов. Для сравнения геномов используется программный пакет BLAST+ (Basic Local Alignment Search Tool) согласно инструкции National Centre for Biotechnological Information (NCBI). Создание БД геномных последовательностей микроорганизмов на основе компьютерной программы BLAST (нуклеотидная группа) и хранение информации в Биобанке промышленных микроорганизмов РКМ будет способствовать повышению качества сервисной функции и безопасности коллекционных культур.

Ключевые слова: коллекция микроорганизмов; биобанк; биоресурсы; база данных; BLAST.