



Systemy restrykcyjno-modyfikacyjne – 50 lat badań

Krzysztof Waleron¹, Joanna Nakonieczna¹, Małgorzata Waleron²,
Anna J. Podhajska¹

¹Katedra Biotechnologii Międzyuczelniany Wydział Biotechnologii
UG i AMG, Gdańsk

²Zakład Ochrony i Biotechnologii Roślin Międzyuczelniany Wydział
Biotechnologii UG i AMG, Gdańsk

50 years of studies of restriction – modification systems

Summary

The basic function of restriction endonucleases and methyltransferases is protection of the host genome against foreign DNA. Their recombination and transposition related functions are still being discussed. Some authors postulate that R-M genes may act as selfish genetic elements. Restriction endonucleases are indispensable tools in molecular biology. As these enzymes recognize DNA sequence very specifically, they serve as a model for protein-DNA interaction. Restriction-modification genes have also played the same role as a model for evolutionary studies as well as protein structure – function relations. So far, there have been more than 3500 bacterial strains studied which possessed R-M genes of more than 280 different specificities towards recognition sequence and cleavage sites. They became a very good commercial product for many biotechnological companies. At present, in a genome sequencing era, R-M genes seem to be much more common than it was thought before.

Key words:

endonuclease, methyltransferase, restriction, modification.

Adres do korespondencji

Krzysztof Waleron,
Katedra Biotechnologii,
Międzyuczelniany Wydział
Biotechnologii UG-AMG,
ul. Kładki 24,
80-824 Gdańsk;
e-mail:
waleron@biotech.univ.gda.pl

1. Wstęp

Endonukleazy restrykcyjne i metylotransferazy występują powszechnie w świecie mikroorganizmów prokariotycznych. Ich podstawową funkcją jest ochrona genomu gospodarza przed obcym DNA wnikającym do komórki. Inne funkcje, związek z re-

kombinacją i transpozycją są wciąż dyskutowane. Jest wiele dowodów, że geny kodujące enzymy restrykcyjne i modyfikujące DNA mogą działać jako elementy pasożytnicze w genomach bakteryjnych. Endonukleazy restrykcyjne w szczególności typu II odegrały niezwykle istotną rolę w badaniach nad genetyką organizmów i biologii molekularnej. Są nieodłącznym narzędziem w analizach genetycznych, klonowaniu, badaniu zróżnicowania genetycznego. Stały się również doskonałym modelem do badań specyficznych oddziaływań białko – kwasy nukleinowe. Posłużyły jako model do poznania zależności pomiędzy strukturą a funkcją białek oraz do badania mechanizmów ewolucji wewnątrz tej ogromnej rodziny funkcjonalnie pokrewnych białek. Enzymy te ulegają intensywnemu horyzontalnemu transferowi pomiędzy genomami, co wywnioskowano na podstawie homologii sekwencji, analizy użycia kodonów i różnic w składzie par GC. Systemy R-M często powiązane są z ruchomymi elementami genetycznymi takimi jak: plazmidy, wirusy, transpozony i integrony. Determinanty genetyczne systemów R-M zachowują się jak mobilne elementy i powodują rearanżacje w genomach. Analiza sekwencji kompletnych genomów bakteryjnych rzuciła nowe światło na zagadnienia związane z dystrybucją tych enzymów wśród mikroorganizmów. Mimo ogromnego zasobu wiedzy na ich temat wciąż zaskakują naukowców, budzą zainteresowanie i kontrowersje.

2. Funkcje systemów R-M

Ponad 50 lat temu zaobserwowano, że bakteriofagi zdolne do infekcji jednego szczepu *E. coli*, bardzo słabo infekują inny badany szczep *E. coli*. Stwierdzono, że większość zakażonych bakteriofagiem bakterii wcale nie produkowało potomstwa fagowego, gdyż fagowy materiał genetyczny ulegał w komórce bakteryjnej zniszczeniu, proces ten nazwano kilka lat później restrykcją. Nieliczne komórki bakterii produkowały normalną liczbę cząsteczek fagowych, dzięki procesowi, który określono jako modyfikacja DNA. Obserwacji tych dokonali w 1952 r. Luria i Human dla bakteriofagów serii T (1) oraz Bertani i Weigle dla fagów P1 i P2 (2).

Wkrótce wykazano, że restrykcja i modyfikacja nie ogranicza się jedynie do bakteriofagów, ale dotyczy każdego rodzaju DNA, wnikającego do komórki bakteryjnej w wyniku transdukcji, koniugacji, bądź transformacji (3-5).

Systemy restrykcyjno-modyfikacyjne (R-M) poza nielicznymi wyjątkami obejmują dwie aktywności enzymatyczne: endonukleolityczną i modyfikacyjną (metylującą). Zaobserwowano również występowanie w obrębie tych systemów dodatkowego elementu, białka C pełniącego funkcje regulatorowe (6,7). Oba enzymy: endonukleaza restrykcyjna i metylotransferaza rozpoznają tę samą sekwencję nukleotydową w DNA, lecz różnią się pod względem specyfiki katalizowanej reakcji i roli jaką pełnią w komórce. Endonukleaza hydrolizuje wiązania fosfodiesterowe w obrębie specyficznych sekwencji niemetylowanego DNA lub w ich sąsiedztwie, zaś metylo-

transferaza modyfikuje je przez przyłączenie do reszty adeniny lub cytozyny grupy metylowej, chroniąc w ten sposób DNA gospodarza przed restrykcją.

Występowanie systemów R-M, jak się wydawało, było ograniczone do królestwa *Procaryota*, zarówno bakterii właściwych jak i sinic, tymczasem ich obecność wykryto również u bakteriofagów oraz wirusów atakujących niższe organizmy eukariotyczne, np. glony z rodzaju *Chlorella*.

Sens istnienia systemów R-M wśród organizmów prokariotycznych uzasadniano potrzebą ochrony komórki przed wnikaniem obcego DNA, głównie DNA fagowego, który pozbawiony specyficznej modyfikacji (metylacji) ulega degradacji, w wyniku działania enzymów restrykcyjnych, podczas gdy metylowany DNA gospodarza jest w pełni chroniony i niepodlega restrykcji. W 1987 r. Kohring i Mayer potwierdzili, że zgodnie ze swoją funkcją cząsteczki endonukleazy są zlokalizowane w peryferyjnych obszarach komórki, zaś metylotransferazy równomiernie rozproszone w cytoplazmie (8). Pośrednim dowodem świadczącym o słuszności tej roli prokariotycznych systemów R-M jest obecność licznych zabezpieczeń antyrestrykcyjnych wytworzonych przez wiele bakteriofagów, prawdopodobnie jako odpowiedź na zjawisko restrykcji (9).

3. Klasyfikacja systemów restrykcyjno-modyfikacyjnych

Początkowo poznane systemy R-M sklasyfikowano w obrębie dwóch typów. W 1971 r. Boyer zaproponował, by podzielić endonukleazy na podstawie kofaktorów wymaganych do reakcji (5). W miarę przybywania nowych danych pojawiała się konieczność modyfikowania zasad klasyfikacji systemów R-M. W 1985 r. Szybalski wyróżnił następujące kryteria pozwalające wyróżnić cztery klasy enzymów:

- struktura cząsteczki enzymu,
- wymagane kofaktory do restrykcji i modyfikacji DNA,
- rozpoznawana sekwencja, która może być: ciągła lub nieciągła, symetryczna lub asymetryczna, unikatowa lub zdegenerowana, złożona z 4 lub większej liczby nukleotydów,
- specyfika cięcia DNA: w lub poza sekwencją rozpoznawaną w dowolnej lub ściśle zdefiniowanej odległości, cięcie sekwencji metylowanych lub niemetylowanych.

3.1. Enzymy zaliczane do klasy I systemów restrykcyjno-modyfikacyjnych

Systemy restrykcyjno-modyfikacyjne klasy I zostały odkryte w komórkach bakterii z rodziny *Enterobacteriaceae*, oraz bakterii należących między innymi do gatunków: *Bacillus subtilis*, *Neisseria gonorrhoeae*, *Citrobacter freundii*, *Mycoplasma pulmonis*, *Klebsiella pneumoniae*, *Lactococcus lactis* i *Staphylococcus aureus*. Dotąd scharakteryzo-

wano kilkadziesiąt endonukleaz typu I (10). Systemy R-M tej klasy zbudowane są z trzech genów: *hsdS*, *hsdM* i *hsdR* (ang. *host specificity for DNA*). Produkt genu *hsdS* jest odpowiedzialny za rozpoznawanie specyficznych sekwencji podczas restrykcji jak i modyfikacji DNA. Produkt genu *hsdM* warunkuje modyfikację DNA, polegającą na przekształceniu adeniny w obrębie specyficznej sekwencji w N⁶-metyloadeninę. Natomiast produkt genu *hsdR* wraz z produktami pozostałych genów (*hsdS* i *hsdM*) jest odpowiedzialny za restrykcję (hydrolizę) DNA. Na podstawie badań porównawczych wyróżniono cztery różne typy systemów tej klasy: IA, IB, IC, ID, różniące się składem nukleotydowym genów strukturalnych (tab. 1). W przypadku wszystkich typów systemów R-M klasy I opisano zarówno monofunkcjonalną metylotransferazę, jak i wielofunkcjonalny kompleks enzymatyczny, spełniający jednocześnie funkcję restryktazy i metylazy DNA.

Tabela 1

Przykłady enzymów R-M typu I

Rodzina	Enzym	Sekwencja rozpoznawana
IA	<i>EcoBI</i>	TGANNNNNNNTGCT
	<i>EcoKI</i>	AACNNNNNGTGC
	<i>EcoDI</i>	TTANNNNNGTCY
	<i>StyLTHI</i>	GAGNNNNNRATYG
	<i>StySPI</i>	AACNNNNNGTRC
IB	<i>EcoAI</i>	GAGNNNNNGTCA
	<i>EcoEI</i>	GAGNNNNNATGC
	<i>CfrAI</i>	GCANNNNNNGTGG
IC	<i>EcoR124I</i>	GAANNNNNRTCG
	<i>EcoR124IΔ</i>	GAANNNNNTTC
	<i>EcoR124II</i>	GAANNNNNRTCG
	<i>EcoDXXI</i>	TCANNNNNR TTC
	<i>EcoprrI</i>	CCANNNNNR TGC
	<i>EcoDXXIΔ</i>	TCANNNNNTGA
ID	<i>StySBLI</i>	CGANNNNNTACC

Wielofunkcjonalny kompleks enzymatyczny zawiera produkty genów *hsdM*, *hsdR* i *hsdS* najprawdopodobniej w stosunku stechiometrycznym: 2:2:1. Siły oddziaływania (asocjacji) pomiędzy produktem genu *hsdR* a enzymem modyfikującym DNA, jak się wydaje, są bardzo słabe, prawdopodobnie z tego powodu nie można oczyścić całego białkowego kompleksu restrykcyjno-modyfikacyjnego. Możliwe jest natomiast przywrócenie aktywności restrykcyjnej kompleksu enzymatycznego R-M po

oddzielnym oczyszczeniu enzymu modyfikującego i endonukleazy, a następnie zmieszaniu obydwu białek. Wykazano, że poszczególne polipeptydy wchodzące w skład kompleksu R-M klasy I zawierają stosunkowo dobrze zdefiniowane domeny białkowe, odpowiedzialne za poszczególne funkcje polipeptydów podczas oddziaływania kompleksu enzymatycznego R-M ze specyficzną sekwencją DNA. Polipeptyd HsdM zawiera konserwatywną domenę w części N-terminalnej białka, która jest wspólna dla wszystkich mN⁶-metylotransferaz i jest odpowiedzialna za wiązanie S-adenozylu-L-metioniny (SAM) będącej donorem grup metylowych (11). Natomiast polipeptyd HsdR zawiera konserwatywną domenę specyficzną dla białek wiążących ATP (12). Wykazano również, że polipeptyd S klasy I R-M typu IA, IB i IC posiada dwie domeny odpowiedzialne za rozpoznawanie specyficznych sekwencji DNA (TRD, ang. *Target Recognition Domain*) (13-16).

3.2. Enzymy zaliczane do klasy II systemów restrykcyjno-modyfikacyjnych

Najliczniejszą i najlepiej poznaną grupą enzymów systemów R-M, są enzymy klasy II. Do tej pory opisano ponad 3500 endonukleaz należących do tej klasy, wykazujących około 240 różnych specyficzności enzymatycznych, w odniesieniu do rozpoznawanej sekwencji DNA (10). Często obserwuje się występowanie kilku systemów R-M w jednym gospodarzu bakteryjnym, np. w jednym ze szczepów *Neisseria gonorrhoeae* zidentyfikowano aktywność aż 14 metylotransferaz i 8 endonukleaz restrykcyjnych w większości zaliczonych do klasy II (17).

Systemy R-M klasy II składają się, z nielicznymi wyjątkami, z dwóch strukturalnie i funkcjonalnie niezależnych enzymów: endonukleazy restrykcyjnej i metylotransferazy. Endonukleaza, w obecności jonów Mg²⁺, trawi cząsteczkę DNA w obrębie rozpoznawanej specyficznej sekwencji lub w ściśle określonej od niej odległości. Drugi składnik systemu R-M, metylotransferaza DNA katalizuje reakcję przeniesienia grupy metylowej z SAM na reszty adeniny lub cytozyny, przekształcając je w N⁶-metyloadeninę, N⁴-metylocytozynę lub N⁵-metylocytozynę. Endonukleazy restrykcyjne klasy II zwykle rozpoznają specyficzne sekwencje DNA, o długości 4-8 par zasad. Endonukleazy restrykcyjne rozpoznające taką samą sekwencję DNA nazywamy izoschizomerami. Jeżeli natomiast rozpoznają tę samą sekwencję, ale różnią się miejscem cięcia cząsteczki DNA w obrębie tej sekwencji to określamy je mianem neoschizomerów. Wśród enzymów zaliczonych do tej klasy na podstawie charakterystycznych właściwości wyróżniono kilka podtypów (tab. 2).

Podtypy typu II endonukleaz restrykcyjnych na podstawie Roberts et al., (18)

Podtyp	Cechy charakterystyczne	Przykłady	Sekwencja rozpoznawana
A	asymetryczna sekwencja rozpoznawana	<i>FokI</i> <i>AciI</i>	GGATG (9/13)* CCGC (-3/-1)*
B	tnie po obydwu stronach sekwencji rozpoznawanej	<i>BcgI</i>	(10/12) CGANNNNNNTGC (12/10)*
C	symetryczna lub asymetryczna sekwencja rozpoznawana, funkcje R i M w jednym polipeptydzie	<i>GsuI</i> <i>HaeIV</i> <i>BcgI</i>	CTGGAG (16/14)* (7/13) GAYNNNNNRTC (14/9)* (10/12) CGANNNNNNTGC (12/10)*
E	dwie rozpoznawane sekwencje docelowe, jedna jako efektor	<i>EcoRII</i> <i>NaeI</i>	↓CCWGG GCC↓GGC
F	dwie rozpoznawane i cięte sekwencje docelowe	<i>SfiI</i> <i>SgrAI</i>	GGCCNNNN↓NGGCC CR↓CCGGYG
G	symetryczna lub asymetryczna sekwencja rozpoznawana, zależność od SAM	<i>BsgI</i> <i>Eco57I</i> <i>MmeI</i>	GTGCAG (16/14)* CTGAAG (16/14)* TCCRAC (20/18)*
H	symetryczna lub asymetryczna sekwencja rozpoznawana, podobieństwo do genów strukturalnych typu I	<i>BcgI</i> <i>AbdI</i>	(10/12) CGANNNNNNTGC (12/10)* GACNNN↓NNGTC
M	podtyp P lub A, wymaga sekwencji metylowanej	<i>DpnI</i>	Gm6A↓TC
P	symetryczne miejsca rozpoznawane i cięcie	<i>EcoRI</i> <i>PpuRI</i> <i>BsI</i>	G↓AATTC RG↓GWCCY CCNNNNN↓NNGG
S	asymetryczne miejsca rozpoznawane i cięcie	<i>FokI</i>	GGATG (9/13)*
T	symetryczna lub asymetryczna sekwencja rozpoznawana, geny R – heterodimery	<i>Bpu10I</i> <i>BsI</i>	CCTNAGC (-5/-2)* CCNNNNN↓NNGG

* – w nawiasach podano odległość miejsca cięcia od sekwencji rozpoznawanej przez enzym odpowiednio na górnej i dolnej nici, liczby odpowiadają ilości nukleotydów na nici DNA.

3.3. Enzymy zaliczane do klasy III systemów restrykcyjno-modyfikacyjnych

Enzymy klasy III stanowią najmniejszą poznaną grupę systemów R-M. Scharakteryzowano zaledwie kilkunastu jej przedstawicieli. Najwcześniej poznanymi systemami tej grupy są *EcoPI* i *EcoP15I* wyizolowane z bakteriofagów *E. coli* P1 i P15, system *HinIII* z *Haemophilus influenzae* Rf oraz system *StyLTI* kodowany przez DNA chromosomalny wielu gatunków z rodzaju *Salmonella*. Wszystkie zaliczone do tej grupy enzymy wykazują bardzo podobną strukturę i mechanizm działania. Poza tym wykazano, że trzy systemy R-M z *Enterobacteriaceae*: *EcoPI*, *EcoP15I* i *StyLTI* charakteryzują się dużym podobieństwem sekwencji aminokwasowych. Na podstawie analizy sekwencji dostępnych genomów różnych mikroorganizmów zidentyfikowano kilkadziesiąt genów mogących potencjalnie pełnić funkcje endonukleaz typu III (tab. 3).

Enzymy klasy III są kodowane przez dwa geny strukturalne *mod* i *res*. Geny *mod* wykazują strukturalne podobieństwo do genów *hsdS* I klasy R-M, zawierają regiony konserwatywne oraz regiony zmienne. W budowie genów *mod* można wyróżnić trzy rejony: dwa wysoce konserwatywne przy końcach 5' i 3', oraz zmienny rejon centralny (19). Uważa się, że rejony konserwatywne kodują domeny białkowe odpowiedzialne za oddziaływanie polipeptydu Mod z polipeptydem Res będącym produktem genu *res*. Centralny rejon zmienny genów *mod* najprawdopodobniej koduje domeny białkowe odpowiedzialne za wiązanie polipeptydu Mod ze specyficzną sekwencją DNA oraz jest przynajmniej częściowo odpowiedzialny za wiązanie SAM (19-23). Endonukleazy restrykcyjne klasy III rozpoznają niesymetryczną sekwencję DNA: AGACC – *EcoPI*; CAGCAG – *EcoP15I*; CGAAT – *HinfIII*; CAGAG – *StyLTI* i przecinają cząsteczkę DNA w odległości 25-26 pz na prawo od rozpoznawanej sekwencji. W przypadku wszystkich poznanych enzymów tej klasy metylotransferaza modyfikuje adeninę tylko w jednej nici rozpoznawanej sekwencji DNA, podczas gdy adenina w obrębie drugiej z nici pozostaje niezmodyfikowana.

3.4. Enzymy zaliczane do klasy IV systemów restrykcyjno-modyfikacyjnych

Systemy klasy IV składają się z jednego lub dwóch genów kodujących białka, które hydrolizują tylko DNA zmodyfikowany (metylacja, hydroksymetylacja, glikozo-hydroksymetylacja). Sekwencje rozpoznawane przez enzymy systemów typu IV nie zostały dobrze zdefiniowane, z wyjątkiem *EcoKMcrBC* pochodzącego z *E. coli*, rozpoznającego dwie sekwencje RmC oddalone od siebie o 40 do 3000 pz (par zasad). Cięcie w tym przypadku odbywa się w odległości ok. 30 pz od jednego z tych miejsc.

4. Klonowanie systemów restrykcyjno-modyfikacyjnych

Obecnie, kiedy masowo sekwencjonowane są genomy różnych mikroorganizmów poszukuje się w nich genów kodujących systemy R-M. Do tej pory poznano sekwencje genów kodujących ponad 300 systemów R-M klasy II, a ciągle odkrywano nowe geny, którym przypisuje się kodowanie białek R-M. Najczęściej kodowane są przez DNA chromosomalny, niekiedy zaś przez DNA plazmidowy, jak np. R-M *EcoRI*, *EcoRII*, *EcoRV* czy *PvuII*. W ostatnim czasie na podstawie analiz sekwencji nukleotydowej zidentyfikowano ponad 1400 genów, które mogą być składnikami systemów R-M typu II (10). W tabeli 3 podane są przykłady zsekwencjonowanych genomów bakteryjnych oraz potencjalna liczba genów kodujących systemy R-M odnaleziona na podstawie podobieństw sekwencji nukleotydowej.

Tabela 3

Liczba potencjalnych systemów R-M zidentyfikowanych na podstawie analiz komputerowych sekwencji DNA organizmów w przypadku których sekwencjonowano cały genom; na podstawie Murray N.E. (96)

Pochodzenie	Wielkość genomu (Mpz)	TypI	TypII	TypIII	TypIV
<i>Aeropyrum pernix</i>	1,67		7		
<i>Aquifex aeolicus</i>	1,25				1
<i>Archaeoglobus fulgidus</i>	2,18	1	2	1	
<i>Bacillus subtilis</i>	4,21		2		1
<i>Borrelia burgdoferi</i>	1,44		2		
<i>Campylobacter jejuni</i>	1,64	1	4		1
<i>Chlamydia muridarum</i>	1,07				
<i>Chlamydia trachomatis</i>	1,05				
<i>Chlamydia pneumoniae</i> AR39	1,23				
<i>Deinococcus radiodurans</i>	2,65		4		3
<i>Escherichia coli</i> K-12	4,60	1			3
<i>Haemophilus influenzae</i>	1,83	2	3	1	
<i>Helicobacter pylori</i> 26695	1,66	3	14	2	
<i>Helicobacter pylori</i> J99	1,64	3	16	2	
<i>Methanobacterium thermoautotrophicum</i>	1,75	1	1		3
<i>Methanococcus jannaschii</i>	1,66	3	8		
<i>Mycobacterium tuberculosis</i>	4,40	1	1		
<i>Mycoplasma genitalium</i>	0,58	1			
<i>Mycoplasma pneumoniae</i>	0,81	1	1		
<i>Neisseria meningitidis</i> A	2,18	3	7	2	
<i>Neisseria meningitidis</i> B	2,27	1	4	1	
<i>Pyrococcus abyssi</i>	1,77	1	4		
<i>Pyrococcus horikoshii</i>	1,74		3		
<i>Rickettsia prowazekii</i>	1,10				
<i>Synechocystis</i> sp.	3,57		1		1
<i>Thermatoga maritima</i>	1,80		1		
<i>Treponema pallidum</i>	1,16				
<i>Ureoplasma urealyticum</i>	0,71	1	1		

Geny kodujące endonukleazę i metylotransferazę we wszystkich poznanych przypadkach zlokalizowane są blisko siebie, w obrębie jednego operonu. Zwykle gen kodujący endonukleazę poprzedza gen kodujący metylotransferazę. Zaledwie

w kilku poznanych systemach R-M geny kodujące restryktazę i metylotransferazę są położone w odwrotnej orientacji (24).

Endonukleazy typu II ze względu na precyzję cięcia znalazły największe zastosowanie w rekombinacji DNA *in vitro*. Możliwość uzyskania białek endonukleaz i metylotransferaz w ilościach znacznie większych niż naturalny poziom enzymów występujący w komórkach gospodarza jak również łatwiejsze niż w przypadku pozostałych klas systemów R-M procedury oczyszczania zachęcały do podejmowania badań nad klonowaniem. Próby klonowania genów endonukleaz i metylotransferaz rozpoczęto w wielu laboratoriach w końcu lat siedemdziesiątych. Pierwszy klonowany w 1980 r. przez Szomolanyi i wsp. gen metylotransferazy pochodził z *Bacillus sphaericus* R. Procedura była niezbyt skomplikowana, dzięki łatwej selekcji klonów z aktywnym genem metylotransferazy (25,26). Fragmenty DNA są ligowane z wektorem (zazwyczaj plazmidem) i transformowane do komórek kompetentnych (najczęściej *E. coli*) w celu skonstruowania biblioteki genów. Transformanty inkubuje się w warunkach optymalnych dla wzrostu przez krótki czas, aby DNA rekombinantów niosących gen metylotransferazy mógł ulec modyfikacji. Następnie izoluje się DNA plazmidowy i poddaje trawieniu enzymem restrykcyjnym, który hydrolizuje tylko niezmetylowany DNA. Alternatywna metoda selekcji populacji plazmidów zawierających gen kodujący metylotransferazę oparta jest na fenotypie restrykcyjnym. Przeprowadza się ją *in vivo*, przez transformację gospodarza, w komórkach którego niezmodyfikowany DNA jest hydrolizowany. Obie drogi selekcji genu metylotransferazy, pozwalają na uzyskanie rekombinantów zawierających oba geny. Izolacja obydwu genów wynika w dużej mierze stąd, że determinanty genetyczne restryktazy i metylotransferazy zlokalizowane są na genomie obok siebie lub w niewielkiej odległości. Opisane metody są skuteczne pod pewnymi zastrzeżeniami, klonowany system R-M jest kompletny i zawiera blisko położone geny kodujące metylotransferazy i endonukleazy oraz metylotransferaza ulega ekspresji w komórkach nowego gospodarza.

Zdarza się, że ekspresja genu metylazy jest słaba lub nie zachodzi, lub też koekspresja genów z sekwencji sąsiadujących z genem metylazy jest letalna dla komórek gospodarza i nie można uzyskać aktywnego genu. Opracowano wiele sposobów pozwalających na ominięcie takich ograniczeń. Jeśli pierwszy klon nie zawiera kompletnego genu endonukleazy lub jeśli metylotransferaza musi być poddana ekspresji w komórkach gospodarza jako pierwsza zanim wprowadzony zostanie gen endonukleazy przeprowadza się tzw. klonowanie dwustopniowe. Próbuje się uzyskać najpierw ekspresję genu metylotransferazy, a potem dopiero przeprowadza się klonowanie genu restryktazy (27-29). W przypadkach kiedy gen metylotransferazy nie może być zidentyfikowany można zastosować dodatkową użyteczną procedurę do identyfikacji genu endonukleazy. Polega ona na oczyszczeniu do homogenności odpowiedniej ilości enzymu z macierzystego mikroorganizmu. Następnie polipeptyd endonukleazy poddaje się N-terminalnemu sekwencjonowaniu, sekwencję aminokwasową polipeptydu wykorzystuje się do ustalenia hipotetycznych sekwencji nu-

kleotydom. Postępowanie to umożliwia zaprojektowanie zestawu oligonukleotydów komplementarnych do danej sekwencji genu endonukleazy. Oligonukleotydy używa się jako startery do amplifikacji fragmentu genu z DNA genomowego, macierzystego organizmu bądź też z biblioteki klonów (30). Klony niosące gen endonukleazy można stabilizować za pomocą heterospecyficzných genów metylotransferaz, niezwiązanych z endonukleazą oryginalnego gospodarza, jeśli rozpoznają tę samą bądź podobną sekwencję i będą mogły pełnić funkcję ochronną przed trawieniem przez klonowaną endonukleazę (28).

Endonukleazy restrykcyjne należą do grupy białek cytotoxycznych. Geny kodujące cytotoxyczne białka są niezwykle trudne do klonowania. Często zdarza się, że mimo zastosowania promotorów pozwalających na zablokowanie niepożądaną ekspresji w komórkach gospodarza, ekspresja zachodzi na poziomie minimalnym, ale wystarczającym by być letalnym dla komórek. Stąd klonowanie do wektorów ekspresyjnych genów kodujących białka toksyczne dla komórek biorcy jest bardzo trudne lub niemożliwe (30).

5. Organizacja genów wchodzących w skład systemów R-M

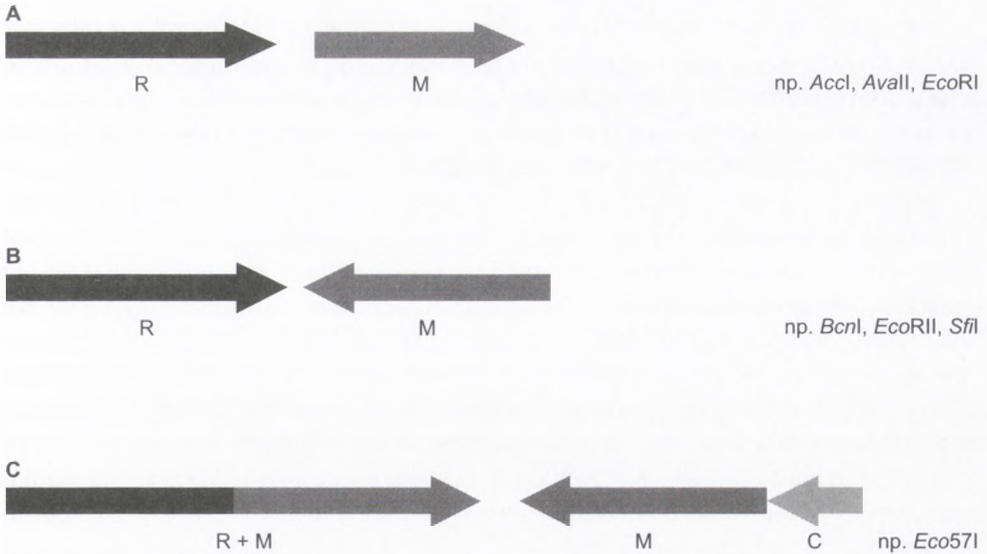
Istnieje duża różnorodność organizacji i położenia genów kodujących systemy R-M.

Geny *hsd* (*hsdS*, *hsdM*, *hsdR*) kodujące system R-M typu I są położone blisko siebie w chromosomie bakteryjnym w tej samej orientacji (rys. 1), *hsdM* zawsze poprzedza *hsdS* i są transkrybowane jako jeden operon, *hsdR* leży przed lub za nimi, posiada swój własny promotor i jest transkrybowany oddzielnie. Zmiany specyficzności podjednostki *hsdS* mają identyczny wpływ zarówno na restrykcję jak i modyfikację. Systemy R-M typu pierwszego mają większe możliwości różnicowania ewolucyjnego niż systemy typu II, ponieważ zmiana w specyficzności jednej aktywności automatycznie powoduje zmianę w drugiej. Zmiany specyficzności w systemach typu I obserwowane były w warunkach laboratoryjnych (24).

Systemy typu II są najprostsze i najpowszechniejsze. Endonukleazy i metylotransferazy są oddzielnymi białkami. Obecne w tych systemach dwa geny strukturalne (wyjątek: *DpnII*), kodujące odpowiednio: endonukleazę i metylotransferazę, położone są w tym samym operonie w różnych orientacjach. Na rysunku 2 przedstawiono wybrane przykłady ułożenia genów kodujących R-M typu II. W obrębie tej



Rys. 1. Organizacja genów systemu R-M typu I.



Rys. 2. Przykłady organizacji genów systemów R-M typu II.

grupy występują również takie białka, które posiadają dwie aktywności: endonukleolityczną i metylującą w jednym polipeptydzie. Przykładem jest system R-M *Eco57I*, który składa się z genu kodującego białko posiadające dwie domeny, tnącą i metylującą oraz genu kodującego metylotransferazę oraz genu białka regulatorowego C (rys. 2C).

W grupie enzymów zaliczanych do typu III występują dwa geny strukturalne: gen *res* poprzedza gen *mod*, a transkrypcja zachodzi ze wspólnego promotora.

Dotychczasowe doniesienia na temat systemów R-M częściej dotyczą odkryć nowych enzymów restrykcyjnych, niż wyników klonowania i ekspresji genów je kodujących. Wynika to m.in. stąd, że nadal stosunkowo słabo poznana jest regulacja ekspresji genów wchodzących w skład systemów R-M, a tym samym często trudno jest uzyskać stabilne klony w szczepach *E. coli*, ponieważ regulacja ekspresji genów może być w nich różna od regulacji zachodzącej w komórkach gospodarza, z którego geny pochodzą (24,31).

Ponieważ obecność endonukleazy restrykcyjnej może być potencjalnie letalna, jej ekspresja podlega ścisłej kontroli. Zanim pojawi się aktywność endonukleazy, metylotransferaza musi zmodyfikować genom komórkowy. Ekspresja genu endonukleazy powinna być regulowana w taki sposób, aby jej poziom był dostosowany do stopnia modyfikacji DNA gospodarza, by w ten sposób zabezpieczyć go przed niszczącym działaniem restryktazy (32,33). Tego typu regulacja jest szczególnie istotna, w przypadku gdy geny kodujące restryktazę i metylazę występują na plazmidach, które na drodze transformacji lub koniugacji mogą być przenoszone do komórek

nowego gospodarza, w których DNA jest niezmodyfikowany (34). Regulacja taka jest również ważna w przypadku systemów R-M naturalnie występujących w komórce podczas chwilowych zaburzeń w syntezie białek, np. w odpowiedzi na głód aminokwasowy, brak donorów grup metylowych, szok termiczny, jak również w wyniku odpowiedzi komórki na utleniacze i uszkodzenia DNA (9).

Regulacja systemów R-M typu I i III wynika z samej organizacji operonów i struktury funkcjonalnej białek. W systemie R-M typu I gen kodujący podjednostkę odpowiedzialną za restrykcję transkrybowany jest z oddzielnego promotora, inaczej niż geny kodujące podjednostki M i S. Natomiast do pojawienia się aktywności endonukleolitycznej niezbędna jest obecność wszystkich trzech podjednostek: R, M i S, co powoduje, że restrykcja zachodzi jedynie, wówczas gdy w komórce nastąpiła modyfikacja DNA. Dodatkowo pojawienie się aktywności endonukleolitycznej jest uwarunkowane obecnością SAM, donora grup metylowych niezbędnych do metylacji. Ma to istotne znaczenie dla zabezpieczenia komórki przed autorestrykcją, w przypadku niedoboru SAM lub produkcji niefunkcjonalnych podjednostek metylazy (32). Podobne zabezpieczenia występują w przypadku systemów R-M typu III, gdzie polipeptyd R, produkt genu *res*, wykazuje aktywność endonukleolityczną jedynie w kompleksie z białkiem M. Ponadto hydroliza DNA następuje tylko, wówczas gdy dwie niezmodyfikowane, rozpoznawane sekwencje ułożone są w odwrotnej orientacji. Zależność ta zapobiega degradacji DNA gospodarza podczas replikacji, gdy DNA nie jest w pełni metylowany (9,32). W ostatnich latach opisano również bardzo ciekawy mechanizm, kontrolujący funkcję R-M typu I, zależny od komórkowej proteazy ClpXP (35). W odpowiedzi na czynniki uszkadzające DNA, w sytuacji, kiedy aktywność metylująca jest zablokowana (36) lub jedynie obniżona (37) dochodzi do proteolizy białka HsdR przeprowadzanej przez protezę ClpX, aby nowo powstający genomowy DNA, który nie został zmodyfikowany przez odpowiednie metylotransferazy nie uległ zniszczeniu. W jaki sposób komórka bakteryjna odróżnia własny niezmodyfikowany DNA od obcego niezmodyfikowanego pozostaje ciągle w sferze badań.

W przypadku systemów R-M typu II problem regulacji ekspresji jest bardziej zróżnicowany. Regulacja może mieć tu charakter bierny lub czynny. Regulacja czynna odbywa się z udziałem specyficznych białek o charakterze represorów lub induktorów. Jeden z mechanizmów regulacji biernej związany jest ze strukturą operonów i wynika z nakładania się pewnych rejonów genów kodujących endonukleazę i genów kodujących metylotransferazę. W systemie R-M *Pst*I, geny kodujące R i M mają orientacje rozbieżne. Silniejszy promotor metylotransferazy pokrywa się ze słabszym, zorientowanym przeciwnie promotorem endonukleazy i współzawodniczy z nim o inicjację transkrypcji (33). Natomiast w systemie R-M *Sma*I, gdzie geny obu enzymów mają orientację zbieżną i zachodzą na siebie rejonami końcowymi, można przypuszczać, że transkrypcja jednego genu powoduje przerwanie transkrypcji drugiego (38). Alternatywny mechanizm bierny wynikałby stąd, że zdecydowana większość endonukleaz to homodimery, a zatem do ujawnienia się aktywności endonukleolitycznej niezbędne jest połączenie się dwóch podjednostek, podczas gdy więk-

szczość metylotransferaz jest aktywna jako monomery. Powoduje to opóźnienie wystąpienia aktywności endonukleolitycznej w stosunku do modyfikacyjnej (33).

Wśród systemów R-M typu II stwierdzono możliwość regulacji na poziomie transkrypcji uwarunkowaną obecnością w pobliżu lub w obrębie sekwencji genów kodujących endonukleazę lub metylotransferazę, miejsc rozpoznawanych przez te enzymy. Dotyczy to np. systemów: *PaeR7I* (dwie sekwencje rozpoznawane przed genem *paeR7IM*), *FokI* (po dwie sekwencje rozpoznawane przed genami: *fokIM* i *fokIR*), *MbolI* (dwie sekwencje rozpoznawane w obrębie genu *mbolIR*), *TaqI* (siedem sekwencji rozpoznawanych w obrębie genu *taqIR*). Takie strategiczne rozmieszczenie rozpoznawanych sekwencji, w miejscach mogących wpływać na transkrypcję, sugeruje możliwość formowania obszarów wiązania czynników regulatorowych takich jak: represory czy induktory, bądź samych białek endonukleaz lub metylotransferaz, których powinowactwo zależałoby od stopnia modyfikacji sekwencji (32).

Ponadto zaproponowano również mechanizmy regulacji potranskrypcyjnej na poziomie inicjacji translacji. Dotyczy to systemów: *EcoRV*, *TaqI*, *FokI*, w których w rejonie promotorowym genów restryktaz stwierdzono obecność struktury typu „spinki do włosów”, która obniża poziom ich ekspresji (39-41). Pewną rolę w opóźnianiu ekspresji genu endonukleazy, a tym samym ograniczeniu jej aktywności w komórce mogłoby również odgrywać występowanie stosunkowo słabych miejsc wiązania rybosomów (RBS, ang. *Ribosomal Binding Site*) dla genu restryktazy, jak ma to miejsce w przypadku systemu *EcoRV* (39).

Kolejny mechanizm regulacji na poziomie mRNA zaproponowano dla systemu *DpnII*. System ten zawiera jedną endonukleazę *DpnB* i dwie metylotransferazy: *DpnM*, która metyluje DNA dwuniciowy i *DpnA*, która preferuje DNA jednoniciowy. Wszystkie trzy geny, w kolejności: *dpnM*, *dpnA*, *dpnB*, mają tę samą orientację i są pod kontrolą wspólnego promotora, który poprzedza gen *dpnM*. Dodatkowo przed genem *dpnM* znajduje się tzw. nietypowe miejsce wiązania rybosomów (ARBS, ang. *Atypical Ribosomal Binding Site*), które bez połączenia z hipotetycznym, pomocniczym białkiem komórkowym nie jest rozpoznawane przez rybosomy. Podobne miejsce ARBS, jak również klasyczne RBS, znajduje się przed genem *dpnA*. Natomiast przed genem *dpnB* obecne jest jedynie RBS. Ze wspólnego promotora powstaje jeden długi transkrypt mRNA. W miejscu ARBS może rozpocząć się translacja tylko w obecności pomocniczego czynnika komórkowego, który oddziałując z ARBS umożliwia wiązanie się rybosomów. Zakłada się, że białko to produkowane jest konstytutywnie. W przypadku gdy w komórce istnieje pojedyncza kopia operonu kodującego system *DpnII* to translacja może zaczynać się zarówno w ARBS jak i w RBS. Wówczas wszystkie trzy białka: *DpnM*, *DpnA*, *DpnB* produkowane są na tym samym poziomie. Jeśli w komórce występuje więcej niż jedna kopia operonu, wówczas ze względu na ograniczoną ilość czynnika pomocniczego, tylko część transkryptu mRNA może ulec translacji z ARBS. W pozostałych przypadkach translacja zachodzi jedynie z RBS, a niechroniony przez rybosomy koniec 5' mRNA kodujący *DpnM* ulega degradacji. Powstają wówczas jedynie białka: *DpnA* i *DpnB* (42).

Sekwencje RBS zaangażowane są również w regulację ekspresji genów systemu *McrA E. coli*. Poziom ekspresji zależny jest tu od tworzenia struktur drugorzędowych mRNA w rejonie inicjacji translacji, co wiąże się ze stopniem dostępności RBS dla rybosomów (43).

Wnikliwa analiza sekwencji nukleotydowych niektórych systemów R-M typu II wykazała obecność dodatkowej ramki odczytu (ORF, ang. *Open Reading Frame*), leżącej zwykle poza genem metylotransferazy, przeważnie tuż przed genem restryktazy (7,33,44). Okazało się, że mutacje w obrębie tej dodatkowej ORF dają fenotyp nierestykcyjny nawet, jeśli gen endonukleazy jest nienaruszony. Efekt ten zostaje zniesiony przez dodanie niezmienionej kopii ORF w układzie *in trans*. Białka kodowane przez te dodatkowe ramki odczytu zostały nazwane białkami C (ang. *controller*) (7). Na podstawie analizy sekwencji tych białek wykazano obecność wysoce konserwatywnego regionu przypominającego strukturę helisa-skręt-helisa (ang. *helix-turn-helix*), charakterystycznego dla wielu białek regulatorowych wiążących DNA. Na tej podstawie założono, że białka C mogą funkcjonować jako represory lub aktywatory transkrypcyjne, analogicznie do represora λ bakteriofaga λ (33).

W przypadku systemu *PvuII* zaobserwowano, że gen kodujący białko C i gen endonukleazy *pvuII*R nakładają się na siebie i mają tę samą orientację. Okazało się, że ekspresja genu endonukleazy zachodzi efektywnie dopiero po nagromadzeniu się białka C, które działa tu jako pozytywny regulator ekspresji, opóźniając jednocześnie pojawienie się aktywności endonukleolitycznej. Natomiast metylotransferaza produkowana jest niezależnie od syntezy białka C. Dodatkowo w obrębie genu metylotransferazy wykryto niewielką, przeciwnie zorientowaną czwartą ramkę odczytu. Kodowany przez nią polipeptyd, tzw. białko W, strukturą swoją przypomina element budowy endonukleazy, w miejscu wiązania się dwóch podjednostek homodimeru. Na podstawie zaobserwowanego podobieństwa założono istnienie dodatkowego mechanizmu regulacji, w którym hipotetyczny produkt czwartej ramki odczytu – białko W, z uwagi na swoją budowę, może działać jako inhibitor łączenia się podjednostek endonukleazy. Współzawodnicząc o miejsce wiązania dwóch podjednostek i wiążąc się z nimi może zapobiegać dimeryzacji niewielkiej liczby cząsteczek endonukleazy, jakie mogłyby powstać z pierwotnego, długiego transkryptu, zanim powstałaby odpowiednia ilość białka C. Jest to dodatkowy mechanizm chroniący przed zbyt wczesnym pojawieniem się aktywności endonukleolitycznej (34,45). Poza tym stwierdzono, że białka C mają zdolność ograniczenia ekspresji genu metylotransferaz, w przypadku gdy cały DNA komórkowy został zmodyfikowany (32).

6. Poglądy na temat pokrewieństwa i ewolucji systemów R-M

Problem pokrewieństwa i ewolucji systemów restrykcyjno-modyfikacyjnych jest wciąż otwarty. Przeprowadzono liczne analizy porównawcze sekwencji aminokwa-

sowych wśród endonukleaz i metylotransferaz w celu znalezienia homologii mogących wskazać na istnienie powiązań ewolucyjnych.

Na podstawie tych analiz wykazano, że metylotransferazy posiadają wiele konserwatywnych motywów w sekwencjach aminokwasowych, które umożliwiają badanie ich pokrewieństwa (46-56). Największe podobieństwo sekwencji stwierdzono w przypadku genów kodujących m⁵C-metylotransferazy. Rodzina metylotransferaz m⁵C jest z ewolucyjnego punktu widzenia szczególna i jej przedstawiciele występują zarówno u *Procarlyota* jak i u *Eucaryota*. Mimo różnej specyficzności, można u nich zlokalizować zestaw dziesięciu wysoce konserwatywnych motywów aminokwasowych występujących w tym samym liniowym porządku, które są niezwykle pomocne podczas identyfikacji nowych enzymów na podstawie sekwencji pierwszorzędowej (49,52,56,57). Również w przypadku dwóch pozostałych klas metylotransferaz: m⁶A-metylotransferaz oraz m⁴C-metylotransferaz, można wyróżnić pewne podobieństwa w sekwencji aminokwasowej białek, jak i nukleotydowej genów kodujących te enzymy. Znaczne różnice w pierwszorzędowej strukturze obserwuje się natomiast w porównaniach sekwencji aminokwasowej adenino-metylotransferaz z sekwencjami cytozyno-metylotransferaz (46,47).

Nie stwierdzono żadnych podobieństw zarówno w sekwencji nukleotydowej genów jak i aminokwasowej białek endonukleazy i metylotransferazy wchodzących w skład tego samego systemu R-M. Jedynym opisanym wyjątkiem, w przypadku którego można dopatrywać się takich podobieństw, jest system *EcoRII* (58).

W przeciwieństwie do metylotransferaz endonukleazy nie posiadają znaczących podobieństw w sekwencjach aminokwasowych z wyjątkiem kilku par izoschizomerów endonukleaz katalizujących identyczną reakcję (rozpoznające identyczną sekwencję nukleotydową i hydrolizujące DNA w identycznym miejscu) jak np. *EcoRI* i *RsrI*, *FnuDI* i *NgoPII*, *XmaI* i *Cfr9I*. Należy nadmienić, że wszystkie wymienione pary izoschizomerów pochodzą z gospodarzy należących do różnych gatunków bakterii.

Natomiast te endonukleazy, które rozpoznają odmienne sekwencje, nie wykazywały, jak się wydawało, żadnej homologii. W ostatnich badaniach wykorzystujących najnowszą wiedzę informatyczną wykazano, że homologia może występować również na poziomie trójwymiarowym – przestrzennej struktury białkowej (47). Nasuwa to sugestię dotyczące drogi ewolucji endonukleaz. Brak podobieństwa sekwencji aminokwasowych sugerowałby, że endonukleazy ewoluowały niezależnie. Niektóre izoschimeryczne systemy są spokrewnione pośrednio (podobne komponenty, lecz odmienna organizacja genów), albo bezpośrednio (zarówno organizacja genów, jak i skład nukleotydowy genów kodujących obydwa enzymy wykazują korelacje). Inne są zaledwie marginalnie podobne, nasuwając pytanie, czy powstały one drogą dywergencji od bardzo dawnego przodka, czy ewoluowały niezależnie od siebie. Przykładem są *HhaI* i *Hinfl*, które mają tę samą organizację genów, lecz odpowiednio enzymy składowe wykazują niewielki procent zgodności sekwencji. Jeszcze inne systemy należące do izoschizomerów prawdopodobnie podlegały wielokrotnym zmianom w toku ewolucji, na co wskazuje, na przykład wysoki stopień pokre-

wieństwa metylotransferaz przy jednoczesnej odmienności endonukleaz lub różnej organizacji genów w poszczególnych systemach (32). Ze względu na fakt, że endonukleazy restrykcyjne nie wykazują podobieństwa w sekwencji aminokwasowej przyjęto początkowo, że nie są ze sobą spokrewnione. Pierwsze doniesienia, że endonukleazy ewoluowały od wspólnego przodka ukazały się w 1995 r., kiedy Jeltsch opublikował dane wskazujące na związek pomiędzy sekwencją aminokwasową enzymu a rozpoznawaną sekwencją DNA (59). Na podstawie analiz krystalograficznych endonukleaz restrykcyjnych typu II wykazano istnienie strukturalnego rdzenia zbudowanego z β -katek otoczonych α -helisami, których ułożenie przestrzenne pozwala na bliskie sąsiedztwo dwóch lub trzech kwaśnych reszt aminokwasowych i jednej lizyny. W niektórych przypadkach lizyna może być zastąpiona przez glutaminę lub kwas glutaminowy. Ten charakterystyczny motyw określany jest jako motyw PDX₁₀₋₃₀D/EXK po raz pierwszy został opisany dla endonukleaz *EcoRI* i *EcoRV* (60). Motyw ten jest odpowiedzialny za wiązanie jonów Mg⁺⁺ niezbędnych do hydrolizy DNA. Istotna rola tego motywu została potwierdzona licznymi badaniami opartymi na miejscowo specyficznej metagenecie. Spostrzeżenia te potwierdziły tezę, że endonukleazy restrykcyjne mają wspólne korzenie ewolucyjne. Podsumowując wyniki przeprowadzonych badań pokrewieństwa wśród enzymów restrykcyjno-modyfikacyjnych, można przytoczyć skrajne przykłady zarówno konwergencyjnej jak i dywergencyjnej ewolucji. Badania nad sekwencjami kompletnych genomów różnych mikroorganizmów przyniosą z pewnością nowe poglądy na temat pokrewieństwa i ewolucji systemów restrykcyjno-modyfikacyjnych.

7. Poszukiwania i dystrybucja systemów R-M

Niemal wszystkie poznane do dziś enzymy restrykcyjno-modyfikacyjne zostały wykryte metodą systematycznego badania ogromnej liczby izolatów bakteryjnych, otrzymanych z kolekcji szczepów lub izolowanych przez badaczy ze środowiska naturalnego. Poszukiwania enzymów R-M zapoczątkował we wczesnych latach siedemdziesiątych H. Smith (10).

Poszukiwanie endonukleaz restrykcyjnych polegało wówczas na hodowli badanego mikroorganizmu, uzyskaniu z niego ekstraktu komórkowego i inkubowaniu próbek ekstraktu z DNA (np. DNA faga lambda) stanowiącego substrat dla endonukleazy. Następnie strawiony DNA faga lambda rozdzielano elektroforetycznie w żelu agarozowym, w celu stwierdzenia, czy uległ on specyficznej fragmentacji w wyniku aktywności jednej lub kilku endonukleaz restrykcyjnych obecnych w komórkach badanego mikroorganizmu.

Wykryty w ten sposób enzym restrykcyjny był oczyszczany najczęściej przy użyciu technik chromatograficznych, a następnie określano jego specyficzność względem DNA substratowego oraz właściwości fizykochemiczne. W ciągu ostatnich 25 lat badań wykazano, że przy zastosowaniu tej metody w 25% badanych ekstraktów

komórek bakteryjnych można stwierdzić aktywność przynajmniej jednej endonukleazy restrykcyjnej typu II (10,61). Aby zwiększyć prawdopodobieństwo wykrycia nowych endonukleaz restrykcyjnych jako substrat do badania aktywności endonukleazy, wykorzystuje się DNA o możliwie jak największej liczbie różnych wariantów sekwencji nukleotydowych rozpoznawanych przez endonukleazy, czyli zarówno DNA pochodzący z organizmów prokariotycznych jak i eukariotycznych (61).

Dzięki ogromnej różnorodności endonukleazy restrykcyjne znalazły również zastosowanie jako marker taksonomiczny na przykład do różnicowania szczepów w obrębie tego samego gatunku (62-66).

Dla uzyskania nowych aktywności endonukleaz podejmowane są próby tworzenia endonukleaz poprzez zmiany poznanych już enzymów tak, aby w efekcie rozpoznawały nowe sekwencje nukleotydowe. W przypadku enzymów rozpoznających sześcionukleotydowe sekwencje, których struktura krystalograficzna jest znana, próbowano zmienić ich specyficzność na drodze mutagenety (67). Pewnym sukcesem było zmodyfikowanie przez Lanio i wsp. enzymu *EcoRV*, efektywniej hydrolizującego rozszerzoną sekwencję 8-nukleotydową niż wyjściową 6-nukleotydową (68). Zmieniono również sekwencję genu kodującego endonukleazę *BamHI* tak, że zmieniony enzym hydrolizował sekwencje zmetylowane (69). Częściowym sukcesem zakończyły się też badania nad endonukleazą *Eco57I* (70). Zmiana specyficzności restrykcyjnej wiąże się prawdopodobnie ze zmianami wielu aminokwasów. Z tego powodu prowadzone dotąd próby nie zakończyły się spektakularnymi wynikami potwierdzającymi możliwość komponowania nowych specyficzności poprzez zmianę sekwencji aminokwasowej.

W połowie lat osiemdziesiątych zapoczątkowano również konstruowanie tzw. chimerycznych enzymów restrykcyjnych, wykorzystując domenową strukturę endonukleazy *FokI* (należącej do klasy IIS). Pierwsze próby dotyczyły zastosowania *FokI* jako uniwersalnego enzymu restrykcyjnego do precyzyjnego cięcia DNA w określonym przez eksperymentatora miejscu (71,72). W badaniach zespołów Podhajskiej i Szybalskiego stwierdzono, że enzym *FokI* wykazuje dużą tolerancję i wiąże się z sekwencją rozpoznawaną występującą na przyłączonym do nici DNA adaptorze i mimo niepełnej hybrydyzacji w obrębie obydwu nici DNA, enzym „odlicza nukleotydy” prawidłowo i przecina DNA heterogennych nici zgodnie z oczekiwaniami a nie 9 i 13 nukleotydów od sekwencji rozpoznawanej, jak to miało miejsce w przypadku niezmodyfikowanego enzymu *FokI*.

Był to enzym, który rokował powodzenie w konstrukcji rekombinowanych aktywności enzymatycznych. Wkrótce Chandrasegaran, Kim i Huang niezależnie podjęli próby tworzenia enzymów fuzyjnych. DNA kodujący domenę tnącą *FokI* rekombinowano *in vitro* z DNA kodującym domeny białkowe wiążące się z DNA. Z powodzeniem wykorzystywano determinanty genetyczne powszechne u *Eucaryota*, kodujące motywy helisa-pętla-helisa (ang. *helix-loop-helix*), motywy palca cynkowego, helisa-skret-helisa białek zawierających motyw zamka leucynowego odpowiedzialnego za wiązanie się z DNA. W wyniku rekombinacji uzyskiwano endonukleazy

o nowej specyficzności. Tak skonstruowane chimeryczne endonukleazy są zdolne przecinać dwuniciowy DNA bardzo blisko oczekiwanego miejsca, choć z niewielką wydajnością (73-80).

7.1. Poszukiwania systemów R-M w genomach bakteryjnych

Nowe możliwości poszukiwania systemów restrykcyjno-modyfikacyjnych pojawiły się wraz z zakończeniem wielu projektów sekwencjonowania całych genomów bakteryjnych. Obecnie dostępne są dane o ponad 387 sekwencjonowanych genomach bakterii i 26 genomach archebakterii (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi).

Dokładne przeszukanie znanych sekwencji genomów zaowocowało niezwykle interesującymi spostrzeżeniami (31). Determinanty genetyczne metylotransferaz są łatwe do zidentyfikowania podczas analizy sekwencji genomu bakteryjnego ze względu na obecność fragmentów sekwencji kodujących charakterystyczne dla tych białek motywy 6-12 aminokwasów. Determinanty genetyczne endonukleaz restrykcyjnych zazwyczaj są znajdowane w bezpośredniej bliskości genu odpowiadającej im metylazy (24). Niestety, jak dotąd, geny kodujące endonukleazy mogą być jednoznacznie identyfikowane, tylko jeśli wykazują bardzo bliskie podobieństwo sekwencji do sekwencji innego znanego enzymu restrykcyjnego o tej samej specyficzności. Najczęściej opisywane są jako otwarte ramki odczytu leżące przed lub za genem metylotransferazy, które nie wykazują jakiegokolwiek podobieństwa do sekwencji obecnych w GenBanku. W wyniku analizy genomu *Helicobacter pylori* zidentyfikowano 23 geny, które mogą potencjalnie kodować metylotransferazy, z czego 14 może być częścią systemu R-M typu II (31). W tym przypadku stosowane wcześniej metody pozwoliły badaczom na detekcję aktywności 1 metylotransferazy i 2 endonukleaz.

Determinanty genetyczne systemów restrykcyjno-modyfikacyjnych, jak widać, są szerzej rozpowszechnione niż można by było przypuszczać na podstawie wcześniej uzyskanych danych. Nie zawsze jednak ulegają one ekspresji w komórce gospodarza i nie zawsze przejawiają swoją aktywność *in vitro*.

Wyjaśnieniem braku ekspresji endonukleazy w badanych przypadkach może być, między innymi, zjawisko zmienności fazowej w ramce odczytu (ang. *phase variation*). Niektóre geny R-M posiadają krótkie powtórzone sekwencje wewnątrz regionów kodujących, które wykazują skłonność do powodowania błędów w replikacji (31,81,82). Błędy te mogą powodować zmianę ramki odczytu i powstawanie nieaktywnych białek. Istnieje zatem możliwość, że pewna część populacji danego gatunku bakterii może posiadać funkcjonalne geny endonukleaz czy metylotransferaz, a część nie. Wiele potencjalnych enzymów restrykcyjnych może być obecnych w komórce w formie nieaktywnej, prawdopodobnie potrzebują jakiegoś czynnika indukującego aktywację genów. Geny wykazujące takie właściwości mają zdolność do bardzo szybkiej ewolucji poprzez stałe przejścia ze stanu aktywnego w nieaktywny (31,82). Jest to

zjawisko, które w pewien sposób może tłumaczyć tak duże zróżnicowanie pomiędzy enzymami restrykcyjnymi. W hipotezach tych, jak się wydaje, potwierdza się obserwacje dokonywane przy poszukiwaniu endonukleaz restrykcyjnych w sinicach. Napotymano na szczepy sinic z rodzaju *Phormidium* i *Anabaena*, które dawały niepowtarzalne wyniki jeśli chodzi o detekcję aktywności endonukleazy restrykcyjnej (dane nie publikowane). Obserwowano aktywność jednej lub dwóch endonukleaz w komórce w zależności od warunków hodowli. Jeżeli geny endonukleaz restrykcyjnych, jak i prawdopodobnie związane z nimi geny metylotransferaz często oscylują między stanem aktywnym i nieaktywnym, to hipoteza pasożytniczych systemów restrykcyjno-modyfikacyjnych proponowana przez Kobayashi jako wyjaśnienie ich utrzymania w komórce gospodarza, jak się wydaje, jest mniej uniwersalna (83). W hipotezie tej zakłada się, że głównym enzymem systemu R-M jest metylotransferaza, gdyż odpowiada za szereg procesów biologicznych w komórce (83). Utrata tego elementu kompleksu restrykcyjno-modyfikacyjnego prowadzi do śmierci komórki gospodarza w wyniku zniszczenia genomu przez odpowiadającą metylotransferazie endonukleazę. Zabicie komórek gospodarza przez geny uszkodzonego systemu R-M nie jest możliwe, gdy w komórce istnieje drugi kompleks genów R-M, lub inna metylotransferaza, modyfikująca tę samą sekwencję rozpoznawaną (23,24,84). Utrata endonukleazy w komórce, która odpowiedzialna jest za „pilnowanie”, stałej aktywności metylotransferazy może pośrednio przyczynić się do śmierci komórki gospodarza. Komórka taka wówczas nie jest chroniona przed utrzymywaniem się w niej obcego DNA na przykład bakteriofagowego, którego ekspresja może prowadzić do śmierci komórki. Model ten jest realny jedynie w przypadku systemów R-M ulegających stałej ekspresji w komórce.

8. Horyzontalny transfer systemów R-M

W coraz bardziej zaawansowanych badaniach genomów bakteryjnych wskazuje się, że horyzontalny transfer genów może mieć istotny wpływ na dystrybucję różnych grup genów wśród mikroorganizmów oraz ewolucję genomów bakteryjnych (31,85). Częstość transferu określonych genów ma ścisły związek z ich funkcją. Bardzo powszechny jest transfer tzw. grup genów operacyjnych (związanych z przetrwaniem gospodarza w danym środowisku), podczas gdy geny informacyjne (związane z transkrypcją, translacją, naprawą DNA i pokrewnymi procesami) niezwykle rzadko ulegają transferowi.

Scharakteryzowano trzy procesy umożliwiające transfer genów: transformację, koniugację i transdukcję. Specyficzne wymagania każdego z nich sugerują różne prawdopodobieństwa dla ich zajścia w różnych warunkach środowiskowych. Proces koniugacji jest najbardziej złożony i ma najwyższe wymagania. Komórki dawcy muszą zawierać element koniugacyjny (plazmid lub transpozon), a komórki dawcy i biorcy muszą mieć ze sobą kontakt fizyczny wystarczająco stabilny by transfer DNA

był możliwy. Obydwie komórki muszą być metabolicznie aktywne, by możliwa była synteza DNA i inne aktywności (86). Transfer genów na drodze transdukcji wymaga aktywnego metabolicznie biorcy, by transdukujące cząsteczki fagowe mogły być produkowane. Biorca może być oddalony od dawcy zarówno przestrzennie jak i w czasie, gdyż informacja genetyczna zawarta w cząsteczkach fagowych może przetrwać długi czas. Fagi często są bardzo odporne na wiele fizycznych i chemicznych czynników i mogą przetrwać w środowisku w stanie nieuszkodzonym (86). Biorca, podobnie jak dawca, musi mieć identyczną wrażliwość na bakteriofaga. Transfer genów na drodze transformacji nie wymaga żywych komórek dawcy, wystarczy by komórki dawcy uległy lizie i uwolniły DNA do środowiska. Zdolność wolnego DNA do przetrwania i rozprzestrzeniania się w środowisku naturalnym determinuje jak daleko w czasie i przestrzeni komórki biorcy mogą być oddalone od dawcy. Biorca musi być fizjologicznie aktywny by móc przyjąć wolny DNA. Dawca i biorca nie muszą być blisko spokrewnieni by zaszła transformacja. Dlatego też naturalna transformacja ma kilka cech umożliwiających występowanie tego procesu w różnych populacjach mikroorganizmów, szczególnie w środowiskach narażonych na ekstremalne zmiany (np. temperatury, pH, zasolenia) lub, w których spotyka się duże fluktuacje dynamiki populacyjnej (86).

Stwierdzono także, że wiele mikroorganizmów może spontanicznie wydzielać DNA na zewnątrz komórki. W środowiskach naturalnych mogą występować znaczne ilości wolnego DNA wydzielanego aktywnie przez żywe komórki mikroorganizmów, jak również pochodzącego z lizujących komórek. Najczęściej taki DNA jest szybko degradowany przez enzymy bakteryjne hydrolizujące DNA (patrz tab. 2 i tab. 3). Najlepiej poznanym przykładem horyzontalnego transferu genów jest przenoszenie genów oporności na antybiotyki i środki chemiczne (86).

Na podstawie porównania i analizy wielu systemów R-M potwierdzono hipotezę, że geny kodujące enzymy restrykcyjno-modyfikacyjne mogą również ulegać horyzontalnemu transferowi (52,85,87,88,94). Geny kodujące białka związane z ruchliwością DNA: transpozazy, integrazy, inwertazy są niekiedy znajdowane w bezpośredniej bliskości determinant genetycznych systemów R-M zlokalizowanych w chromosomalnym DNA (88-93). Białka te mogą ułatwić transfer genów restrykazy i metylotransferazy pomiędzy różnymi mikroorganizmami.

Po porównaniu 8 izoschizomerów endonukleaz wyizolowanych z komórek bakterii z gatunków należących do rodzaju *Thermus* okazało się, że izoschizomery izolowane ze szczepów pochodzących z bliskich geograficznie środowisk były identyczne lub prawie identyczne, natomiast enzymy izolowane ze szczepów pochodzących ze środowisk odległych geograficznie wykazywały większe zróżnicowanie. Podobieństwo sekwencji aminokwasowej pomiędzy izoschizomerami ze Stanów Zjednoczonych i Japonii wynosiło 75%, USA i Wysp Azorskich 67%, USA i Nowej Zelandii 59%. W przypadku wszystkich 8 izoschizomerów zaobserwowano, w środkowej części ramki odczytu, wysoce konserwowany ewolucyjnie, region 34 aminokwasowy (87-120); zmiany aminokwasów w tym obszarze dotyczą jedynie pozycji 98 i 113 (95).

Autorzy sugerują, że system R-M *TaqI* mógł być w przeszłości przeniesiony do rodzaju *Thermus*. Na potwierdzenie, zwracają oni uwagę na znaczące różnice w proporcjach par G+C w sekwencjach kodujących determinanty genetyczne izoschimerów systemu R-M *TaqI* w stosunku do pozostałej części genomu przedstawicieli rodzaju *Thermus*. Również użycie specyficznych kodonów, odmiennych niż powszechnie występujące w rodzaju *Thermus*, odpowiadających poszczególnym aminokwasom w sekwencji genów kodujących determinanty systemów z rodziny *TaqI*, wskazuje na obce pochodzenie tych genów (85,95). Hipotezę tę, jak się zdaje, potwierdza stwierdzenie po raz pierwszy w mezofilnej sinicy *Phormidium papyraceum*, mikroorganizmie nie należącym do rodzaju *Thermus*, obecności termofilnej endonukleazy *PpaAll* będącej izoschimerem *TaqI* (66). Co więcej, nie stwierdzono w przypadku szczepu *Phormidium papyraceum* towarzyszącej endonukleazy *PpaAll* aktywnej metylotransferazy.

Jeltsch i Pingoud postulują, że badanie występowania specyficznych kodonów w determinantach genetycznych kodujących systemy restrykcyjno-modyfikacyjne pozwala na wskazanie, które geny systemów R-M przywędrowały do danego gospodarza (85). Zaobserwowano, że w poszczególnych gatunkach mikroorganizmów występują charakterystyczne preferencje w użyciu kodonów, kodujących poszczególne aminokwasy. Jeżeli w obrębie badanej determinanty genetycznej zaobserwuje się użycie kodonów innych niż typowe dla genomu gospodarza, to może oznaczać, że gen ten pochodzi z innego organizmu (85).

9. Podsumowanie

Od pierwszej obserwacji aktywności endonukleaz restrykcyjnych i towarzyszących im metylotransferaz dokonanej przez Luriię i Bertaniego minęło ponad 50 lat. Ich odkrycie przyczyniło się do szybszego rozwoju biologii molekularnej i genetyki. Enzymy te stały się jednym z ważniejszych narzędzi badawczych. W efekcie prowadzonych, przez ponad 50 lat, poszukiwań enzymów restrykcyjnych, w wielu światowych laboratoriach przebadano kilkanaście tysięcy szczepów bakterii i scharakteryzowano ponad 3565 różnych szczepów bakteryjnych posiadających endonukleazy restrykcyjne i metylotransferazy, wykazujące ponad 280 różnych specyficzności pod względem rozpoznawanej sekwencji nukleotydowej i miejsca hydrolizy DNA (23). Poznano mechanizm ich działania, regulację ekspresji, budowę przestrzenną cząsteczek. Badano ewolucję i pokrewieństwo oraz mechanizmy dystrybucji systemów restrykcyjno-modyfikacyjnych w świecie mikroorganizmów. Enzymy te stały się także źródłem dochodów wielu firm biotechnologicznych. Analizy sekwencji kompletnych genomów bakteryjnych z pewnością przyniosą jeszcze wiele interesujących obserwacji na temat tej grupy białek.

Literatura

1. Luria S. E., Human M. L., (1952), *J. Bacteriol.*, 64, 557-569.
2. Bertani G., Weigle J. J., (1953), *J. Bacteriol.*, 65, 113-123.
3. Arber W., (1968), *Symp. Soc. Gen. Microbiol.*, 18, 295-314.
4. Arber W., Linn S., (1969), *Annu. Rev. Biochem.*, 38, 467-500.
5. Boyer H. W., (1971), *Annu. Rev. Microbiol.*, 25, 153-176.
6. Tao T., Blumenthal R. M., (1992), *J. Bacteriol.*, 174, 3395-3398.
7. Lunen K., Chang Z., Morgan R., Nwankwo D. O., Slatko B., Wilson G., (1997), *4th NEB Workshop on Biological DNA Modification*, Innsbruck p.6.03, 48.
8. Kohring G. W., Mayer F., (1987), *FEBS Lett.*, 216, 207-210.
9. Bickle A. T., Kruger D. H., (1993), *Microbiol. Rev.*, 57, 434-450.
10. Roberts R. J., Vincze T., Posfai J., Macelis D., (2003), *Nucleic Acids Res.*, 31, 418-420.
11. Powell L. M., Dryden D. T. F., Willcock D. F., Pain R. G., Murray N. E., (1993), *J. Mol. Biol.*, 234, 60-71.
12. Loenen W. A. M., Daniel A. S., Braymer H. D., Murray N. E., (1987), *J. Mol. Biol.*, 198, 159-170.
13. Boyer H. W., Roulland-Dussoix D., (1969), *J. Mol. Biol.*, 41, 459-472.
14. Glover S. W., Colson C., (1969), *Genet Res.*, 13, 227-240.
15. Fuller-Pace F. V., Cowan G. M., Murray N. E., (1985), *J. Mol. Biol.*, 186, 65-75.
16. Fuller-Pace F. V., Murray N. E., (1986), *Proc. Natl. Acad. Sci. USA*, 83, 9368-9372.
17. Stein D. C., Gunn J. S., Radlinska M., Piekarowicz A., (1995), *Gene*, 157, 19-22.
18. Roberts R. J., (2003), *Nucleic Acids Res.*, 31, 1805-1812.
19. Humbelin M., Suri B., Rao D. N., Hornby D. P., Eberle H., Tripfle T., Kenel S., Bickle T. A. (1988), *J. Mol. Biol.*, 200, 1-18.
20. Rosner J. L., (1973), *Virology*, 52, 213-222.
21. Rao D. N., Eberle H., Bickle T. A., (1989), *J. Bacteriol.*, 171, 2347-2352.
22. Rao D. N., Page M. G. P., Bickle T. A., (1989), *J. Mol. Biol.*, 209, 599-606.
23. Ahmad I., Rao D. N., (1994), *J. Mol. Biol.*, 242, 378-388.
24. Wilson G. G., (1991), *Nucleic Acids Res.*, 19, 2539-2566.
25. Mann M. B., Rao R. N., Smith H. O., (1978), *Gene*, 3, 97-112.
26. Kiss A., Baldauf F., (1983), *Gene*, 21, 111-119.
27. Brooks J. E., Benner J. S., Heiter D. F., Silber K. R., Szynter L. A., Jager-Quinton T., Moran L. S., Slatko B. E., Wilson G. G., Nwankwo D. O., (1989), *Nucleic Acids Res.*, 17, 979-997
28. Wilson G. G., Meda M. M., (1993), US Patent Office US 5179015.
29. Brooks J. E., Howard K. A., (1994), US Patent Office US 5320957.
30. Noren C. J., Roberts R. J., Byrd D. R., Morgan R. D., Patti J., (1999), International Patent Office WO 9911821.
31. Roberts R. J., (1998), *The NEB Transcript* 9, 1-4.
32. Wilson G. G., Murray N. E., (1991), *Annu. Rev. Genet.*, 25, 585-627.
33. Tao T., Bourne J. C., Blumenthal R. M., (1991), *J. Bacteriol.*, 173, 1367-1375.
34. Adams G. M., Blumenthal R. M., (1995), *Gene*, 157, 193-199.
35. Makovets S., Doronina V. A., Murray N. E., (1999), *Proc. Natl. Acad. Sci. USA*, 96, 9757-9762.
36. Doronina V. A., Murray N. E., (2001), *Mol. Microbiol.*, 39, 416-428.
37. O'Neill M., Powell L., Murray N. E., (2001), *J. Mol. Biol.* 307, 951-963.
38. Heitman J., (1993), *Genetic Engineering*, 15, 57-108.
39. Bougueleret L., Schwarzstein M., Tsugita A., Zabeau M., (1984), *Nucleic Acid Res.*, 12, 3659-3677.
40. Barany F., (1988), *Gene*, 65, 167-177.
41. Looney M. C., Moran L. S., Jack W. E., Feehery G. R., Benner J. S., Slatko B. E., Wilson G. G., (1989), *Gene*, 80, 193-208.
42. Lacks S. A., Greenberg B., Sabelnikov A. G., (1995), *Gene*, 157, 209-212.
43. Shivapriya R., Prasad R., Narayanan I. L., Krishnaswamy S., Dharmalingam K., (1995), *Gene*, 157, 201-207.

44. Lubys A., Jurenaite S., Janulaitis A., (1999), *Nucleic Acids Res.*, 27, 4228-4234.
45. Adams G. M., (1996), *Diss. Abstr.*, 57, 108.
46. Bujnicki J. M., (1999), *In Silico Biology*, 1, e0016.
47. Bujnicki J. M., (2000), *J. Mol. Evol.*, 50, 39-44.
48. Bujnicki J. M., Radlinska M., (1999), *Nucleic Acids Res.*, 27, 4501-4509.
49. Bujnicki J. M., Radlinska M., (1999), *Acta Microbiol. Pol.*, 48, 19-30.
50. Lauster R., Kriebardis A., Guschlbauer W., (1987), *FEBS Lett.*, 220, 167-176.
51. Lauster R., (1988), *Gene*, 74, 243.
52. Lauster R., (1989), *J. Mol. Biol.*, 206, 313-321.
53. Lauster R., Trautner T. A., Noyer-Weidner M., (1989), *J. Mol. Biol.*, 206, 305-312.
54. Jeltsch A., Kroger M., Pingoud A., (1995), *Gene*, 160, 7-16.
55. Jeltsch A., (1999), *J. Mol. Evol.*, 49, 161-164.
56. Posfai J., Bhagwat A. S., Posfai G., Roberts R. J., (1989), *Nucleic Acids Res.*, 17, 2421-2435.
57. Kumar S., Cheng X., Klimasauskas S., Sha M., Posfai J., Roberts R. J., Wilson G. G., (1994), *Nucleic Acids Res.*, 22, 1-10.
58. Kossykh V., Repyk A., Kaliman A., Buryanov Y., (1989), *Biochim. Biophys. Acta*, 1009, 290-292.
59. Jeltsch A., Kroger M., Pingoud A., (1995), *Gene*, 160, 7-16.
60. Venclovas C., Timinskas A., Siksnyš V., (1994), *Proteins*, 20, 279-282.
61. Raleigh E. A., Vaisvila R., Morgan R. D., (1999), International Patent Office WO 9964632.
62. de Waard A., Duyvesteyn M., (1980), *Arch. Microbiol.*, 128, 242-247.
63. Janulaitis A., Kazleuskiene R., Lazareviciute L., Gilvonauskaitė D., Steponaviciene D., Jagelavicius M., Petrusyte M., Bitinaite J., Vezeviciute Z., Kiuduliene E., Butkus V., (1988), *Gene*, 74, 229-232.
64. Piechula S., Piosik J., Bielawski K., Podhajska A. J., (1996), *Mol. Biotech.*, 5, 97-99.
65. Waleron M., Waleron K., Podhajska A. J., Łojkowska E., (1998), *J. Appl. Genetics.*, 168.
66. Piechula S., Waleron K., Świątek W., Biedrzycka I., Podhajska A. J., (2001), *FEMS Microbiol. Lett.*, 198, 135-140.
67. Szybalski W., (1985), *Gene*, 40, 169-173.
68. Lanio T., Jeltsch A., Pingoud A., (2000), *Protein Enz.*, 13, 275-281.
69. Dorner L. F., Bitinaite J., Whitaker R. D., Schildkraut I., (1999), *J. Mol. Biol.*, 285, 1515-1523.
70. Rimseliene R., Maneliene Z., Lubys A., Janulaitis A., (2003), *J. Mol. Biol.*, 327, 383-391.
71. Podhajska A. J., Szybalski W., (1985), *Gene*, 40, 175-182.
72. Podhajska A. J., Kim S. C., Szybalski W., (1992), *Methods Enzymol.*, 216, 303-309.
73. Chandrasegaran S., (1994), International Patent Office WO 9418313, US Patent Office US 5356802.
74. Chandrasegaran S., (1998), US Patent Office US 5792640.
75. Chandrasegaran S., Smith J., (1999), *Biol. Chem.*, 380, 841-848.
76. Huang B., Schaeffer C. J., Li Q., Tsai M. D., (1996), *J. Protein Chem.*, 15, 481-489.
77. Kim J. J., Min K. T., Kim M. H., Augh S. J., Kim B. D., Lee D. S., (1996), *Gene*, 171, 129-130.
78. Kim Y., Chandrasegaran S., (1994), *Proc. Natl. Acad. Sci. USA*, 91, 883-887.
79. Kim Y. G., Cha J., Chandrasegaran S., (1996), *Proc. Natl. Acad. Sci. USA*, 93, 1156-1160.
80. Kim Y. G., Kim P. S., Herbert A., Rich A., (1997), *Proc. Natl. Acad. Sci. USA*, 94, 12875-12879.
81. Hood D. W., Deadman M. E., Jennings M. P., Bisercic M., Fleischmann R. D., Venter J. C., Moxon E. R., (1996), *Proc. Natl. Acad. Sci. USA*, 93, 11121-11125.
82. Dybvig K., Sitaraman R., French C. T., (1998), *Proc. Natl. Acad. Sci. USA*, 95, 13923-13928.
83. Kobayashi I., (1996), in: *Epigenetic Mechanisms of Gene Regulation*, Eds. Russo V. E. A., Martienssen R. A., Riggs A. D., Cold Spring Harbor Laboratory Press, New York, 155-172.
84. Stankevicius K., Povilionis P., Lubys A., Menkevicius S., Janulaitis A., (1995), *Gene*, 157, 49-53.
85. Jeltsch A., Pingoud A., (1996), *J. Mol. Evol.*, 42, 91-96.
86. Lorenz M. G., Wackernagel W., (1994), *Microbiol. Rev.*, 58, 563-602.
87. Kobayashi I., (2001), *Nucleic Acids Res.*, 29, 3742-3756.
88. Kita K., Tsuda J., Kato T., Okamoto K., Yanase H., Tanaka M., (1999), *J. Bacteriol.*, 181, 6822-6827.
89. Rowe-Magnus D. A., Guerout A. M., Mazel D., (1999), *Research in Microbiology*, 150, 641-651.
90. Karreman C., de Waard A., (1988), *J. Bacteriol.*, 170, 2527-2532.

91. Brassard S., Paquet H., Roy P. H., (1995), *Gene*, 157, 69-72.
92. Anton B. P., Heiter D. F., Benner J. S., Hess E. J., Greenough L., Moran L. S., Slatko B. E., Brooks J. E., (1997), *Gene*, 187, 19-27.
93. Lee K. F., Shaw P. C., Picone S. J., Wilson G. G., Lunen K. D., (1998), *Biol. Chem.*, 379, 437-441.
94. Kobayashi I., Nobusato A., Kobayashi-Takahashi N., Uchiyama I., (1999), *Curr. Opin. Genet. Dev.*, 9, 649-656.
95. Cao W., Lu J., Barany F., (1997), *Gene*, 197, 205-214.
96. Murray N. E., (2002), *Microbiology*, 148, 3-20.