

Mikrobiologické báze dat a databanky pro biotechnologie

663 579

Ing. RICHARD BUREŠ, CSc., Mikrobiologický ústav ČSAV, Praha

Klíčová slova: *biotechnologie, informační systém, databanka, mikroorganismus*

Mezi důležité zdroje informací, které podmiňují rozvoj biotechnologie, patří informace o vlastnostech průmyslově využitelných kmenů různých druhů mikroorganismů. Tyto informace ve své původní a úplné formě jsou obsaženy v záznamech různých více či méně specializovaných pracovišť, která shromažďují, uchovávají, testují a na požádání zasílají vybrané kmeny mikroorganismů. Takto získané kmeny mohou sloužit jednak jako modelové systémy, dále pro další šlechtění a genové manipulace a nebo jako přímo průmyslově využitelné kmeny pro produkci určité látky požadované kvality a v požadovaném množství. (V těchto případech jde o patentované kmeny uložené ve sbírkách, které splňují podmínky tzv. Budapeštské dohody o ukládání patentovaných kmenů, z roku 1976).

Hlavním informačním úkolem je lokalizovat pracoviště (sbírku), které vlastní kmen požadovaných vlastností. K tomu slouží různé typy katalogů a jiných více či méně podrobných popisů uložených kmenů. Problém je v tom, že ve formě spolehlivého katalogu je zpracována jen menší část mikrobiologického materiálu uloženého ve sbírkách a vhodného pro průmyslové použití.

Pro zvýšení organizovanosti a stupně standardizace v této oblasti se vytvářejí různé regionální i celosvětové organizace. Z nich nejvýznamnější je Světová federace sbírek kultur (WFCC) a Evropská organizace sbírek kultur (ECCO). Za jejich podpory a spoluúčasti řady dalších organizací, zejména z oblasti standardizace vědeckotechnických dat — Výboru pro vědeckotechnická data (CODATA), Organizace pro vzdělání, vědu a kulturu (UNESCO), EHS a řady dalších, vznikají v posledních letech projekty zaměřené na zvýšení dostupnosti existujících mikrobiologických dat za pomoci výpočetní techniky a mezinárodních přenosových telekomunikačních sítí.

Základní myšlenka při vytváření mikrobiologických

informačních systémů využívajících metody automatizovaného zpracování dat pomocí počítačů spočívá v tom, že místo sběru a zpracování těchto dat nemusí být nutně totožné s umístěním sbírky samotné. Vzniká potom jakýsi distribuovaný systém zpracování mikrobiologických informací, kdy kompletní data jsou v příslušné sbírce v tištěné formě nebo v paměti počítače, avšak stručná informace o jejich existenci a základním obsahu je přístupná uživatelům v centrálním databázovém středisku.

Takto organizované informační systémy je možno podle rozsahu údajů uložených v centrální databázi rozdělit do dvou skupin:

1. Informační systémy, které poskytují kromě základních údajů o lokalizaci určitého kmene, jména kurátora a adresy sbírky, ještě popis základních fyziologických a morfologických vlastností, způsobu kultivace, produkčních vlastností atp. Do této skupiny patří například informační systém o kvasinkách provozovaný Ústavem pro výzkum potravin (IFRN) v Norwich, Velká Británie, pod názvem COMPASS (tento systém poskytuje ještě možnost automatické identifikace neznámých kmenů kvasinek, automatické objednávání kmenů a využívání tzv. E-mail, tj. zasílání zpráv podobné telexu), a dále systém MICIS (Microbial Culture Information Service) zahrnující doposud data z 9 sbírek bakterií, kvasinek, řas atp. Provozovatelem tohoto systému je Laboratory of Government Chemist, Londýn. Tento systém umožňuje nalézt neznámý kmen požadovaných vlastností na základě údajů o substrátu, metabolitech, enzymatické aktivitě, citlivosti, průmyslových vlastnostech atp., nebo vytisknout poměrně rozsáhlou informaci o kmenu známém, definovaném jménem nebo číslem. Ukázka známých ze systému MICIS je uvedena v *tabulkách 1 a 2*.

2. Informační systémy obsahující pouze základní údaje

Tab. 1. Přehled skupin informací o vybraných kmenech mikroorganismů v systému MICIS

1. Nomenclature	11. Bacteria Morphology Colony
2. Synonyms	12. Bacteria Morphology Liquid
3. Growth and Preservation	13. Fungal Morphology
4. Special Applications	14. Yeast Morphology
5. Sensitivity and Tolerance	15. Algae and Protozoa Morphology
6. Substrate Utilisation	16. Source and History
7. Industrial Properties	17. Literature References
8. Enzymes	18. Request Printout
9. Metabolites Produced	19. Enter New Reference Number
10. Bacteria Morphology Cell	

Tab. 2. Ukázka obsahu části záznamu o vybraném kmenu [BRUCELLA ABORTUS] v systému MICIS

Nomenclature		Growth and Preservation	
Name BRUCELLA ABORTUS Reference NCTC 10093 Genus BRUCELLA Species ABORTUS Authority Subclassification Category BACTERIUM		Gaseous requirements CARBON DIOXIDE REQUIRED Lighting conditions pH range To Optimum 7.0 Temperature range To Optimum 37 Culture media 5 % HORSE BLOOD AGAR Maintenance: Long term FREEZE-DRYING Short term Unsuitable	
Date last amended 86/01/27			
Substrate utilisation			
Substrate	NITRIC ACID		
Type of reaction	REDUCTION	Result	POSITIVE
Substrate	GLUCONIC ACID		
Type of reaction	OXIDATION	Result	NEGATIVE
Substrate	SALICIN		
Type of reaction	ANAEROBIC FERMENTATION	Result	NEGATIVE
Substrate	SORBITOL		
Type of reaction	ANAEROBIC FERMENTATION	Result	NEGATIVE
Substrate	MANNITOL		
Type of reaction	ANAEROBIC FERMENTATION	Result	NEGATIVE
Substrate	RAFFINOSE		
Type of reaction	ANAEROBIC FERMENTATION	Result	NEGATIVE
Enzymes produced			
Enzyme		Result	
CASEINASE		NEGATIVE	
GELATINASE		NEGATIVE	
CYTOCHROME C OXIDASE		POSITIVE	
ORNITHINE DECARBOXYLASE		NEGATIVE	
LYSINE DECARBOXYLASE		NEGATIVE	
ARGININE DEIMINASE		NEGATIVE	
PHENYLALANINE AMINOTRANSFERASE		NEGATIVE	
A-AMYLASE		NEGATIVE	
CATALASE		POSITIVE	
UREASE		NEGATIVE	
Source and history			
Type Strain Y Depositors name PAYNE DJH Year 1960			
Previous history		Habitat type HUMAN	
Country of origin		Hazard status YES	
Habitat description			
Availability RESTRICTED			
Literature references			
Author MEYER ME and MORGAN WJB Journal INT J SYST BACT			
Volume 13 Year 1973 Pages 135 To			
Author SUBCOMMITTEE Journal INT BULL BACT NOMENC			
Volume TAXON 13 Year 1963 Pages 153 To			
Author Journal WHO TECH REP SER			
Volume 67 Year 1953 Pages 20 To			

o mikrobiologickém materiálu, který mají ve svých depozitářích zúčastněně sbírky. Výsledkem rešerše je informace o sbírce, která vlastní mikrobiologický materiál požadovaných vlastností, avšak nikoliv podrobnější popis vlastností jednotlivých kmenů. Do této skupiny patří především MSDN (Microbial Strain Data Network), kde se používá pro popis fyziologických a morfologických vlastností 6číslicový tzv. RKC kód. Tento systém vzniká na základě společného úsilí CODATA, World Federation of Culture Collection a International Union of Microbiological Societies i za spoluúčasti řady dalších nevládních organizací. Kódový systém je otevřený a bude se doplňovat podle potřeby o další údaje. V současné době je přístupný pouze vzorek báze dat a realizace celého projektu je plánována do roku 1989. Pro přístup k uloženým informacím se používá telekomunikační systém DIALCOM, který je řízen počítačem umístěným ve Washingtonu v USA. Ukázka záznamu o sbírce mikroorganismů v systému MSDN s využitím RKC kódu je uvedena v tabulce 3.

Tab. 3. Ukázka záznamu o sbírce mykobakterií v Pasteurově ústavu

CL MYCOBACTERIA DI Institut Pasteur AD Service de la Tuberculose et des Mycobacteries AD 25 Rue du Docteur Roux; 75724 Paris Cedex 15; France TE (16)306-1919 x3820 TLX PASTEUR 250609F Co Dr. Veronique Levy-Frebault ID DIALXOM 42:CDT0002						
003008	005026	015024	016030	016031	016181	016253
016262	016263	016267	016274	016276	016318	016321
016323	016332	016337	017015	017033	017035	017046
018006	020008	020018	024424	024425	024426	025309
025310	025312	026515	026524	028152	030170	030188
030190	030258	034123	034125	034135	034136	034137
034142	034143	040367				

Použité číselné kódy odpovídají charakteristikám dané skupiny mikroorganismů nebo jednotlivých kmenů. [Viz následující vybrané příklady]:
003008: Cells are rod-shaped (cylindrical).
003026: Longer axis of rod is less than twice the shorter axis (coccobacillary).
015024: Growth tends to be in serpentine, cord-like masses in which the cells show a parallel orientation.
016030: Colony surface is smooth.
016031: Colony surface is rough.
016253: Organism grows in media containing 1 ug/ml p-aminosalicylate (PAS).
016263: Organism grows in media containing 550.0 ug/ml hydroxylamine hydrochloride.
016265: Organism grows in media containing 500.0 ug/ml hydroxylamine hydrochloride.

Do této skupiny dále patří vývojový projekt financovaný v rámci rozvojových programů zemí EHS. Jde o evropskou síť mikrobiologických informací MINE (Microbial Information Network Europe). V tomto systému se bude používat tzv. minimální soubor dat (Minimum Data Set) obsahující pouze základní informace o kmenech typu katalogu. Uživatelé budou mít přístup k centrální databázi prostřednictvím národních uzlových bodů. V současné době jsou ustaveny takovéto národní pobočky ve Velké Británii, NSR, Belgii a Holandsku. Plánuje se napojení tohoto systému také na informační systémy patřící do první skupiny, tj. MICIS a COMPASS a řadu dalších.

V tomto výčtu nelze opomenout jeden z prvních informačních systémů o uložení mikroorganismů ve sbírkách, který je nyní soustředěn v Japonsku v Ústavu pro fyzikální a chemický výzkum (RIKEN) pod názvem World Data Centre (WDC). Uvedený systém obsahuje informace o více než 320 sbírkách na celém světě, a to ve formě základních údajů o zaměření sbírky, včetně seznamu rodů a druhů uložených mikroorganismů, doplněných o jména a adresy kurátorů.

Kromě zmíněných systémů informujících o kmenech mikroorganismů existují systémy obsahující informace o hybridomech (Hybridoma Data Bank) a databanky obsahující údaje o sekvencích nukleových kyselin a aminokyselin (např. GenBank a Protein Sequence Data-

bank). Zejména v oblast těchto posledně jmenovaných informačních systémů se v současné době vyvíjí úsilí o vzájemnou koordinaci a standardizaci.

Při vytváření zde popsaných mikrobiologických informačních systémů je nejdůležitějším úkolem shromažďování kvalitních vstupních informací. To není možné bez vzájemné spolupráce řady organizací schopných poskytovat spolehlivá data v požadovaném množství a formátu. Proto vznikají v jednotlivých státech národní střediska pro sběr a předávání mikrobiologických dat. Aktivní účast v tomto systému je základní podmínkou pro zajištění možnosti přístupu k centrálním informačním bankám dat. Proto bude důležité, zda se podaří propojit tyto záměry s řešením obdobné problematiky v rámci Komplexního programu vědeckotechnického pokroku (KP VTP) členských zemí RVHP. Za ČSSR se tato problematika zpracovává v Mikrobiologickém ústavu ČSAV a je rovněž součástí řešení jednoho hlavního úkolu státního vědeckotechnického programu rozvoje biotechnologie v 8. pětiletém plánu.

Literatura

- [1] ROGOSA, M., KRICHEVSKY, M. I., COLWELL, R. R.: Coding Microbiological Data for Computers, Springer-Verlag: New York, 1986
- [2] KRICHEVSKY, M. I., SUGAWARA, H., FABRICIUS, B. O.: Culture Collections as Information Sources for Biotechnology, Proceedings Conference: Biotechnology Information 86, Brighton, U. K. 22.—25. September 1986
- [3] BINARA Newsletter, č. 11, 1987
- [4] NCYC Computer Services Handbook
- [5] MICIS News, 1987, č. 1—7
- [6] Budapeštská smlouva o mezinárodním uznání uložení mikroorganismů k účelům patentového řízení. Překlad. Ústř. knihovna patentové literatury, Úřad pro vynálezy a objevy, Praha, 1978, řada C, č. 16
- [7] KOMAGATA, K.: Relocation of the World Data Centre, MIRCEN Journal 1987, 3, s. 337—342

Lektoroval doc. Ing. M. Kocúr, CSC.

Bureš, R.: Mikrobiologické báze dat a databanky pro biotechnologie. Kvas. prům., 34, 1988, č. 11, s. 331—333.

Příspěvek se zabývá otázkou vytváření informačních systémů o mikroorganismech, zejména ve forměází dat a databank. Popisuje funkci a zaměření projektů

mezinárodní spolupráce v rámci obdobných projektů členských zemí RVHP.

Буреш, Р.: Микробиологические базы данных и банки данных для биотехнологий. Квас. прум., 34, 1988, № 11, стр. 331—333.

Статья занимается вопросом создания систем информации о микроорганизмах, особенно в форме баз данных и банков данных. Описывает функцию и направление проектов международного сотрудничества в этой области, особенно МСДН, МИЦИС, МИНЕ и КОМПАСС. Также обсуждается возможность международного сотрудничества в рамках аналогичных проектов стран-членов СЭВ.

Bureš, R.: Microbiological Databases and Databanks for Biotechnology. Kvas. prům., 34, 1988, No. 11, pp. 331—333.

The article briefly discusses organization of international activity with the aim to create information systems on microorganism, especially in the form of databases and databanks.

Attention is focused on formation and main purpose of international projects in this field, with the brief description of MSDN, MICIS MINE and COMPASS. Also is mentioned importance of international cooperation in the framework of similar project within COMECON countries.

Bureš, R.: Mikrobiologische Datenbasen und Datenbanken für Biotechnologie. Kvas. prům., 34, 1988, Nr. 11, s. 331—333.

Der Beitrag beschäftigt sich mit der Frage der Aufbau von Informations-systems über Mikroorganismen, hauptsächlich in der Form von Datenbasen und Datenbanken. Beschrieben sind die Funktionen und Richtungen der Projekt der internationalen Zusammenarbeit auf diesem Gebiet, innsbesondere MSDN, MICIS, MINE und COMPASS.

Gleichzeitig ist auch die Notwendigkeit der internationaler Zusammenarbeit im Rahmen ähnlichen Projekte der Länder, Mitglieder der RWG diskutiert.