



Wyznaczenie miejsca inicjacji replikacji na podstawie sekwencji chromosomu bakteryjnego

Jolanta Zakrzewska-Czerwińska¹, Paweł Mackiewicz²,
Anna Zawilak¹, Stanisław Cebrat²

¹Zakład Mikrobiologii, Instytut Immunologii i Terapii Doświadczalnej
im. L. Hirszfelda, Polska Akademia Nauk, Wrocław

²Zakład Genomiki, Instytut Genetyki i Mikrobiologii,
Uniwersytet Wrocławski, Wrocław

Determination of replication initiation site

Summary

Initiation of bacterial chromosome replication is mediated by a single initiator protein – DnaA which interacts specifically with multiple DnaA boxes located within the origin of replication *oriC*. We have applied *in silico* methods: DNA asymmetry, DnaA box distribution and *dnaA* gene location to identify the putative replication origins in bacterial chromosomes. The three methods identify the same region as a putative origin in more than half of the analyzed chromosomes. The most universal method of putative *oriC* identification in bacterial chromosomes is DNA asymmetry, although in some cases it is necessary to apply all three methods. Interestingly, most bacterial chromosomes exhibit an overrepresentation of DnaA boxes; they contain at least one cluster of DnaA boxes in the vicinity of the *oriC* region that is probably involved in controlling replication initiation.

Key words:

initiation of replication, bacterial chromosome, DnaA protein, *oriC* region, DNA asymmetry.

1. Wstęp

Podział każdej komórki pro- i eukariotycznej musi być poprzedzony podwojeniem materiału genetycznego, tak by komórki potomne otrzymały pełen zestaw genów (1,2). Replikacja DNA

Adres do korespondencji

Jolanta
Zakrzewska-Czerwińska,
Zakład Mikrobiologii,
Instytut Immunologii
i Terapii Doświadczalnej
im. L. Hirszfelda,
Polska Akademia Nauk,
ul. Weigla 12,
53-114 Wrocław;
e-mail: zakrzew@iitd.
pan.wroc.pl

zapewnia wierne przekazywanie informacji genetycznej z komórek rodzicielskich do komórek potomnych. W przeprowadzonych badaniach nad inicjacją replikacji bakteryjnych chromosomów oraz chromosomów archebakterii i eukariontów pokazano, że mechanizm inicjacji u organizmów nawet tak ewolucyjnie odległych, jak bakterie i ssaki, przebiega w podobny sposób. Polega on na związaniu się inicjatorowego białka lub białek z regionem inicjacji *ori* (*origin*), co prowadzi do utworzenia kompleksu nukleoproteinowego, a następnie do miejscowego rozplecenia helisy DNA i powstania oczka replikacyjnego (3). W miejscach rozdzielania dwóch nici DNA tworzy się para widełek replikacyjnych, w których syntetyzowane są nowe nici DNA na matrycy rodzicielskiego DNA. Różnice w przebiegu procesu inicjacji replikacji DNA dotyczą liczby miejsc *ori* i ich cech strukturalnych oraz liczby i rodzaju białek uczestniczących w inicjacji replikacji.

U bakterii, replikacja rozpoczyna się w jednym określonym miejscu, natomiast u eukariontów występuje wiele miejsc inicjacji replikacji – nawet kilka tysięcy *ori*. U archebakterii synteza DNA jest inicjowana z jednego (np. u *Pyrococcus abyssi*, (4)) lub z dwóch (np. u *Sulfolobus solfataricus*, (5)) miejsc na chromosomie. Region *ori* u bakterii i archebakterii jest swoście rozpoznawany przez jedno białko, natomiast u eukariontów przez kompleks kilku inicjatorowych białek (3).

Głównymi elementami inicjacji replikacji u bakterii są: inicjatorowe białko DnaA i region *oriC* (*origin of chromosomal replication*), w którym rozpoczyna się replikacja chromosomu (6,7). Proces replikacji DNA został najlepiej poznany u *Escherichia coli* – modelowego organizmu w biologii molekularnej mikroorganizmów. Białko DnaA, wiążąc się z regionem *oriC* tworzy rdzeń, wokół którego owija się podwójna nić DNA. W wyniku powstałych naprężeń dochodzi do miejscowego rozplecenia regionu *oriC* i utworzenia oczka replikacyjnego.

Proces replikacji DNA jest regulowany na poziomie inicjacji replikacji. Inicjacja replikacji jest, jak się wydaje, najważniejszym etapem kontrolującym częstość podziałów komórkowych, a tym samym tempo wzrostu komórki i dlatego istotne jest określenie reguł umożliwiających wyznaczenie regionu *oriC* na chromosomie bakteryjnym.

Do niedawna, regionów *oriC* poszukiwano za pomocą metod *in vivo*: analizy replikacji minichromosomów i badania wbudowywania analogów nukleotydów, zwykle znakowanych radioaktywnie, w chromosomalny DNA synchronizowanych hodowli komórkowych (1). Stosowano również techniki *in vitro*: dwukierunkowej elektroforezy DNA, 2-D (*two dimensional gel*) (8) i wyznaczania miejsca inicjacji replikacji metodą RIP (*replication initiation point mapping*) (9,10). Minichromosom jest plazmidem zawierającym chromosomalny region *ori* jako jedyne miejsce inicjacji replikacji DNA. Jego niezależna replikacja w komórkach danej bakterii świadczy o tym, że jest to miejsce, w którym na chromosomie rozpoczyna się powielanie DNA. Badanie replikacji plazmidów stanowiących swoistą bibliotekę fragmentów chromosomalnego DNA pozwala na identyfikację regionu *oriC* chromosomu bez uprzedniego wyznaczenia przypuszczalnych miejsc inicjacji replikacji. Druga metoda *in vivo* opiera się

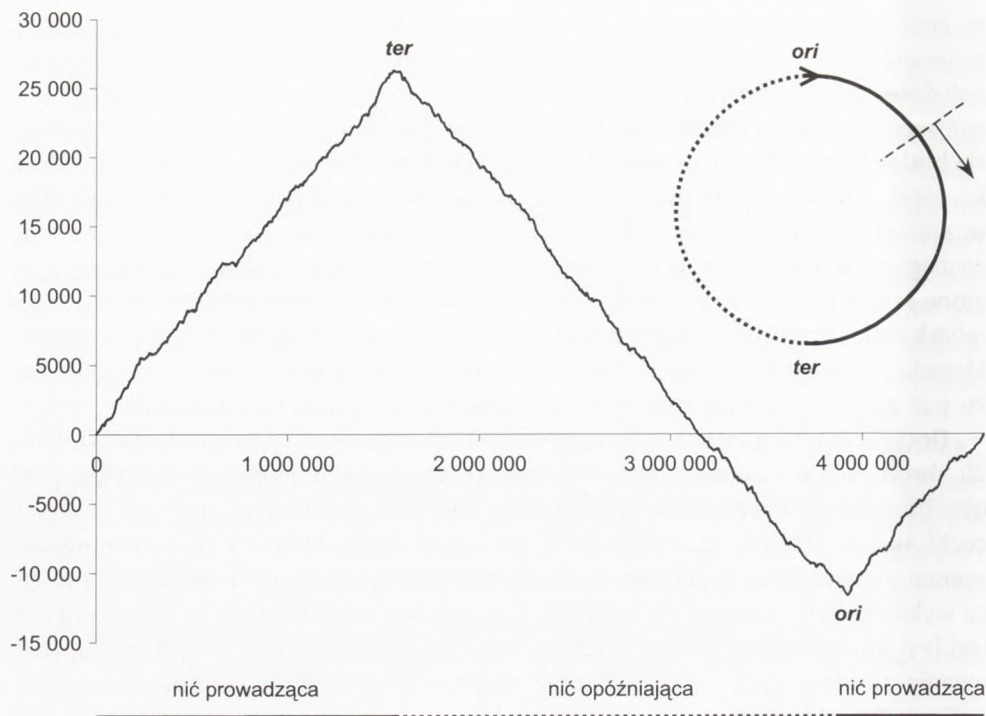
na analizowaniu regionów DNA, w które zostały wbudowane odpowiednio znakowane nukleotydy, i porównania intensywności wyznakowania DNA w zależności od położenia na chromosomie. Region DNA najwcześniej ulegający „zaznaczeniu” jest miejscem, w którym rozpoczyna się replikacja DNA. W metodzie tej głównie stosuje się hodowle zsynchronizowane. Można ją wykorzystać również do badania hodowli komórek niezsynchronizowanych jednak w takich przypadkach niezbędne jest wcześniejsze, przybliżone wyznaczenie regionu *oriC* innymi metodami. Podobne wymagania dotyczą technik *in vitro*, 2-D i RIP: znajomość przybliżonej lokalizacji regionu *oriC* oraz izolowanie DNA z synchronizowanych hodowli (DNA izolowany z komórek z niezsynchronizowanych hodowli jest uboższy w intermediały replikacyjne). Metoda RIP pozwala wyznaczyć miejsce startu replikacji DNA z dokładnością do kilku par zasad, natomiast pozostałe z dokładnością do kilkuset par zasad.

Dotychczas, za pomocą opisanych technik wyznaczono regiony *oriC* dla zaledwie 20 chromosomów bakteryjnych. Techniki te są pracochłonne i charakteryzują się określonymi ograniczeniami. Przykładowo, konstrukcja minichromosomów jest pracochłonna ze względu na konieczność konstruowania bibliotek i ma pewne ograniczenia: nie wszystkie organizmy tolerują dodatkową kopię *oriC* obecną w komórce, co wyklucza zastosowanie tej techniki. Podstawową wadą większości opisanych metod jest konieczność synchronizacji wzrostu komórek (co zwykle wymaga konstruowania zmutowanych szczepów) oraz znajomość przybliżonej lokalizacji regionu *oriC* na chromosomie.

Do niedawna wyznaczenie prawdopodobnego miejsca *oriC* bez badań *in vivo* lub/i *in vitro* nie było możliwe. Od kilku, lat dzięki poznaniu sekwencji nukleotydowych chromosomów ponad 200 bakterii oraz opracowaniu algorytmów komputerowych analizujących te sekwencje, możliwe jest wyznaczanie regionów chromosomów bakteryjnych, w obrębie których dochodzi do inicjacji replikacji DNA.

2. Asymetria w składzie nukleotydowym między różnie replikowanymi niciami DNA

Już na podstawie analizy składu nukleotydowego pierwszych zsekwencjonowanych bakteryjnych chromosomów zauważono, że nici DNA, prowadząca i opóźniająca różnią się składem – są asymetryczne pod względem występowania określonych nukleotydów, np. nic prowadząca zawiera statystycznie więcej guaniny niż nic opóźniająca (11,12). Zaobserwowane zależności zostały potwierdzone w większości zsekwencjonowanych chromosomach bakteryjnych przez innych badaczy (np. (13-18)). Dlaczego nici DNA, prowadząca i opóźniająca są asymetryczne, różnią się składem nukleotydowym? Istnieje wiele przyczyn tej asymetrii, ale najważniejszą, jak się wydaje, jest odmienna presja mutacyjna (wzór podstawień nukleotydowych) związana z replikacją nici syntetyzowanych w sposób ciągły i poprzez fragmenty Okazaki (prace przeglądowe: (19-22)).



Rys. 1. Analiza asymetrii chromosomu *E. coli* metodą spacerów po DNA. Oś X odzwierciedla położenie na chromosomie, a oś Y miarę asymetrii – różnicę między liczbą komplementarnych nukleotydów [G-C]. Ekstrema asymetrii związane z przejściem z nici prowadzącej na opóźniającą lub odwrotnie odpowiadają miejscu inicjacji (*ori*) i terminacji (*ter*) replikacji. Przerwana linia ze strzałką na schemacie chromosomu (wstawka) wskazuje początek spaceru po DNA.

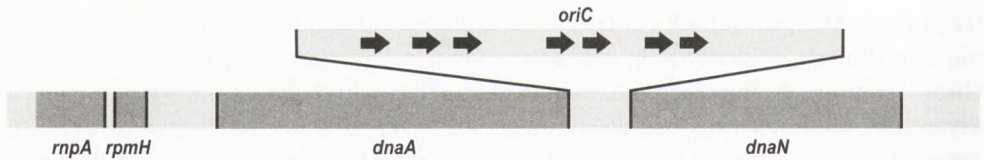
Analizę asymetrii chromosomu można zilustrować graficznie (rys. 1), stosując metodę spacerów po DNA (17). Spacerując po chromosomie obserwujemy kolejne nukleotydy komplementarne i dokonujemy odpowiednich zmian na wykresie, na którym oś X odzwierciedla położenie na chromosomie, a oś Y miarę asymetrii. Jeżeli natrafimy na guaninę, to przesuwamy się o jedną jednostkę do góry, jeżeli na cytozynę, to w dół, a jeżeli na adeninę lub tyminę, to przesuwamy się do kolejnego nukleotydu. W ten sposób uzyskujemy wykres przedstawiający różnicę między liczbą komplementarnych nukleotydów (np. [G-C]) wzdłuż chromosomu. Przy „spacerowaniu” można jeszcze dodatkowo odejmować globalny trend w składzie całego chromosomu, dzięki czemu wykres nie tylko rozpoczyna się, ale i kończy w punkcie dla $Y=0$, co wyraźniej pokazuje lokalne zmiany asymetrii na chromosomie. Zastosowanie spaceru po DNA wygładza losowe fluktuacje asymetrii, pokazując trendy globalne. Obserwowane na wykresie minimum i maksimum, związane z przejściem z nici prowadzącej na opóźniającą i odwrotnie, odpowiada miejscom na chromosomie, w których rozpoczyna i kończy się replikacja chromosomu. Analizując zatem *in silico*

(za pomocą komputera) sekwencje chromosomów, pod kątem częstości występowania określonych nukleotydów, możemy wyznaczyć miejsca inicjacji replikacji tych chromosomów. W analizach można uwzględniać również inne parametry opisujące asymetrię, takie jak częstość występowania poszczególnych nukleotydów (17,18), znormalizowane różnice między komplementarnymi nukleotydami, np. $GCskew = [G-C]/[G+C]$ (12,14,15,19,20), przewagę puryn nad pirymidynami i zasad ketonowych nad aminowymi (13), używalność kodonów i aminokwasów (17,18) rozkład oligomerów (16). Poza prostymi spacerami po DNA (11) i pozbawionymi trendu (17) stosowano również inne metody do opisu i oceny asymetrii wzdłuż chromosomu, jak: analiza odchyień skew w zachodzących oknach (12,15,20), kumulowanie odchyień skew (14,19), analiza wariancji (19), liniowa analiza dyskryminacyjna (18) i transformacja falkowa (23). Analizy można przeprowadzać osobno dla każdej z pozycji w kodonie sekwencji kodujących białko, skupiając się szczególnie na trzecich pozycjach w kodonie, które najbardziej odzwierciedlają asymetrię, ponieważ stosunkowo łatwo akumulują substytucje wprowadzane przez presję mutacyjną (15,17-19).

Metody badające asymetryczność składu nici DNA są obecnie często używane do wyznaczania regionów *oriC*. Opracowano wiele programów komputerowych, które umożliwiają wyznaczanie *oriC*, np. program komputerowy Oriloc (24) lub DNAskew (25). Opierają się one na metodzie podobnej do opisanej. Jednakże, zastosowanie tylko tego typu analiz może prowadzić do błędnej lokalizacji regionu *oriC*, np. u *Helicobacter pylori* wyznaczony *in silico* region *oriC* jest oddalony o około 2500 pz od regionu, w którym białko DnaA swoiście oddziałuje ze zgrupowaniem boksów DnaA (26). Analiza asymetrii DNA niektórych chromosomów wykazuje obecność więcej niż dwóch ekstremów, np. u *Streptomyces avermitilis*, (patrz rys. 3G) występują dwa minima i trudno jest przewidzieć, które z nich odpowiada miejscu inicjacji replikacji. Dlatego analiza asymetrii chromosomu powinna być uzupełniona o poszukiwanie głównych elementów zaangażowanych w inicjację replikacji.

3. Poszukiwanie zasadniczych elementów inicjacji replikacji chromosomów

Bakteryjne regiony *oriC* są zawsze zlokalizowane w odcinkach międzygenowych, często w bliskim sąsiedztwie genów kodujących białka zaangażowane w replikację (*rnpA-rmpH-dnaA-dnaN-recF-gyrB-rnpA*) zwykle przed 5' lub za 3' końcem genu *dnaA* kodującego inicjatorowe białko DnaA (rys. 2) (7). Ustalenie zatem położenia genu *dnaA* na chromosomie może być użyteczne we wstępnej fazie poszukiwania regionu *oriC*. Gen *dnaA* został zidentyfikowany w chromosomach wszystkich badanych mikroorganizmów, poza dwoma wyjątkami – na chromosomach *Blochmannia floridanus* (27) i *Wigglesworthia glossinidia* (28) nie znaleziono genu *dnaA*. Są to gramujemne, endosymbiotyczne organizmy, należące do rodziny enterobakterii, blisko spokrewnione



Rys. 2. Organizacja bakteryjnego regionu *oriC* na przykładzie *Mycobacterium tuberculosis*. Strzałkami zaznaczono położenie i orientację boksów DnaA.

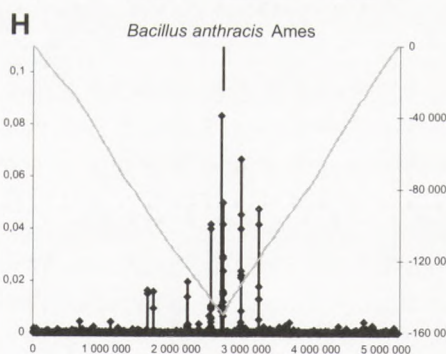
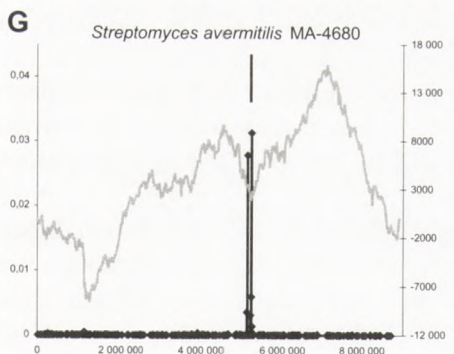
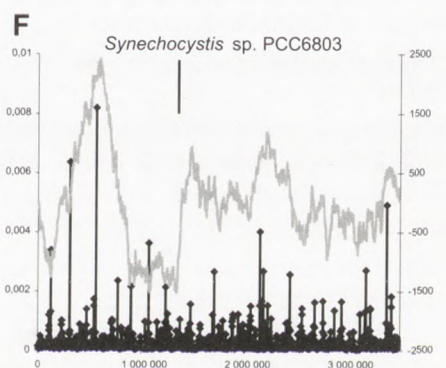
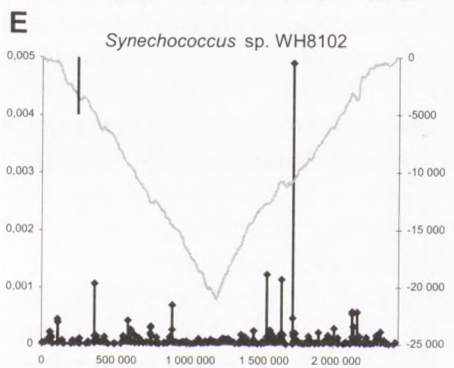
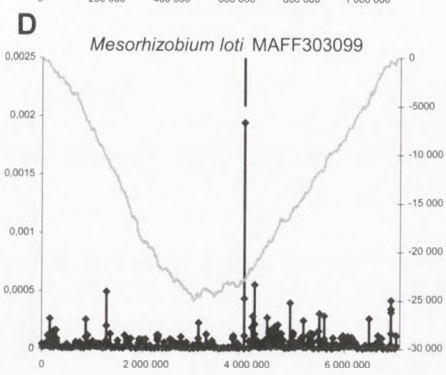
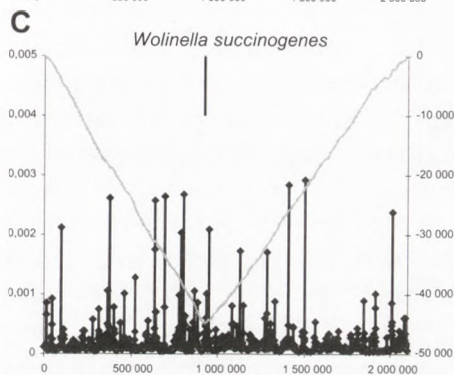
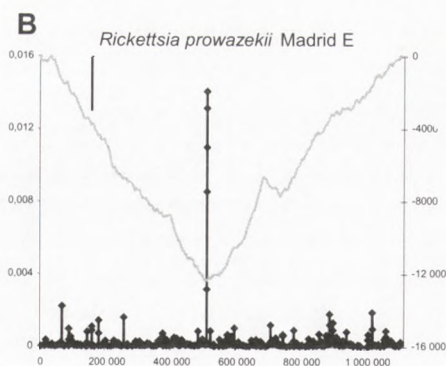
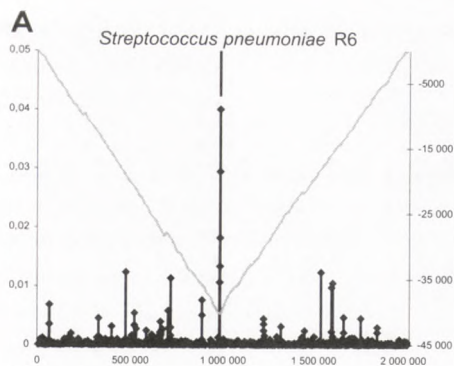
z *E. coli*. Przymuszczalnie inicjacja replikacji chromosomów tych bakterii, podobnie jak ich metabolizm, są w dużym stopniu zależne od gospodarza (mrówka *Camponotus* spp. w przypadku *B. floridanus*; mucha tse-tse w przypadku *W. glossinidia*).

U bakterii w regionie *oriC* występują 9-nukleotydowe, niepalindromowe sekwencje wiążące swoiście inicjatorowe białko DnaA, nazywane dalej boksami DnaA (ang. DnaA box). Długość regionu *oriC* wynosi od 180 pz (*Helicobacter pylori*) do 1000-2000 pz (*Bacillus subtilis*, *Streptomyces*) (29). W tych regionach występuje od kilku (np. *E. coli* – 5 boksów DnaA) do kilkunastu (np. *Streptomyces* – 19 boksów DnaA) sekwencji wiążących inicjatorowe białko DnaA oraz sekwencje bogate w pary AT, w obrębie których dochodzi do rozplecenia podwójnej nici DNA. Liczba boksów DnaA i ich rozmieszczenie w regionie *oriC* mogą być różne, ale przyjmuje się, że do zachowania aktywności replikacyjnej powinno być ich co najmniej 4 na odcinku DNA nie dłuższym niż 200 pz. Wyjątkowe pod tym względem są regiony *oriC* *Xanthomonas campestris* (30), *Coxiella burnetti* (31) i *Caulobacter crescentus* (32), które zawierają odpowiednio tylko 3, 2 lub 1 boks DnaA. Obecność kilku sekwencji DnaA, zwłaszcza w sąsiedztwie genu *dnaA*, może sugerować istnienie regionu *oriC*.

U *E. coli*, na podstawie stałych powinowactwa białka DnaA do DNA, ustalono sekwencję najwyższej zgodności: 5'-TT^A/TNCACA-3' (33). U organizmów charakteryzujących się wysoką zawartością GC (np. *Streptomyces*, % GC ~70), w pozycji 3. boks DnaA najczęściej występuje G lub C. Niektóre pozycje nukleotydowe w sekwencji boks DnaA są bardziej konserwatywne niż inne. W przeprowadzonej analizie oddziaływania białko DnaA-DNA wykazano, że nukleotydy w pozycjach 2, 4, 7 i 9 odgrywają istotną rolę w wiązaniu się z białkiem, natomiast nukleotyd w pozycji 5 prawdopodobnie nie jest zaangażowany w oddziaływanie z tym białkiem. Ostatnio w poznanej strukturze krystalicznej domeny wiążącej DNA (motyw wiążący helisa-skręt-helisa) białka DnaA *E. coli* z oligonukleotydem zawierającym „silny” boks DnaA

Rys. 3. Asymetria DNA, rozmieszczenie sekwencji boksów DnaA i położenie genu *dnaA* na chromosomach bakterii. Oś X – położenie na chromosomie w parach zasad, w A-E i H sekwencje chromosomów rozpoczynają się i kończą w przypuszczalnym miejscu terminacji replikacji, a przypuszczalny region *oriC* znajduje się w połowie wykresu. W F i G położenia na chromosomie są takie, jak w oryginalnie adnotowanych sekwencjach. Szara linia i prawa oś Y pokazuje spacer po DNA opisujący asymetrię ([A-T] w F; dla pozostałych chromosomów – [G-C]). Czarne linie z punktami i lewa oś Y pokazują rozmieszczenie sekwencji boksów DnaA, objaśnienie w tekście. Krótka linia na górze wykresu wskazuje na położenie genu *dnaA*.





(TTATCCACA) wyjaśniono molekularne podstawy oddziaływania poszczególnych aminokwasów z nukleotydami boksów DnaA (34). Okazało się, że białko DnaA tworzy wiązania ze wszystkimi (oprócz nukleotydu w pozycji 5) nukleotydami 9-merowego boksów DnaA.

Poszukując zgrupowania boksów DnaA należy pamiętać, że pozycja 5. boksów DnaA nie jest ewolucyjnie konserwatywna, a u organizmów charakteryzujących się wysoką procentową zawartością par GC, w pozycji 3. boksów DnaA występuje przeważnie G lub C (29). Dodatkowo, w niektórych przypadkach, trzeba mieć świadomość, że w analizowanym organizmie białko DnaA może wykazywać odmienną swoistość substratową niż analogiczne białko u *E. coli* (35). Poszukując zatem zgrupowania boksów DnaA, należy uwzględnić możliwość, że ich sekwencje mogą różnić się nawet do 3 nukleotydów od sekwencji boksów DnaA *E. coli*. Analizę położenia boksów DnaA na chromosomie można zilustrować graficznie: dla każdego boksów DnaA wyliczamy odwrotność jego odległości do przyległych dwóch boksów i przedstawiamy te wartości na diagramie. Im bliżej położone są boksów DnaA (ich zgrupowanie) tym dłuższy jest odcinek, w którym zaznaczone punkty reprezentują położenie boksów DnaA. Długości większe niż 0,005 wskazują na potencjalny klaster trzech boksów, między którymi średnia odległość jest mniejsza niż 100 pz. Przykładowo (rys. 3A) na chromosomie *Streptococcus pneumoniae* R6 występuje tylko jedno zgrupowanie boksów DnaA i w tym miejscu prawdopodobnie rozpoczyna się replikacja. To zgrupowanie boksów DnaA jest położone w bliskim sąsiedztwie genu *dnaA* i ekstremum asymetrii, co dodatkowo potwierdza, że wytypowany region chromosomu jest miejscem inicjacji syntezy DNA. Na chromosomie *E. coli*, podobnie jak na chromosomach wielu innych organizmów (rys. 3H) występuje więcej niż jedno zgrupowanie boksów DnaA i bez uzupełniającej analizy (badanie asymetrii DNA, lokalizacja genu *dnaA*) trudno byłoby przewidzieć, które ze zgrupowań boksów DnaA jest regionem *oriC*.

4. Zastosowanie metod *in silico* do lokalizacji regionu *oriC*

Ostatnio przeprowadzono analizę sekwencji chromosomów 120 bakterii w celu wyznaczenia miejsca inicjacji replikacji (29). Do identyfikacji potencjalnych regionów *oriC* zastosowano trzy niezależne metody komputerowej analizy sekwencji: badanie asymetrii w składzie nukleotydowym między różnie replikowanymi nićmi, wyznaczenie rozkładu boksów wiążących inicjatorowe białko DnaA oraz określenie położenia genu *dnaA*. Spośród 120 analizowanych chromosomów, trzy zastosowane metody zgodnie identyfikowały ten sam region jako *oriC* (rys. 3A) u ponad połowy analizowanych mikroorganizmów (67 organizmów). W 30% chromosomów, położenie tylko dwóch z trzech elementów analizy pokrywa się; na rysunku 3B-D przedstawiono reprezentatywne przykłady. W tej grupie mikroorganizmów ekstremum asymetrii jest w większości przypadków położone w bliskim sąsiedztwie jednego z elementów zaangażowanych w inicjację replikacji (zgrupowania boksów DnaA lub

genu *dnaA*): region ten jest zapewne miejscem inicjacji replikacji DNA. Zaledwie w trzech z analizowanych chromosomów ekstremum jest oddalone od potencjalnego miejsca startu replikacji DNA, w którym występują gen *dnaA* i zgrupowanie boksów DnaA. W 17 przypadkach wszystkie trzy elementy położone są w różnych miejscach na chromosomie (rys. 3E).

W przypadku niektórych chromosomów badanie asymetrii DNA nie wskazuje jednego globalnego ekstremum. Przykładem może być chromosom *Synechocystis* sp. PCC6803 (rys. 3F), w którym można się jednak dopatrzeć lokalnego ekstremum skorelowanego ze zgrupowaniem sekwencji boksów DnaA. Do dziś, pomimo poznania sekwencji nukleotydu chromosomu, nie udało się wyznaczyć eksperymentalnie regionu *oriC* tej bakterii (36). Być może zaproponowane miejsce jest rzeczywistym miejscem inicjacji replikacji.

Należy pamiętać, że niektóre bakterie (np. *Vibrio cholerae*) mają dwa chromosomy, z których jeden ma typowy region *oriC* i gen *dnaA*, natomiast w drugim nie występują te elementy. Przykładowo, u *V. cholerae* replikacja jednego z chromosomów rozpoczyna się w *oriC*, natomiast drugiego w regionie, którego struktura przypomina plazmidowy region *ori* (37). W tym ostatnim chromosomie w pobliżu *ori* zidentyfikowano gen *rep*, którego produkt wykazuje stosunkowo wysoką homologię do plazmidowego białka inicjatorowego.

W większości chromosomów stwierdzono, że boksy DnaA pojawiają się częściej niż wynikałoby to z ich losowego występowania na chromosomie. Ponadto w wielu chromosomach zaobserwowano obecność więcej niż jednego zgrupowania boksów DnaA w okolicy regionu *oriC*. Zgrupowania te mogą uczestniczyć w regulacji procesu replikacji przez miareczkowanie cytoplazmatycznego białka DnaA (29).

Wszystkie lub niektóre z elementów związane z inicjacją replikacji (asymetria nici DNA, zgrupowanie boksów DnaA, gen *dnaA*) zostały utracone u bakterii, które prowadzą wewnątrzkomórkowy pasożytniczy lub symbiotyczny tryb życia, a ich chromosomy uległy znacznej redukcji (29,38).

5. Podsumowanie

Od 1995 r., w którym poznano pierwszą kompletną sekwencję chromosomu bakteryjnego *Haemophilus influenzae*, zsekwencjonowano ponad 200 genomów bakteryjnych. Na podstawie ich analizy wykazano, że chromosom bakteryjny jest asymetryczny, tzn. nić DNA syntetyzowana jako ciągła różni się składem nukleotydowym od nici opóźniającej. Obserwacja ta zaowocowała opracowaniem metody wyznaczania przypuszczalnego regionu inicjacji replikacji. Analiza asymetrii DNA okazała się najbardziej uniwersalną metodą, jednak większą skuteczność przewidywania uzyskuje się uwzględniając wyniki wszystkich trzech opisanych metod. Na podstawie analizy sekwencji nukleotydowych bakteryjnych chromosomów wykazano, że w regionie, w którym rozpoczyna się replikacja chromosomu bakteryjnego, wy-

stępuje zwykle kilka charakterystycznych elementów, wspólnych nawet dla mikroorganizmów odległych filogenetycznie. Podstawowym wyznacznikiem jest obecność zgrupowanych na niewielkim odcinku DNA boksów DnaA swoiście wiążących inicjatorowe białko DnaA. Zwykle *oriC* położone jest w pobliżu genu *dnaA*. Znalezienie na chromosomie bakteryjnym fragmentu DNA zawierającego w bliskim sąsiedztwie kilka do kilkunastu boksów DnaA, gen *dnaA* oraz miejsce zmiany asymetrii między niemi DNA pozwala na wyznaczenie przypuszczalnego miejsca *oriC*. U większości bakterii te trzy elementy sąsiadują ze sobą. Istnieje grupa bakterii, u których tylko dwa z trzech elementów bezpośrednio ze sobą sąsiadują: przeważnie ekstremum asymetrii z jednym z elementów inicjacji replikacji (zgrupowaniem boksów DnaA lub genem *dnaA*). Opracowana metoda umożliwia wyznaczanie miejsc *oriC* w nowo sekwencjonowanych chromosomach bakteryjnych, których liczba przekroczy kilka tysięcy w ciągu najbliższych kilku lat.

Praca finansowana przez Ministerstwo Nauki i Informatyzacji (KBN, projekty badawcze nr 3PO4A07922, 2PO4A06228).

Literatura

1. Kornberg A., Baker T., (1992), *The DNA replication*, 2nd ed. Freeman W. H., and Co., New York, NY.
2. Baker T. A., Bell S. P., (1998), *Cell*, 92, 295-305.
3. Kelman L. M., Kelman Z., (2003), *Mol. Microbiol.*, 48, 605-615.
4. Matsunaga F., Forterre P., Ishino Y., Myllykallio H., (2001), *Proc. Natl. Acad. Sci. USA*, 98, 11152-11157.
5. Robinson N. P., Dionne I., Lundgren M., Marsh V. L., Bernander R., Bell S. D., (2004), *Cell*, 116, 25-38.
6. Messer W., Weigel C., (1996), *Initiation of chromosome replication*, in: *Escherichia coli and Salmonella Cellular and Molecular Biology*, 2nd ed., Ed. Neidhardt F. C., 1579-1601, ASM Press Washington, D.C.
7. Messer W., (2002), *FEMS Microbiol. Rev.*, 26, 355-374.
8. Brewer B. J., Fangman W. L., (1987), *Cell*, 51, 463-471.
9. Gerbi S. A., Bielinsky A. K., (1997), *Methods*, 13, 271-280.
10. Bielinsky A. K., Gerbi S. A., (2001), *J. Cell Sci.*, 114, 643-651.
11. Lobry J. R., (1996a), *Biochimie*, 78, 323-326.
12. Lobry J. R., (1996b), *Mol. Biol. Evol.*, 13, 660-665.
13. Freeman J. M., Plasterer T. N., Smith T. F., Mohr S. C., (1998), *Science*, 279, 1827.
14. Grigoriev A., (1998), *Nucl. Acids Res.*, 26, 2286-2290.
15. McLean M. J., Wolfe K. H., Devine K. M., (1998), *J. Mol. Evol.*, 47, 691-696.
16. Salzberg S. L., Salzberg A. J., Kerlavage A. R., Tomb J-F., (1998), *Gene*, 217, 57-67.
17. Mackiewicz P., Gierlik A., Kowalczyk M., Dudek M. R., Cebrat S., (1999), *Genome Res.*, 9, 409-416.
18. Rocha E. P., Danchin A., Viari A., (1999), *Mol. Microbiol.*, 32, 11-16.
19. Tillier E. R., Collins R. A., (2000), *J. Mol. Evol.*, 50, 249-257.
20. Mrazek J., Karlin S., (1998), *Proc. Natl. Acad. Sci. USA*, 95, 3720-3725.
21. Frank A. C., Lobry J. R., (1999), *Gene*, 238, 65-77.
22. Kowalczyk M., Mackiewicz P., Mackiewicz D., Nowicka A., Dudkiewicz M., Dudek M. R., Cebrat S., (2001), *J. Appl. Genet.*, 42, 553-577.
23. Song J, Ware A, Liu S-L (2003), *BMC Genomics*, 4, 17.
24. Frank A. C., Lobry J. R., (2000), *Bioinformatics*, 16, 560-561.

25. Ma X-R., Xiao S-B., Guo A-Z., Lu J-Q., Chen H-C., (2004), *Acta Biochimica et Biophysica Sinica*, 36, 16-20.
26. Zawilak A., Cebrat S., Mackiewicz P., Król-Hulewicz A., Jakimowicz D., Messer W., Gościński G., Zakrzewska-Czerwinska J., (2001), *Nucl. Acids Res.*, 29, 2251-2259.
27. Gil R., Silva F. J., Zientz E., Delmotte F., Gonzalez-Candelas F., Latorre A., Rausell C., Kamerbeek J., Gadau J., Holldobler B., et al., (2003), *Proc. Natl. Acad. Sci. USA*, 100, 9388-9393.
28. Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M., Aksoy S., (2002), *Nat Genet.*, 32, 335-336.
29. Mackiewicz P., Zakrzewska-Czerwinska J., Zawilak A., Dudek M. R., Cebrat S., (2004), *Nucl. Acids Res.*, 32, 3781-3791.
30. Yen M. R., Lin N. T., Hung C. H., Choy K. T., Weng S. F., Tseng Y. H., (2002), *Appl. Environ. Microbiol.*, 68, 2924-2933.
31. Suhan M., Chen S. Y., Thompson H. A., Hoover T. A., Hill A., Williams J. C., (1994), *J. Bacteriol.*, 176, 5233-5243.
32. Marczyński G. T., Shapiro L., (2002), *Annu. Rev. Microbiol.*, 56, 625-656.
33. Schaefer Ch., Messer W., (1991), *Mol. Gen. Genet.*, 226, 34-40.
34. Fujikawa N., Kurumizaka H., Nureki O., Terada T., Shirouzu M., Katayama T., Yokoyama S., (2003), *Nucleic Acids Res.*, 31, 2077-2086.
35. Zawilak A., Durrant M. C., Jakimowicz P., Backert S., Zakrzewska-Czerwinska J., (2003), *J. Mol. Biol.*, 334, 933-947.
36. Richter D., Hess W. R., Krause M., Messer W., (1998), *Mol. Gen. Genet.*, 257, 534-541.
37. Egan E. S., Waldor M. K., (2003), *Cell*, 114, 521-530.
38. Moran N. A., (2002), *Cell*, 108, 583-586.