

УДК 579.61:616.34-008.6-078

Використання інформаційних технологій при ідентифікації мікроорганізмів

М. Д. Кац, С. А. Деркач, О. М. Давиденко, А. І. Носатенко,
І. А. Крилова, О. А. Бочаров
Інститут мікробіології та імунології ім. І. І. Мечникова
АМН України, Харків

Резюме

Побудовані математичні моделі шляхом розробки тест-комплексів та створена комп'ютерна програма для спрощеної ідентифікації ентеробактерій.

Ключові слова: ідентифікація мікроорганізмів, математичне моделювання, програмне забезпечення для мікробіології.

Клин. информат. и Телемед.
2006. Т.3. Вып.4. с.82–85

Вступ

Діагностика завжди є однією із найважливіших завдань медицини, а точність діагностики, в основному, визначає рівень розвитку науки в конкретні історичні періоди.

З виникненням електронно-обчислювальної техніки і розвитком прикладної математики розпочався бум робіт, пов'язаних з намаганням формалізувати процес діагностики за допомогою математичних моделей. Принципові труднощі моделювання «великих» систем (до яких відносяться і медично-діагностичні системи) привели до пошуку інших шляхів.

Однією з найбільш багаточисельних родин мікроорганізмів, історія вивчення якої нараховує більше 170 років, є родина ентеробактерій. Внутрішньовидова ідентифікація представників цієї родини є досить тривалою і важкою тому, що потребує постановки багатьох лабораторних тестів (згідно класифікації Берджі – 48 тестів), результати яких не однозначні і тому їх оцінка викликає певні труднощі [1].

Нами розроблена і експериментально перевірена універсальна інтелектуальна методологія вивчення складних об'єктів, в основі якої закладена ідея штучного інтелекту [2–4]. В цій методології усі етапи побудування математичної моделі (крім вибору переліку вхідних і вихідних перемінних), формалізовані та відтворюють об'єктивність виявлених залежностей.

Матеріали та методи

Особливістю методичного підходу при виконанні даної роботи є використання

принципово нового методу мозаїчного портрету (ММП), розробленого в рамках інтелектуальної технології вивчення і удосконалення складних систем (ІНТВСС) [6].

Практично всі недоліки відомих методів у тому числі і математичного моделювання, можна здолати у тому разі, якщо для ідентифікації використовувати не окремі тести, а тест-комплекси, специфічні для кожного мікроорганізму.

Виділення таких специфічних тест-комплексів за допомогою ММП проілюстровано на прикладі родини ентеробактерій.

У таблицю родини ентеробактерій внесені показники Берджі, кожний рядок якої містить результати тестів, а кожний стовпчик відповідає одному виду мікроорганізму.

Для створення математичної моделі вихідна таблиця Берджі перебудовується наступним чином:

- кожний рядок створеної таблиці описує бактерію за номером, відповідним її порядковому номеру в таблиці Берджі;
- в кожному стовпчику створеної таблиці наведені порядкові номери та результати відповідних тестів:

«-» – імовірність негативного результату 90–100%;

«+» – імовірність позитивного результату 90–100 %.

У випадку, коли результати тестів мають імовірність менш ніж 90 % ([–], d, [+]), вони кодуються крапкою. При побудові моделі враховуються тільки тести з оцінками (–) та (+). Специфічні для кожного мікроорганізму тест-комплекси визначаються наступним чином. Шукають такі сполучення тестів в першій по порядку таблиці бактерії, які зустрічаються тільки у неї і не зустрічаються ні в одному з рядків, відповідних іншим 114 бактеріям даної родини. Одержані сполучення тлумачать як специфічні тест-

комплекси для цього мікроорганізму. Приклад визначення тест-комплексів уже був опублікований нами раніше [5].

Результати

Оцінка інформативності тестів

Імовірний характер ідентифікації визначається, насамперед, неоднозначністю тестів для різних штамів одного і того ж мікроорганізму, тоді один і той же тест в різних випадках може приймати значення як (-) так і (+), і вже потім ураховується близькість властивостей деяких груп бактерій усереднені родини.

Якби тест для якоїсь бактерії зустрічався у 50% випадків з кодом (-) і у 50% з кодом (+), то інформативність даного тесту склала б 100%. Коли зустрічаються тести з кодами (-) і (+) від 0 до 50%, оцінка їх повторення відповідає інформативності даних тестів. Наприклад, якщо частота тесту з кодом (-) складає 35%, то інформативність цього тесту по даному коду також складає 35%. Якщо тести з кодом (-) і (+) зустрічаються від 50 до 100%, інформативність тесту визначається за формулою:

$I = (100 - a)\%$, де «а» – % частоти тесту з одним із кодів (-) чи (+). Наприклад, якщо тест з кодом (-) зустрічається у 75%, то його інформативність по даному коду складає $100\% - 75\% = 25\%$. Загальна інформативність тесту визначається по сумі інформативності кодів зі значенням (-) і кодів зі значенням (+).

У таблиці 1 наведений перелік з указівкою інформативності кожного з них рангу по інформативності.

Для збільшення точності ідентифікації використовується надмірна кількість тест-комплексів специфічних для конкретного мікроорганізму. У випадку, якщо який-небудь тест наявний переважно для даного мікроорганізму зі значенням (+), буде оцінений як (-), це призведе до зниження всіх специфічних для нього тест-комплексів, в який включено і «помилковий» тест, і до появи в опису деяких тест-комплексів інших бактерій. Оскільки вірогідність появи «помилкової» оцінки конкретного тесту мала (максимум 10%), то за рахунок використання надмірної кількості тест-комплексів (в більшість яких «провокаційний» тест не входить) забезпечується надійність ідентифікації.

Табл. 1. Інформативність тестів ідентифікації (згідно визначника Берджі) після їх кодування.

N п/п	Назва тесту	Інформативність тесту	Інформативність тесту
1	2	3	4
1	Фарбування по Граму (24 г)	0	(45-46)
2	Оксидаза	0	(45-46)
3	Утворення індолу	47,9	27
4	Проба з метиловим червоним	55,6	22
5	Реакція Фогес-Проскауера	66,9	17
6	Цитрат (середовище Сіменса)	60,0	21
7	Утворення H ₂ S	46,1	29
8	Гідроліз сечовини	25,2	40
9	Фенілаланіндезаминаза (24 години)	27,0	39
10	Лізиндекарбоксилаза	69,6	15
11	Аргининдігідролаза	37,7	35
12	Орнітиндекарбоксилаза	78,3	6
13	Рухомість	65,3	18
14	Гідроліз желатини (22 °С)	40,9	34
15	Зростання в присутності KCN	80,0	5
16	Використання малонату	74,8	11
17	Утворення кислоти з D-глюкози	3,5	44
18	Утворення газу з D-глюкози	77,4	7
19	Утворення кислоти з: D-адонітолу	34,7	36
20	L-арабінози	64,4	19
21	Гліцерилу	55,6	23
22	Дульцитолу	29,5	38
23	Міо-інозитолу	52,2	26
24	D-ксилози	63,6	20
25	Лактози	70,0	14
26	Мальтози	55,2	24
27	D-манітолу	46,9	26
28	D-маннози	21,7	41
29	Мелібіози	82,6	4
30	α-метіл-D-глюкозиду	44,3	30
31	L- рамнози	85,3	3
32	Рафінози	75,7	10
33	Саліцину	76,6	8
34	Сахарози	74,8	12
35	D-сорбітолу	86,0	2
36	Трегалози	53,1	25
37	Целлобіози	89,5	1
38	Мукату	68,7	16
39	Тартрат (середовище Джорданса)	43,4	33
40	Гідроліз ескуліну	76,5	9
41	Використання ацетату	43,5	(31 - 32)
42	Відновлення нітрату	30,4	37
43	Дезоксирибонуклеаза (25 °С)	17,4	42
44	Ліпаза	43,5	(31 - 32)
45	ОНРГ	71,3	13
46	Пігмент	-	-
47	Розташування джугіків	-	-
48	Утворення каталази (24 години)	7,0	43

Табл. 2. Ідентифікація *Rahnella N77* за комп'ютерною програмою з використанням усіх тестів згідно визначника Берджі.*

N п/п	Назва тесту	Оцінка тесту за класифікацією Берджі	Варіанти оцінки тестів позначених дослідником			
			I	II	III	IV
1	2	3	4	5	6	7
1	Фарбування по Граму (24 г)	-	-	-	-	-
2	Оксидаза	-	-	-	-	-
3	Утворення індолу	-	-	-	-	-
4	Проба з метиловим червоним	[+]	?	+	-	+
5	Реакція Фогес-Проскауера	+	+	+	+	+
6	Цитрат (середовище Сіменса)	+	+	+	+	+
7	Утворення H ₂ S	-	-	-	-	-
8	Гідроліз сечовини	-	-	-	-	-
9	Фенілаланіндезаминаза (24 години)	+	+	+	+	+
10	Лізиндекарбоксилаза	-	-	-	-	-
11	Аргининдигідролаза	-	-	-	-	-
12	Орнітиндекарбоксилаза	-	-	-	-	-
13	Рухомість	-	-	-	-	-
14	Гідроліз желатини (22 °C)	-	-	-	-	-
15	Зростання в присутності KCN	-	-	-	-	-
16	Використання малонату	+	+	+	+	+
17	Утворення кислоти з D-глюкози	+	+	+	+	+
18	Утворення газу з D-глюкози	+	+	+	+	+
19	Утворення кислоти з: D-адонітолу	-	-	-	-	-
20	L-арабінози	+	+	+	+	+
21	Гліцерилу	[-]	?	-	-	+
22	Дульцитолу	[+]	?	+	-	+
23	Міо-інозитолу	-	-	-	-	-
24	D-ксилози	+	+	+	+	+
25	Лактози	+	+	+	+	+
26	Мальтози	+	+	+	+	+
27	D-манітолу	+	+	+	+	+
28	D-маннози	+	+	+	+	+
29	Мелібіози	+	+	+	+	+
30	α-метіл-D-глюкозиду	-	-	-	-	-
31	L-рамнози	+	+	+	+	+
32	Рафінози	+	+	+	+	+
33	Саліцину	+	+	+	+	+
34	Сахарози	+	+	+	+	+
35	D-сорбітолу	+	+	+	+	+
36	Трегалози	+	+	+	+	+
37	Целлобіози	+	+	+	+	+
38	Мукату	d	?	+	-	+
39	Тартрат (середовище Джорданса)	-	-	-	-	-
40	Гідроліз ескуліну	+	+	+	+	+
41	Використання ацетату					
42	Відновлення нітрату	+	+	+	+	+
43	Дезоксирибонуклеаза (25 °C)	-	-	-	-	-
44	Ліпаза	-	-	-	-	-
45	ОНРГ	+	+	+	+	+
46	Пігмент	-	-	-	-	-
47	Розташування джутиків	p	?	?	?	?
48	Утворення каталази (24 години)	+	+	+	+	+
49	Окислення-бродиння	f	?	?	?	?
Висновок по результатах ідентифікації			=17т.к.	=26т.к.	=17т.к.	=26т.к.

Експериментальна перевірка ефективності використання створеної комп'ютерної програми

Було проведено визначення видової належності цілого ряду мікроорганізмів з урахуванням їх біохімічної активності, як традиційним візуальним методом, так і з використанням розробленої нами методики ідентифікації.

При внесенні в комп'ютер значень усіх сорока дев'яти тестів (згідно визначнику Берджі) з урахуванням усіх можливих варіантів комбінації їх значень («+», «-», [+], [-], d), ми отримували майже 100 % вірні висновки за кілька хвилин (як приклад табл. 2).

При використанні меншої кількості тестів для бактеріологічної ідентифікації мікроорганізмів, виділених нами від хворих та музейних штамів, що і має місце у роботі більшості лабораторій, результати були також інформативні для коректного підтвердження видової належності бактерій за допомогою комп'ютерної програми (як приклад – табл. 3).

Висновки

Таким чином, ідентифікація мікроорганізмів родини ентеробактерій за допомогою специфічних для кожної бактерії тест-комплексів, відібраних за результатами біохімічних тестів можлива. Очевидно, що у порівнянні з візуально-аналітичним методом (шляхом пошуку відповідних значень тестів, характерних для мікроорганізму, ідентифікація якого проводиться лікарем-лаборантом по сторінках визначника бактерій Берджі), запропонований нами метод комп'ютерної ідентифікації є значно спрощеним, надійним, позбавленим випадкових помилок, тобто, вискооефективним. Переваги такого способу стають особливо відчутні при отриманні нетипових значень біохімічних тестів та їх поєднання.

Отже побудування математичних моделей шляхом розробки тест-комплексів

*Примітка:

«?» – видова відповідність ідентифікованого мікроорганізму; т.к. – специфічні тест-комплекси; [-] – 21-25% штамів позитивні; d – 26-75% штамів позитивні; [+] – 76-89% штамів позитивні.

Табл. 3. Ідентифікація *Citrobacter freundii* за комп'ютерною програмою з використанням лише частини тестів.

N п/п	Тест	Оцінка тесту за класифікацією Берджі	Варіанти оцінки тестів, позначених дослідником			
			I	II	III	IV
1	2	3	4	5	6	7
1	Індол	-	-	-	-	-
2	Метилрот	+	+	+	+	+
3	Фогес-Проскауер	-	-	-	-	-
4	Цитрат Сімонса	+	+	+	+	+
5	H ₂ S (сірководень)	[+]	?	+	-	-
6	Сечовина	d	?	+	-	+
7	Феніланін	-	-	-	-	-
8	Лізин	-	-	-	-	-
9	Аргинин	d	?	+	-	+
10	Орнітин	[-]	?	+	-	-
11	Рухомість	-	-	-	-	-
12	Желатина	-	-	-	-	-
13	Малонат натрію	[-]	?	+	-	-
14	Газ із глюкози	+	+	+	+	+
15	Лактоза	d	?	+	-	-
16	Сахароза	d	?	+	-	-
17	Маніт	+	+	+	+	+
18	Дульцит	d	?	+	-	-
19	Саліцин	-	-	-	-	-
20	Адоніт	-	-	-	-	-
21	Інозіт	-	-	-	-	-
22	Сорбіт	+	+	+	+	+
23	Рафіноза	d	?	+	-	-
24	Рамноза	+	+	+	+	+
25	Арабіноза	+	+	+	+	+
Висновок по результатах ідентифікації			=1т.к.	=13т.к.	=7т.к.	=7т.к.

та створення пакетів комп'ютерних програм є основою для подальшого розвитку цього науково-практичного напрямку.

Література

1. Bergey's Manual of Systematic of Bacteriology. – 1986. – P. 13.
2. Kats M., Kestelman V., Davidenko A. Intelligent technology of complex systems study as a methodological background for development of high medical technologies // Scientific Izrael – Technological advantages. – Issues 1-2. 2004. – Vol.6. – №.55.
3. Mayorov O., Kats M. Intellectual Systems for Differential Diagnostics

within Groups of hardly distinguished Diseases //Medical in fobahn Europe Proceedings of MIE2000 and GMDS2000/ – P.180–184.

4. Кац М. Д. Использование искусственного интеллекта для разработки методов дифференциальной диагностики внутри групп трудно-различимых заболеваний //Клиническая информатика и телемедицина. – 2004. – № 1. – С.86–89.
5. Новый подход к идентификации микроорганизмов / Кац М. Д., Давиденко А. М., Деркач С. А., Носатенко А. И., Похил С. И., Крылова И. А., Строгая Е. И., Габышева Л. С. / Вісник Харківського національного університету ім. В. Н. Каразіна. – Медицина. – 2004. – Вип.8. – С.100–101.

6. Кац М. Д., Ардашев В. Н., Яковлев Г. М. Метод мозаичного портрета в прогнозировании осложненный инфаркта миокарда // Кардиология. – 1981. – № 6. – С. 40–44.

Use of information technologies for identification of microorganisms

M. Kats, S. Dercach, O. Davidenko, A. Nosatenco, I. Krilova, A. Bocharov
Mechnikov Institute of Microbiology of Ukraine Medical Sciences Academy, Kharkiv

Abstract

Mathematical model is constructed by means of elaboration of the tests-complexes and computer programme for simplified identification of Enterobacteria is created.

Key words: identification of microorganisms, mathematical modeling, software for microbiology.

Использование информационных технологий при идентификации микроорганизмов

М. Д. Кац, С. А. Деркач, О. М. Давиденко, А. И. Носатенко, И. А. Крилова, А. А. Бочаров
Институт микробиологии и иммунологии им. И. И. Мечникова АМН Украины, Харьков

Резюме

Построена математическая модель путем разработки тест-комплексов и создана компьютерная программа для упрощенной идентификации энтеробактерий.

Ключевые слова: идентификация микроорганизмов, математическое моделирование, программное обеспечение для микробиологии.

Переписка

д.т.н., профессор **М. Д. Кац**
ул. Донецкая, 37, кв. 24
Северодонецк, Луганская обл.
93400, Украина
e-mail: kats@zfs.lg.ua