



## Genomika *Rhizobium leguminosarum* – badanie genów syntezy polisacharydów powierzchniowych

Anna Skorupska, Jarosław Król, Andrzej Mazur, Małgorzata Marczak  
Zakład Mikrobiologii Ogólnej, Instytut Mikrobiologii i Biotechnologii,  
Uniwersytet Marii Curie-Skłodowskiej, Lublin

### Genomic approach to study surface polysaccharide genes in *Rhizobium leguminosarum*

#### Summary

Symbioses are natural biotechnological systems in which soil bacteria, named rhizobia, form interactions with legumes, resulting in nitrogen fixation. *Rhizobium leguminosarum* bv. *trifolii* TA1 (*RtTA1*) genome is partitioned into chromosome and four megaplasmids. The macrorestriction map of *RtTA1* genome was constructed and over 100 genetic markers on chromosome and plasmids were localized. In the *RtTA1* genome, five large regions and many separate potential polysaccharide biosynthesis genes were found. All regions showed high degree of sequence identity and conserved synteny between *R. leguminosarum* bv. *viciae* (*Rlv*) and *Rhizobium etli* (*Rhe*). Comparative analyses of genetic maps of close relatives *RtTA1*, *Rlv* and *Rhe* showed significant chromosome colinearity, providing a basis for comprehensive analysis of *R. leguminosarum* genome evolution.

#### Key words:

*Rhizobium*, genome map, synteny, comparative genomics.

#### Adres do korespondencji

Anna Skorupska,  
Zakład Mikrobiologii  
Ogólnej,  
Instytut Mikrobiologii  
i Biotechnologii,  
Uniwersytet Marii  
Curie-Skłodowskiej,  
ul. Akademicka 19,  
20-033 Lublin;  
e-mail:  
genet@biotop.umcs.lublin.pl

### 1. Symbioza *Rhizobium* z roślinami motylkowatymi

Glebowe bakterie symbiotyczne, nazywane ogólnie rizobiami, indukują na korzeniach lub łodygach roślin motylkowatych rozwój nowych organów, brodawek, w których zachodzi wiązanie azotu atmosferycznego. Rizobia, korzystając z roślinnych

związków organicznych, dostarczają im w zamian łatwo przyswajalnych, zredukowanych związków azotu, uniezależniając rośliny od nawożenia azotowego. Proces wiązania azotu atmosferycznego, przeprowadzany przez bakterie symbiotyczne oraz inne organizmy prokariotyczne, jest główną drogą wprowadzania zredukowanych związków azotu do biosfery. Globalny poziom biologicznego wiązania azotu określa się na 150-180 mln t, podczas gdy światowa produkcja przemysłowa nawozów azotowych wynosi około 80 mln t (1). Symbioza jest zatem wydajnym, naturalnym systemem biotechnologicznym pozwalającym na znaczne ograniczenie nawożenia azotowego i utrzymanie tzw. zrównoważonego rolnictwa. Badania nad bakteriami symbiotycznymi, poza znaczeniem poznawczym, mają również za zadanie podniesienie wydajności wiązania azotu w różnych układach symbiotycznych.

Pierwszym sygnałem w procesie nawiązania symbiozy są związki flawonoidowe wydzielane przez korzenie roślin motylkowatych, które wiążą się z bakteryjnym białkiem NodD należącym do rodziny regulatorów transkrypcyjnych LysR (2-4). Zaktwowane białko NodD wiąże się do promotorów symbiotycznych *nod*-box i indukuje transkrypcję genów *nod* (ang. *nodulation*) odpowiedzialnych za syntezę związków lipochitoooligosacharydowych, nazwanych czynnikami Nod. Bakteryjne czynniki Nod inicjują różnorodne odpowiedzi ze strony rośliny-gospodarza, z których najwcześniejszą jest pulsacyjna zmiana wewnątrzkomórkowego poziomu wapnia i zmiany cytoskieletu komórek włósnikowych korzenia, które prowadzą do zwinięcia włósnika, charakterystycznego dla zgodnego rozpoznania partnerów symbiozy. Równocześnie, czynnik Nod produkowany przez bakterie znajdujące się w zwiniętym włósniku stymuluje komórki kory korzenia do podziałów mitotycznych dając początek brodawce korzeniowej (primordium). Odpowiedź rośliny na czynniki Nod jest możliwa dzięki wielodomenowym białkom receptorowym należącym do rodziny LysM, a także dzięki innym białkom roślinnym uczestniczącym w transdukcji sygnałów, które są niezbędne zarówno do skręcenia włósnika, jak i do powstania tzw. nici infekcyjnej. Nić infekcyjna jest rurkowatą strukturą tworzoną przez roślinę, w której rizobia rozmnażają się i penetrują do głębszych warstw kory korzenia. Po dotarciu do primordium brodawki, bakterie są uwalniane z nici infekcyjnej do komórek roślinnych brodawki w procesie endocytozy i otoczone błoną peribakteroidalną, tworzą symbiosomy (5). W symbiosomach, w warunkach obniżonego stężenia tlenu, następuje indukcja ekspresji genów kompleksu nitrogenazy odpowiedzialnej za redukcję azotu cząsteczkowego. Zredukowane związki azotu są transportowane z symbiosomów do komórek roślinnych, natomiast rizobia korzystają z roślinnych produktów fotosyntezy tworząc mutualistyczny związek symbiotyczny.

Rozwój efektywnej symbiozy jest zależny nie tylko od roślinnych sygnałów molekularnych i bakteryjnych czynników Nod, ale także od bakteryjnych polisacharydów powierzchniowych. Co najmniej cztery typy polisacharydów, mają znaczenie w symbiozie: lipopolisacharydy (LPS) – strukturalne składniki błony zewnętrznej, kapsularne polisacharydy (CPS lub KPS) – zwykle związane z błoną zewnętrzną, egzopolisacharydy (EPS) – słabo lub w ogóle nie związane z błoną zewnętrzną oraz cykliczne



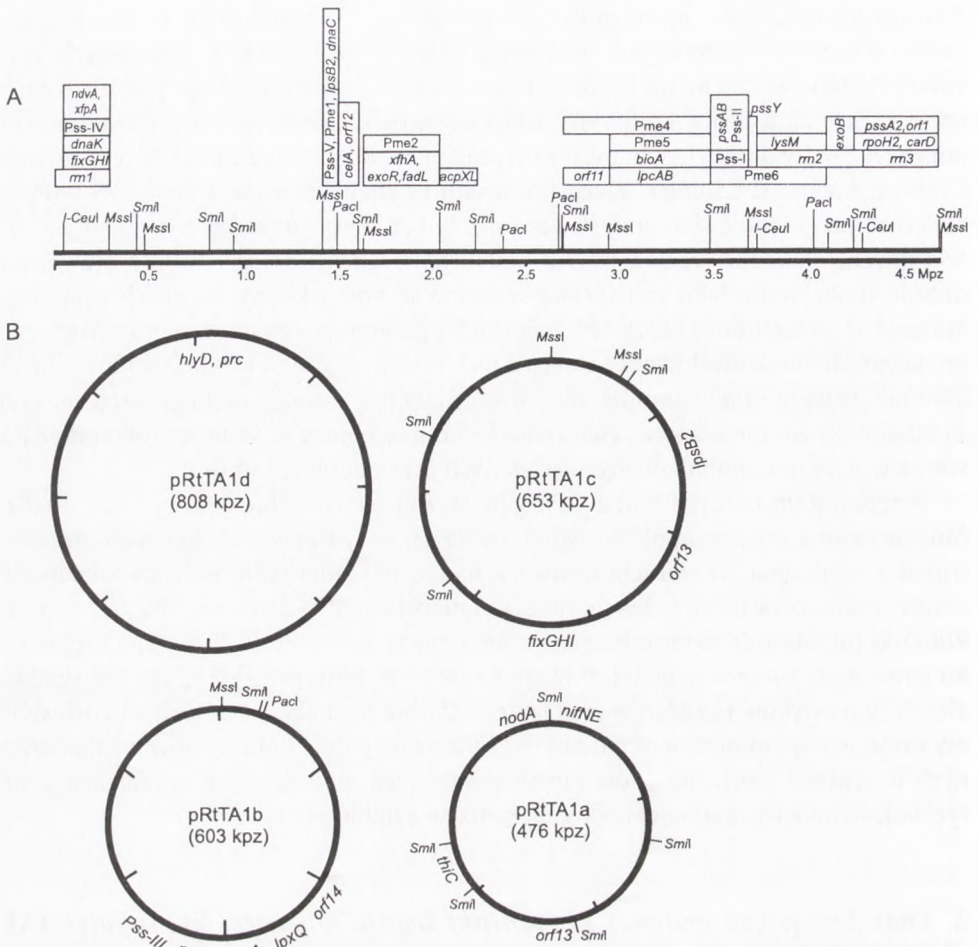
$\beta$ -(1,2)-glukany zlokalizowane głównie w peryplazmie komórki (4,5). Szczególne znaczenie w rozwoju symbiozy ma egzopolisacharyd, gdyż powstanie efektywnych, czyli zdolnych do wiązania azotu brodawek korzeniowych na ważnych gospodarczo roślinach, takich jak lucerna, koniczyna, wyka czy groch, zależy od prawidłowo funkcjonujących genów odpowiedzialnych za syntezę EPS. Rola biologiczna EPS jest złożona i nie do końca wyjaśniona. Zarówno struktura chemiczna jak i ilość EPS wpływa znacząco na przebieg symbiozy. Uważa się, że EPS bierze udział w tworzeniu nici infekcyjnych, unikaniu przez bakterie roślinnych reakcji obronnych oraz prawdopodobnie funkcjonuje jako cząsteczka sygnałna w początkowych etapach symbiozy. Mutacje w genach biosyntezy EPS najczęściej powodują zahamowanie rozwoju nici infekcyjnych i brak uwalniania bakterii z nici, w wyniku czego brodawka korzeniowa, albo nie zawiera symbiosomów, albo bakteroidy nie osiągają pełnego rozwoju i nie są zdolne do wiązania azotu. Znaczenie EPS dla symbiozy zostało szeroko przedstawione w ostatnio opublikowanych artykułach przeglądowych (4,6).

Przedmiotem naszych badań jest genetyczna kontrola biosyntezy EPS i analiza funkcji genów *pss* (ang. *polysaccharide synthesis*) w symbiozie *R. leguminosarum* bv. *trifolii* z koniczyną. W dotychczasowych badaniach zidentyfikowaliśmy kilkanaście genów zaangażowanych w biosyntezę, polimeryzację i wydzielanie EPS na zewnątrz komórki (6). Metodą ukierunkowanej mutagenyzy skonstruowaliśmy mutanty w kilku genach *pss* i poznaliśmy ich funkcję zarówno w biosyntezie EPS jak i w symbiozie. W tym artykule przedstawiamy nasze ostatnie badania, w których zastosowaliśmy podejście genomiczne dla identyfikacji możliwie dużej puli genów zaangażowanych w syntezę polisacharydów powierzchniowych, poznania ich organizacji genetycznej, lokalizacji genomowej i znaczenia w symbiozie.

## 2. Charakterystyka genomu *Rhizobium leguminosarum* bv. *trifolii* TA1 (*R<sup>t</sup>TA1*)

Gatunek *R. leguminosarum* zawiera trzy biotypy (ang. *biovar*): *trifolii*, który tworzy skrajnie specyficzny układ symbiotyczny z koniczyną, *viciae*, tworzący efektywną symbiozę m.in. z wyką i grochem oraz *phaseoli* zakażający fasolę. Ze zróżnicowanego genetycznie biotypu *R. leguminosarum* bv. *phaseoli* wyodrębniono gatunek *R. etli*, który również tworzy układ symbiotyczny z fasolą (7).

Wszystkie znane gatunki rizobiów mają duże genomy, złożone z kilku replikonów, tj. chromosomu oraz kilku megaplazmidów, które mogą stanowić nawet 40-50% całego genomu (8). Ostatnio określono sekwencję nukleotydową genomów kilku gatunków rizobiów i stwierdzono, że geny *nod* odpowiedzialne za tworzenie brodawek i geny *nif-fix* (ang. *nitrogen fixation*) odpowiedzialne za wiązanie azotu, są zwykle zlokalizowane na megaplazmidach, nazywanych plazmidami symbiotycznymi (pSym). Tak jest w przypadku *Sinorhizobium meliloti* (9), *Rhizobium* sp. NGR234 (10), *Rhizobium etli* (7), czy *R. leguminosarum* bv. *viciae* (11). Utrata plazmidu symbio-



Rys. 1. Genetyczna i fizyczna mapa genomu *R. leguminosarum* bv. *trifolii* TA1. A) Kolistą mapą chromosomu *RtTA1*, została otworzona w arbitralnym punkcie, tj. w miejscu restrykcyjnym *Mssl* (*Pme1*) pomiędzy operonami *rrn1* i *rrn3*. Markery genetyczne przypisano poprzez hybrydyzację Southerna do najmniejszego fragmentu restrykcyjnego. Kolejność markerów w grupie zlokalizowanej na tym samym fragmencie jest przypadkowa i nie odzwierciedla rzeczywistego porządku genów w danym regionie. B) Kolisty mapy plazmidów *RtTA1* (pRtTA1d – pRtTA1a). Wielkości plazmidów przedstawiono w skali (zmodyfikowane wg 17).

tycznego zawsze wiąże się z utratą zdolności do tworzenia brodawek na roślinach. W innych gatunkach rizobiów, takich jak *Mesorhizobium loti* (12) czy *Bradyrhizobium japonicum* (13), geny symbiotyczne są włączone do chromosomu tworząc tzw. wyspy symbiotyczne. W genomach kilku gatunków rizobiów zidentyfikowano duże zgrupowania genów odpowiedzialnych za biosyntezę polisacharydów powierzchni-



wych zlokalizowane na megaplazmidzie – pSymb w *S. meliloti* (14), lub na chromosomie – *M. loti* i *B. japonicum* (12,13).

Genom *RtTA1* składa się z 5 replikonów: kolistego chromosomu i czterech dużych plazmidów: pRtTA1a (497 kpz), pRtTA1b (598 kpz), pRtTA1c (646 kpz) i pRtTA1d (798 kpz) (15). Najmniejszy plazmid pRtTA1a został oznaczony jako plazmid symbiotyczny, pSym, ponieważ na tym plazmidzie zidentyfikowano geny regionu *nod-nif*, poprzez hybrydyzację Southerna z sondami zawierającymi geny symbiotyczne (rys. 1). Nieoczekiwanie, stwierdzono, że nie wszystkie geny symbiotyczne znajdują się na plazmidzie pSym. Dwie kopie genów *fixGHI* odpowiedzialnych za wiązanie azotu zlokalizowano na chromosomie i na niesymbiotycznym plazmidzie pRtTA1c. Plazmidy szczepu *RtTA1* mają prawdopodobnie system replikacji typu *repABC*, co stwierdzono przy użyciu heterologicznej sondy zawierającej gen *repC* pochodzącej z genomu *R. leguminosarum* bv. *viciae*. Replikon chromosomowy zidentyfikowano używając sond zawierających gen *rrs*, dla rybosomalnego 16S rRNA oraz gen *dnaK*, który koduje istotne dla bakterii białko opiekuńcze.

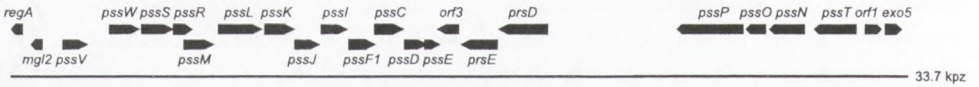
Identyfikacja i mapowanie genów *pss* okazało się bardzo trudne bez sporządzenia mapy restrykcyjnej genomu *RtTA1*. Biorąc pod uwagę wysoką zawartość par G+C w zsekwencionowanych genomach rizobiów [np. 64% G+C w *B. japonicum* (13), 61% G+C w *R. leguminosarum* bv. *viciae* (11)] do konstrukcji mapy fizycznej *RtTA1* wybrano rzadkotnące enzymy restrykcyjne rozpoznające sekwencje bogate w pary A+T oraz enzym *I-Ceurom*, który rozpoznaje konserwatywne 26 nt sekwencje w obrębie operonów dla rybosomalnego RNA (*rrn*) (16). Relatywnie niewielką liczbę fragmentów restrykcyjnych uzyskano używając enzymów *MssI* (*Pmel*), *PaclI*, *SmiI* (*Swal*) oraz *I-CeuI*. Na podstawie sumy średnich długości fragmentów restrykcyjnych genomowego DNA *RtTA1*, uzyskanych w wyniku trawienia poszczególnymi enzymami (pojedynczymi jak również w kombinacjach podwójnych) wielkość genomu *RtTA1* określono na  $7,257 \pm 8$  kpz (15). W dalszych etapach pracy, zidentyfikowano fragmenty makrorestrykcyjne należące do poszczególnych plazmidów, uporządkowano fragmenty chromosomu i sporządzono mapę fizyczną i genetyczną genomu *RtTA1*, na której zlokalizowano ponad 100 markerów genetycznych (rys. 1). Podobnie jak w innych genomach bakteryjnych (18,19), na chromosomie *RtTA1* w pobliżu miejsca startu replikacji (*ori*) zidentyfikowano trzy operony *rrn* oraz inne geny istotne dla podstawowego metabolizmu, takie jak *dnaK* czy *rpoH*.

### 3. Identyfikacja genów odpowiedzialnych za biosyntezę EPS

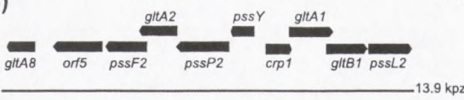
W genomach rizobiów o znanej sekwencji nukleotydowej, geny odpowiedzialne za biosyntezę polisacharydów powierzchniowych zwykle są rozproszone (7,10, 11). W celu identyfikacji tej grupy genów w *RtTA1* skonstruowano bibliotekę genomową w wektorze pIndigoBac5/*oriV* (20), która zawierała 2591 klonów o średniej wielkości wstawki ~40 kpz. Metodami hybrydyzacji z sondami zawierającymi



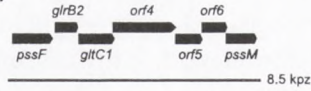
A)



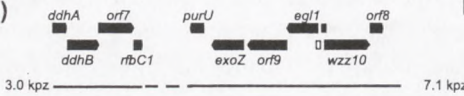
B)



C)



D)



E)



Rys. 2. Genetyczna mapa regionów Pss *R. leguminosarum* bv. *trifolii* TA1. A) Pss-I, B) Pss-II, C) Pss-III, D) Pss-IV, E) Pss-V (zmodyfikowane wg 15).

poprzednio opisane geny *pss*, a także poprzez amplifikację genów potencjalnie związanych z produkcją EPS, metodą PCR ze starterami zaprojektowanymi na podstawie sekwencji nukleotydowych opublikowanych genomów *R. leguminosarum* bv. *viciae* (*Rlv*) i *R. etli* (*Rhe*), zidentyfikowaliśmy 5 dużych zgrupowań genów zaangażowanych w syntezę powierzchniowych polisacharydów (rys. 2), a także kilkanaście genów rozproszonych w genomie i nie tworzących większych zgrupowań (rys. 1).

Największy region nazwany Pss-I, wielkości 33,7 kbp, zlokalizowany na chromosomie zawierał 24 geny, z których większość to geny wcześniej opisane przez nas: *pssT*, *pssNOP*, *prsDEorf3*, *pssCDE* oraz *pssL* (21-26), lub przez innych autorów (27-29), kodujące glikozylotransferazy, endoglikanazy, enzymy modyfikujące podstawową podjednostkę EPS poprzez przyłączanie niecukrowych podstawników oraz białka odpowiedzialne za polimeryzację i wyprowadzanie EPS poza komórkę (rys. 2A). Ten region uznano za główne zgrupowanie genów odpowiedzialnych za biosyntezę egzopolisacharydu w *R. leguminosarum* bv. *trifolii*, ponieważ opisano w nim szereg mutacji, których efektem fenotypowym był brak syntezy EPS (6,27-29).

Chromosomalny region Pss-II (14 kbp), zawiera 10 genów (*gltA8*, *orf5*, *pssF2*, *gltA2*, *pssP2*, *pssY*, *crp1*, *gltA1*, *gltB1*, *pssL2*), które w większości kodują potencjalne glikozylotransferazy o podobieństwie do białek związanych z biosyntezą i polimeryzacją lipopolisacharydu (LPS) (rys. 2B). Produkt genu *pssL2* jest podobny do translokazy WzxC. Gen *pssF2* koduje hipotetyczne białko błony zewnętrznej (OMA, ang. *outer membrane auxiliary*) zaangażowane w biosyntezę lub/i transport polisacharydów. Gen *pssP2* koduje białko należące do rodziny białek Wzc odpowiedzialnych za transport i polimeryzację polisacharydów.

Region Pss-III jest jedynym do tej pory zidentyfikowanym plazmidowym regionem związanym z biosyntezą powierzchniowych polisacharydów *RtTA1* (rys. 2C). Zawiera 7 genów: *pssF*, *gltB2*, *gltC1*, *orf4*, *orf5*, *orf6* i *pssM*, których produkty białkowe to przede wszystkim enzymy należące do rodziny glikozylotransferaz. PssF zaliczono do grupy białek błony zewnętrznej OMA.

Region Pss-IV (~ 10 kpz) zlokalizowany na chromosomie, jest podzielony na dwa oddzielne skupiska genów oznaczone odpowiednio Pss-IVA (3 kpz), który zawiera 4 geny: *ddhA*, *ddhB*, *orf7* i *rbcC1* oraz Pss-IVB (7,1 kpz) z genami *purU*, *exoZ*, *orf9*, *egl1*, *wzz10*, *orf8* (rys. 2D). Przewidywana komputerowo funkcja białek kodowanych przez geny Pss-IVA, sugeruje udział tego regionu w biosyntezie LPS *RtTA1*. W regionie Pss-IVB znajduje się m.in. acetylotransferaza *ExoZ*, enzym odpowiedzialny m.in. za modyfikacje niecukrowe polisacharydów, *Wzz10* uczestniczące w transporcie i/lub polimeryzacji polisacharydów (rodzina białek *Wzz*) oraz enzym *Egl1* o potencjalnej funkcji endoglukanazy. Lokalizacja tego regionu w pobliżu Pss-IVA oraz przewidywane funkcje genów tego regionu wskazuje, że również Pss-IVB może być związany z biosyntezą LPS.

Chromosomalny region Pss-V (*rfb*) (3 kpz) zawiera 4 geny *rfbADBC* (rys. 2E). Na podstawie podobieństwa białek *RfbADB* do *RmlADB* z *Escherichia coli* oraz *RfaC* do *RmlC* z *Shigella boydii* zaangażowanych w biosyntezę części O-swoistej LPS, można z dużym prawdopodobieństwem przypuszczać, że Pss-V jest kolejnym regionem związanym z biosyntezą LPS w *RtTA1*.

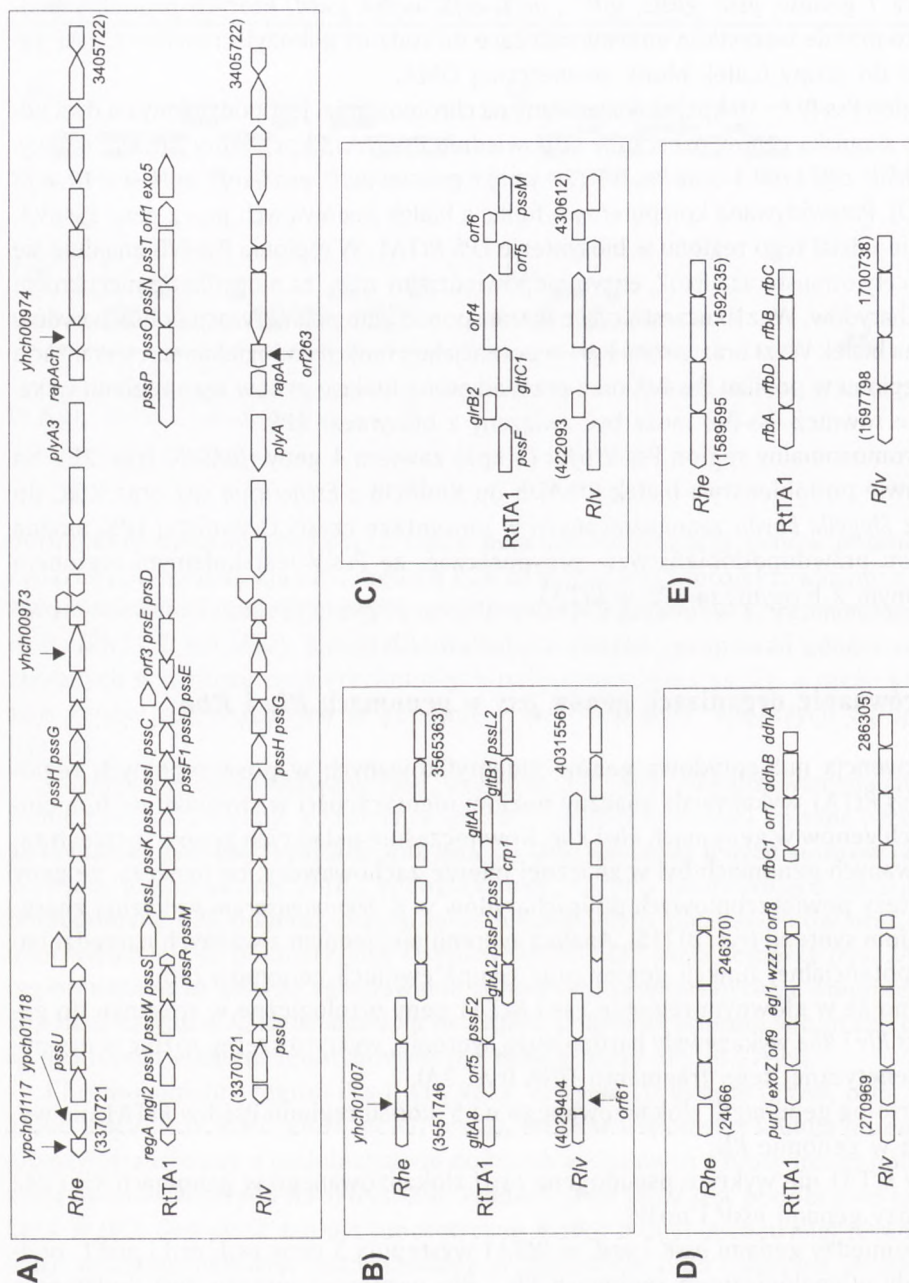
#### 4. Porównanie organizacji genów *pss* w genomach *Rlv* i *Rhe*

Sekwencja nukleotydowa genów zidentyfikowanych w poszczególnych regionach Pss *RtTA1* wykazywała znaczny poziom identyczności w stosunku do homologicznych genów w genomach *Rlv* i *Rhe*. Równocześnie układ tych genów w trzech zanalizowanych genomach był w znacznej mierze zachowawczy, co oznacza, że geny biosyntezy powierzchniowych polisacharydów w *R. leguminosarum* wykazują znaczny poziom syntenii (rys. 3) (15). Analiza syntenii jest jednym z ważnych narzędzi badania potencjalnej funkcji genów oraz tempa ewolucji genomów (30).

Mimo że w głównym regionie Pss-I *RtTA1* geny ortologiczne w stosunku do genów *pss* *Rlv* i *Rhe* wykazywały bardzo dużą syntenię, wykryto szereg różnic w organizacji genetycznej tego fragmentu DNA (rys. 3A):

- ortolog genu *mgI2* zlokalizowanego na 5' końcu regionu Pss-I w *RtTA1* nie występuje w genomie *Rlv*;
- w *RtTA1* nie wykryto pseudogenu *pssU* zlokalizowanego w genomach *Rlv* i *Rhe* pomiędzy genami *pssV* i *pssW*;
- pomiędzy genami *pssK* i *pssC* w *RtTA1* występują 3 geny *pssJ*, *pssI* i *pssF1*, podczas gdy w analogicznych regionach *Rlv* i *Rhe* występują jeszcze dwa dodatkowe geny *pssG* i *pssF* kodujące glikozylotransferazy, których nie wykryto w całym genomie *RtTA1*;





Rys. 3. Syntenia regionów Pss RTA1, Rhe i Rlv. W nawiasach podano pozycje regionów Pss na odpowiednich replikonach w Rhe oraz Rlv. A) Pss-I, B) Pss-II, C) Pss-III, D) Pss-IV, E) Pss-V (zmodyfikowane wg 15).



– w genomie *Rhe* poniżej genu *pssC* występuje ORF o przypuszczalnej funkcji WzyC polimerazy O-antygeny. Genu tego nie wykryto w *RtTA1* i *Rlv*;

– w regionie pomiędzy genami *prsD* i *pssP* (4,4 kpz) w *Rlv* i *Rhe* występują geny *plyA* i *rapA* kodujące odpowiednio, endoglukanazę i białko adhezyjne (27). W analogicznym regionie *RtTA1* stwierdzono jedynie obecność szczątkowych ORF-ów wykazujących podobieństwo do *plyA* i *rapA*.

Opisane w dwóch gatunkach *R. leguminosarum* i *R. etli* różnice w sekwencji regionów skupiających główne geny biosyntezy EPS mogą wskazywać zarówno na unikatowość gatunkową zewnątrzkomórkowego EPS, jak i na jego swoistą rolę w procesie symbiozy.

Porównanie organizacji Pss-II z odpowiednimi regionami w *Rlv* i *Rhe* wykazało znaczny poziom syntenii (rys. 3B). Jedyna różnica dotyczy braku w genomie *RtTA1* otwartej ramki odczytu występującej pomiędzy genami *gltA8* i *orf5* w genomach *Rlv* i *Rhe*. Plazmidowy region Pss-III wykazuje 95% identyczności sekwencji nukleotydowej i absolutną syntenie w porównaniu z odpowiednim regionem zlokalizowanym na plazmidzie pRL11 w *Rlv* (rys. 3C). Co ciekawe, w genomie *Rhe* nie wykryto segmentu odpowiadającego Pss-III na żadnym z replikonów.

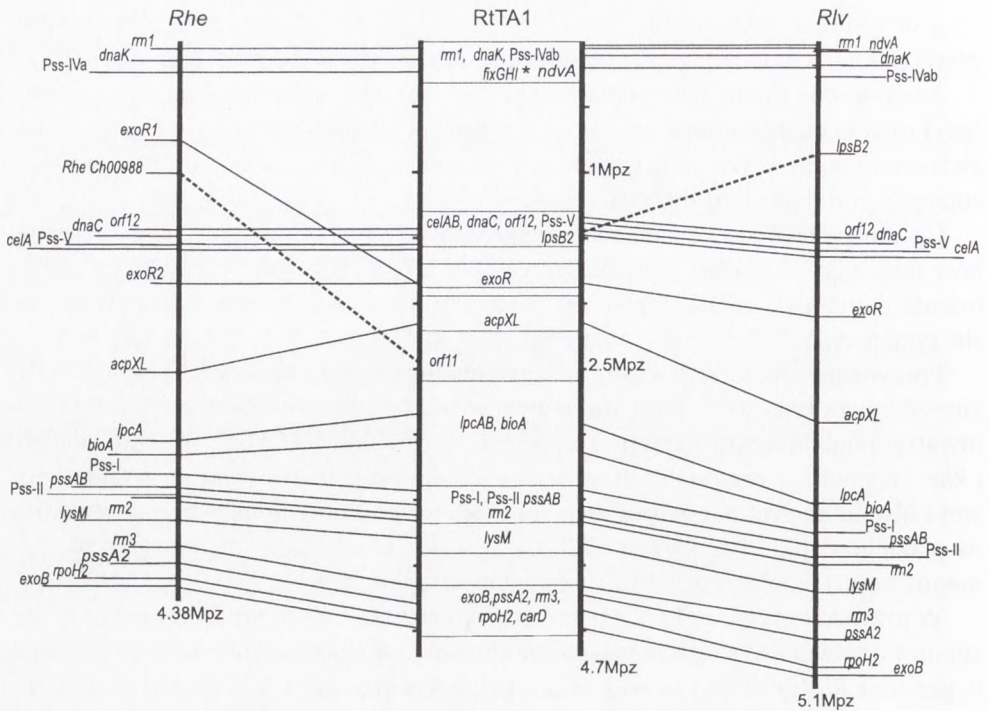
W przypadku regionu Pss-IV podzielonego na część A i B, ortologi Pss-IVA o wysokim podobieństwie i zachowawczym ułożeniu na chromosomie wykryto jedynie w genomie *Rlv* (rys. 3D). Jest to więc drugi (po Pss-III), region biosyntezy powierzchniowych polisacharydów unikatowy dla *RtTA1* i *Rlv*. Z kolei ortologi regionu Pss-IVB wykryto zarówno w genomie *Rlv* jak i *Rhe*.

Organizacja regionu Pss-V wykazuje niemal absolutną syntenie, oraz powyżej 95% identyczności sekwencji DNA z odpowiednimi odcinkami genomu *Rlv* i *Rhe* (rys. 3E).

Należy podkreślić, że zidentyfikowane w *RtTA1* regiony Pss-I – V nie wykazują syntenii w porównaniu do innych genomów, (poza gatunkami *Rlv* i *Rhe*) jeszcze raz wskazując na unikatowość polisacharydowych struktur powierzchniowych rizobów. Przeprowadzone analizy genomu *RtTA1* wykazały obfitość regionów związanych z powierzchniowymi polisacharydami, zwłaszcza genów biosyntezy LPS. Nie tworzą one jednak wspólnego dużego zgrupowania na żadnym z replikonów i są rozproszone głównie na chromosomie *RtTA1*.

Young i wsp. (11) zaproponowali analizę zawartości %G+C, a szczególnie %G+C w trzeciej, zmiennej pozycji kodonu (GC3S), w celu przypisania poszczególnych regionów DNA do tzw. rdzeniowych („core”) lub dodatkowych („accessory”) elementów genomu. Część rdzeniowa genomu, zwykle chromosomalna, ma wyższą wartość %G+C, podczas gdy dodatkowe elementy genomu mają niższą zawartość %G+C i są zlokalizowane głównie na plazmidach i na wyspach chromosomalnych. Geny określane jako rdzeniowe wykazują zwykle większą tendencję do zachowawczości sekwencji i organizacji genetycznej (syntenii), co wskazuje na ich istotną rolę w metabolizmie komórki bakteryjnej. Dodatkowe elementy genomu stanowią tę część informacji genetycznej, która jest wykorzystywana przez komórkę tylko oka-





Rys. 4. Porównanie organizacji chromosomów trzech blisko spokrewnionych bakterii *RtTA1*, *Rlv* i *Rhe*. Geny *RtTA1*, których homologi występują na chromosomie zarówno *Rlv* jak i *Rhe*, połączono ciągłymi liniami. *RtTA1 lpsB2* i *orf11*, których homologi znaleziono na chromosomie tylko jednej bakterii połączono liniami przerywanymi. Chromosomalne markery genetyczne specyficzne tylko dla *RtTA1* oznaczono gwiazdką. Kolistą mapą chromosomu *RtTA1* została przekształcona w formę liniową dla ułatwienia porównania organizacji (zmodyfikowane wg 17).

zjonalnie, w zależności od warunków środowiskowych, w jakich znajduje się bakteria (11).

Analizę zawartości procentowej G+C i GC3/S przeprowadzono dla regionów i skupisk genów *pss RtTA1* i odpowiadających im regionów w genomach *Rlv* i *Rhe* (tab.). Stwierdzono, że chromosomalne regiony *Pss-II*, *Pss-IV*, *Pss-V* wykazują wartości tych parametrów typowe dla tzw. rdzeniowych elementów genomu, podczas gdy parametry plazmidowego regionu *Pss-III* wskazują na typowy charakter dodatkowego elementu genomu *RtTA1*. Co ciekawe, w przypadku chromosomalnego regionu *Pss-I*, który jest wysoce zachowawczy i znajduje się we wszystkich trzech genomach, wartości %GC i GC3S, są niższe niż plazmidowego regionu *Pss-III*. Takie cechy, w przypadku chromosomalnych, rdzeniowych części genomu, mogą wskazywać, że region *Pss-I* jest wyspą genomową, która została przeniesiona do chromosomu z innego replikonu np. przez elementy mobilne, bakteriofaga albo w wyniku wewnętrznych rearanżacji genomu, np. rekombinacji pomiędzy plazmidem i chromosomem.



Tabela

Zawartość G+C oraz GC3S<sup>a</sup> regionów Pss zlokalizowanych na chromosomie i na plazmidach *RtTA1*, *Rlv* i *Rhe*

Nazwa regionu/lokalizacja	<i>RtTA1</i>	<i>Rlv</i>	<i>Rhe</i>
	% G+C/GC3S	% G+C/GC3S	% G+C/GC3S
Pss-I / chromosom	56/54.1	56.8/55.3	57.5/56.7
Pss-II / chromosom	59.8/59.0	59.9/60.5	64.6/60.1
Pss-III / plazmid	60.0/57.3	60.1/56.2	n.d.
Pss-IVA / chromosom	63.5/66.2	63.7/64.4	n.d.
Pss-IVB / chromosom	61.4/61.9	62.5/60.7	61.7/62.3
Pss-V / chromosom	61.6/66.1	60.9/64.0	60.6/61.7

<sup>a</sup> GC3S – zawartość G+C w trzeciej pozycji kodonu; n.d. – region nie wykryty w tym gatunku bakterii

## 5. Analiza porównawcza genomów *RtTA1*, *Rlv* i *Rhe*

Na podstawie analizy genów odpowiedzialnych za biosyntezę polisacharydów powierzchniowych zlokalizowanych na chromosomie wskazano na znaczną zachowawczość ich organizacji. Badania rozszerzono o porównanie organizacji całych chromosomów trzech blisko spokrewnionych bakterii, tj. *RtTA1*, *Rlv* i *Rhe* (rys. 4) wykorzystując mapę makrorestrykcyjną i genetyczną chromosomu *RtTA1* (rys. 1). Taki sposób podejścia do genomiki porównawczej okazał się bardzo użyteczny i pozwolił na wysunięcie kilku hipotez dotyczących przypuszczalnych zmian ewolucyjnych tych blisko spokrewnionych bakterii oraz wykazanie pewnego stopnia plastyczności genomów rizobiów (31-33). Badanie podobieństwa układu genów w genomach, czyli kolinearności genomów, jest jedną z metod określania ich relacji ewolucyjnych (34,35). Analiza porównawcza map genetycznych bakterii należących do tego samego rodzaju lub gatunku, zwana genomiką mikroewolucyjną, pozwala uzyskać cenne informacje dotyczące zachowania określonego układu genów na chromosomach, a także wskazuje na mechanizmy genetyczne odpowiedzialne za rearanżacje genomowe (36-38). W przypadku genomów rizobiowych ocena stopnia ich kolinearności jest szczególnie ciekawa, ze względu na ich niestabilność i skłonność do rearanżacji zarówno na skutek wewnętrznych przegrupowań, jak i na drodze horyzontalnego transferu DNA (33,38).

W pierwszym etapie, wykorzystując program WebACT, porównano sekwencje chromosomów *Rlv* i *Rhe*, wykazując ich znaczną kolinearność na całej długości. Niewielkie różnice obserwowane jako przerwy zaburzające kolinearność, mogły wynikać m.in. z różnic wielkości obydwu chromosomów, gdyż chromosom *Rlv* jest większy od *Rhe* o 675 kpz. Porównanie sekwencji plazmidów *Rlv* i *Rhe* wykazało zupełny brak kolinearności, co potwierdza fakt, że jest to bardzo zmienna część genomu, która może również przyczyniać się do rearanżacji w obrębie chromosomu, np. poprzez homologiczną rekombinację pomiędzy sekwencjami powtórzonymi (34,39).



W kolejnym etapie, markery genetyczne umieszczone na mapie chromosomu *RtTA1* zlokalizowano w sekwencji chromosomów *Rlv* i *Rhe* (rys. 4). Porównanie tych odmiennie skonstruowanych map ponownie wykazało znaczną kolinearność analizowanych replikonów. Rozmieszczenie markerów genetycznych w badanych chromosomach zostało w dużej mierze zachowane w ewolucji, podczas gdy w sekwencji DNA wystąpiły zmiany, na co wskazuje odmienny profil makrorestrykcyjny chromosomów *RtTA1*, *Rlv* i *Rhe*. Przypuszczalnie, pewne czynniki selekcyjne, sprzyjają raczej substytucjom pojedynczych nukleotydów niż przegrupowaniom dużych fragmentów chromosomów. Powoduje to różnice wzorców makrorestrykcyjnych przy zachowaniu porządku układu genów chromosomalnych.

Pomimo znacznej kolinearności stwierdzono także pewne różnice w organizacji badanych chromosomów, np. w pobliżu centralnej części chromosomu obejmującej fragmenty restrykcyjne uzyskane po trawieniu enzymem *MssI* (rys. 1A, 4). W tym obszarze, na chromosomie *Rhe* znajduje się jedna z dwóch kopii genu regulatorowego *exoR*, a na chromosomie *Rlv* paralog *lpsβ2*. W *RtTA1* nie wykryto dodatkowej kopii genu *exoR*, a chromosomalna kopia *lpsB2* jest nieznacznie przesunięta w porównaniu z *Rlv* (rys. 4). Obecność *lpsB2* w tym regionie, i jego drugiej kopii na plazmidzie p*RtTA1c*, może oznaczać, że obserwowane zaburzenie kolinearności jest efektem rearanżacji chromosomu z udziałem plazmidowych powtórzeń genów. Różnice organizacji dotyczą także lokalizacji genu *acpXL*, który jest przesunięty na chromosomie *RtTA1* w porównaniu do *Rhe* i *Rlv* (rys. 4).

Ważnym czynnikiem kształtującym genomy bakterii, poza transferem horyzontalnym genów, jest ich utrata lub duplikacja (40,41). Duplikacja genów może zachodzić jako odpowiedź na różne czynniki selekcyjne zmiennego środowiska, jakim jest np. gleba. W genomach rizobiów również można obserwować taką zmienność (41). W *RtTA1* zidentyfikowano wiele paralogów, głównie w genach biosyntezy polisacharydów (np. *pssP* i *pssP2*, *pssA* i *pssA2*, powtórzenia *lpsB*), oraz w genach symbiotycznych (np. *fixGHI*). Porównanie organizacji plazmidów nie było możliwe ze względu na zbyt małą liczbę markerów genetycznych zlokalizowanych na megaplazmidach *RtTA1*.

## 6. Uwagi końcowe

Genomy bakterii glebowych, w tym rizobiów, są duże i zwykle podzielone na kilka replikonów: chromosom i megaplazmidy stanowiące znaczną część genomu. Analiza porównawcza genomów trzech blisko spokrewnionych bakterii należących poprzednio do trzech biotypów *R. leguminosarum*, o różnej specyficzności zakażenia roślin motylkowatych, pozwoliła na wysunięcie kilku ogólnych uwag. Obserwowana zachowawczość porządku genów w chromosomach *RtTA1*, *Rlv* i *Rhe* wskazuje na ich wspólne pochodzenie i jednocześnie umiarkowaną plastyczność, przynajmniej na poziomie chromosomu. Przypuszczalnie większość obserwowanych istotnych zmian



w tych genomach zaszła w ewolucji przed ich rozdzieleniem na trzy biotypy. Jednakże różnice wielkości genomów i podziału informacji genetycznej pomiędzy poszczególne replikony sugerują ich odmienne drogi ewolucji, z prawdopodobnym udziałem pozachromosomalnego DNA. Można przypuszczać, że zróżnicowana architektura genomów *RtTA1*, *Rlv* i *Rhe*, zwłaszcza w obrębie plazmidowego DNA, przyczyniła się do odmiennych dla każdego biotypu zdolności adaptacyjnych, w tym również różnic w swoistości symbiotycznych oddziaływań z partnerem roślinnym. Można przypuszczać, że nie bez wpływu na te właściwości pozostaje zarówno obfitość genów biosyntezy powierzchniowych polisacharydów, jak również różnice w regionach skupiających główne geny biosyntezy EPS, które decydują o unikatowości gatunkowej zewnątrzkomórkowego EPS, a także o jego swoistej roli w symbiozie. Chromosomalna lokalizacja większości zidentyfikowanych regionów biosyntezy powierzchniowych polisacharydów dodatkowo podkreśla istotność tych genów zarówno w glebowym jak i symbiotycznym stylu życia rizobiów.

Praca finansowana z projektu 2 P04A 034 26 Ministerstwa Nauki i Szkolnictwa Wyższego.

## Literatura

1. Bezdicsek D. F., Kennedy A. C., (1998), *Microorganisms in Action*, Eds. Lynch J. M., Hobbie J. E. Blackwell Scientific Publications, Oxford, UK.
2. Perret X., Staehelin C., Broughton W., (2000), *Microbiol. Mol. Biol. Rev.*, 64, 180-201.
3. Spaink H. P., (2000), *Ann. Rev. Microbiol.*, 54, 257-288.
4. Jones K. M., Kobayashi H., Davies B. W., Taga M. E., Walker G. C., (2007), *Nature Rev. Microbiol.*, 5, 619-633.
5. Margaert P., Uchiumi T., Alunni B., Evanno G., Cheron A., Catrice O., Mausset A., Barloy-Hubler F., Galibert F., Kondorosi A., Kondorosi E., (2006), *Proc. Natl. Acad. Sci. USA*, 103, 5230-5235.
6. Skorupska A., Janczarek M., Marczak M., Mazur A., Król J., (2006), *Microb. Cell Fact.*, 16, 7.
7. Gonzalez V., Santamaria R. I., Bustos P., Hernandez-Gonzalez I., Medrano-Soto A., Moreno-Hagelsieb G., Janga S. C., Ramirez M. A., Jimenez-Jacinto V., Collado-Vides J., Davila G., (2006), *Proc. Natl. Acad. Sci. USA*, 103, 3834-3839.
8. Honeycutt R. J., McClelland B. W., (1993), *J. Bacteriol.*, 175, 6945-6952.
9. Barnett M. J., Fisher R. F., Jones T., Komp C., Pia Abola A., Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J., Gurjal M., Hong A., Huizar L., Hyman R. W., Kahn D., Kahn M. L., Kalman S., Keating D. H., Palm C., Peck M. C., Surzycki R., Wells D. H., Yeh K-Ch., Davis R. W., Federespiel N. A., Long S. R., (2001), *Proc. Natl. Acad. Sci. USA*, 98, 9883-9888.
10. Freiberg C., Fellay R., Bairoch A., Broughton W. J., Rosenthal A., Perret X., (1997), *Nature*, 387, 394-401.
11. Young J. P., Crossman L. C., Johnston A. W., Thomson N. R., Ghazoui Z. F., Hull K. H., Wexler M., Curson A. R., Todd J. D., Poole P. S., Mauchline T. H., East A. K., Quail M. A., Churcher C., Arrowsmith C., Cherevach I., Chillingworth T., Clarke K., Cronin A., Davis P., Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Mungall K., Norbertczak H., Rabinowitsch E., Sanders M., Simmonds M., Whitehead S., Parkhill J., (2006), *Genome Biol.*, 7, R34.
12. Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S., (2000), *DNA Res.*, 7, 331-338.



13. Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T., Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K., Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M., Tabata S., (2002), *DNA Res.*, 9, 189-197.
14. Finan T. M., Weider S., Wong K., Buhrmester J., Chain P., Vorholter F. J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J., Golding B., Pühler A., (2001), *Proc. Natl. Acad. Sci. USA*, 98, 9889-9894.
15. Król J. E., Mazur A., Marczak M., Skorupska A., (2007), *Genomics*, 89, 237-247.
16. Liu S. L., Hessel A., Sanderson K. E., (1993), *Proc. Natl. Acad. Sci. USA*, 90, 6874-6878.
17. Król J. E., Mazur A., Marczak M., Skorupska A., (2007), *Mol. Genet Genomics*, DOI 10.1007/s00438-007-0299-9.
18. Le Bourgeois P., Daveran-Mingot M. L., Ritzenthaler P., (2000), *J. Bacteriol.*, 182, 2481-2491.
19. Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J., Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S., Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D., Pohl T., Portetelle D., Puhler A., Purnelle B., Ramsperger U., Renard C., Thebault P., Vandenberg M., Weidner S., Galibert F., (2001), *Proc. Natl. Acad. Sci. USA*, 98, 9877-9882.
20. Wild J., Hradecna W., Szybalski W., (2002), *Genome Res.*, 12, 1434-1444.
21. Król J., Skorupska A., (1997), *Microbiology*, 143, 1389-1394.
22. Król J., Wielbo J., Mazur A., Kopcińska J., Łotocka B., Golinowski W., Skorupska A., (1998), *Mol. Plant Microbe Interact.*, 11, 1142-1148.
23. Mazur A., Król J. E., Wielbo J., Urbanik-Sypniewska T., Skorupska A., (2002), *Mol. Plant Microbe Interact.*, 15, 388-397.
24. Mazur A., Krol J., Marczak M., Skorupska A., (2003), *J. Bacteriol.*, 185, 2503-2511.
25. Mazur A., Marczak M., Król J. E., Skorupska A., (2005), *Arch. Microbiol.*, 184, 1-10.
26. Marczak M., Mazur A., Krol J. E., Gruszecki W. I., Skorupska A., (2006), *J. Bacteriol.*, 188, 6943-6952.
27. Finnie C., Zorreguieta A., Hartley N. M., Downie J. A., (1998), *J. Bacteriol.*, 180, 1691-1699.
28. Ivashina T. V., Khmelitsky M. I., Shlyapnikov M. G., Kanapin A. A., Ksenzenko V. N., (1994), *Gene*, 150, 111-116.
29. van Workum W. A. T., Canter Cremers H. C. J., Wijfies A. H. M., van der Kolk C., Wijffelman C. A., Kijne J. W., (1997), *Mol. Plant Microbe Interact.*, 10, 290-301.
30. Guerrero G., Peralta H., Aguilar A., Diaz R., Villalobos M. A., Medrano-Soto A., Mora J., (2005), *BMC Evol. Biol.*, 5, 55-74.
31. Mavingui P., Flores M., Guo X., Davila G., Perret X., Broughton W. J., Palacios R., (2002), *J. Bacteriol.*, 184, 171-176.
32. Davila G., Gonzalez V., Ramirez-Romero M. A., Rodriguez O., (2005), *Genomes and genomics of nitrogen-fixing organisms*, Eds. Palacios R., Newton W. E., 133-142, Springer, Dordrecht.
33. Palacios R., Flores M., (2005), *Genomes and genomics of nitrogen-fixing organisms*, Eds. Palacios R., Newton W. E., 183-200, Springer, Dordrecht.
34. Huyen M. A., Bork P., (1997), *Proc. Natl. Acad. Sci. USA*, 95, 5849-5856.
35. Dempsey J. A., Wallace A. B., Cannon J. G., (1995), *J. Bacteriol.*, 177, 6390-6400.
36. Romling U., Schmidt K. D., Tummler B., (1997), *J. Mol. Biol.*, 271, 386-404.
37. Jordan I. K., Rogozin I. B., Wolf Y. I., Koonin E. V., (2002), *Theor. Popul. Biol.*, 61, 435-447.
38. Brom S., Garcia de los Santos A., de Lourdes Girard M., Davila G., Palacios R., Romero D., (1991), *J. Bacteriol.*, 173, 1344-1346.
39. Eisen J. A., (2000), *Curr. Opin. Microbiol.*, 3, 475-480.
40. Hooper S. D., Berg O. G., (2003), *Mol. Biol. Evol.*, 20, 945-954.
41. Gevers D., Vandepoele K., Simillon C., van de Peer C., (2004), *Trends Microbiol.*, 12, 148-154.