



## Wpływ replikacji na organizację i ewolucję genomów bakterii

Maria Kowalczyk, Dorota Mackiewicz, Paweł Mackiewicz,  
Natalia Polak, Stanisław Cebrat

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

### Influence of replication on organization and evolution of bacterial genomes

#### Summary

In bacterial chromosomes, strong bias in nucleotide composition has been observed between differently replicated DNA strands (leading and lagging ones), and also in many species between the regions proximal and distal to the origin of replication (*ori*). This bias is also reflected in composition and distribution of genes along the chromosome. Several phenomena connected with the replication of the chromosome are responsible for such polarization, especially mutational pressure, repair mechanisms and recombinations, and also selection pressure. All these phenomena are not indifferent for gene evolution and their rearrangements which are strictly connected with the organization of bacterial chromosome.

#### Key words:

DNA asymmetry, leading strand, lagging strand, bacterial chromosome, replication, evolution, rearrangements.

#### Adres do korespondencji

Maria Kowalczyk,  
Zakład Genomiki,  
Instytut Genetyki  
i Mikrobiologii,  
Uniwersytet Wrocławski,  
ul. Przybyszewskiego 63/77,  
51-148 Wrocław;  
e-mail:  
kowal@microb.uni.wroc.pl

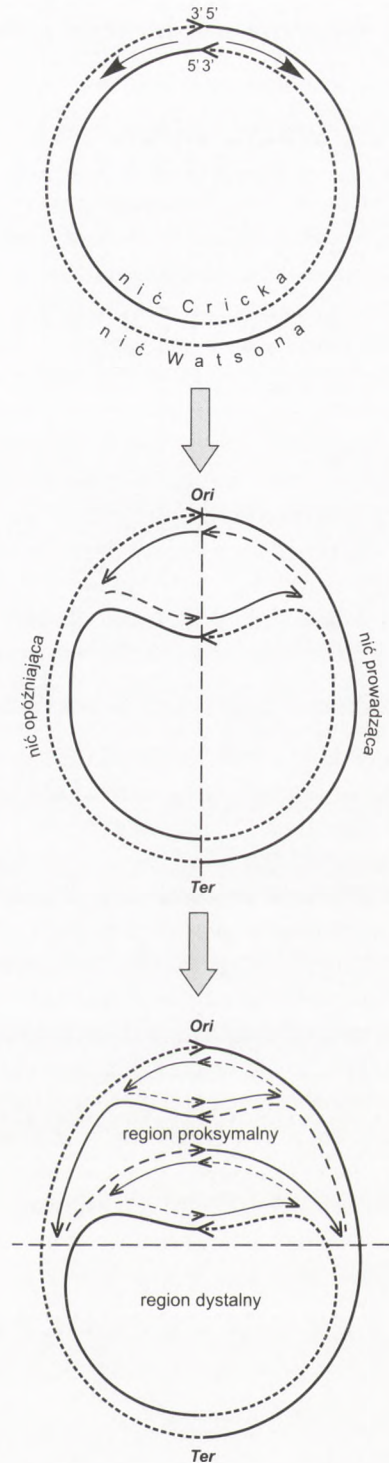
### 1. Replikacja chromosomu bakteryjnego

Dzięki rosnącej liczbie zsekwencjonowanych genomów bakterii pojawia się coraz więcej informacji dotyczących ich organizacji, a także czynników zwiększających lub ograniczających ich zmienność genetyczną, a w konsekwencji wpływających na ich ewolucję. Odkrywanie zasad organizacji genomów pomaga połączyć wyniki badań eksperymentalnych z wiedzą o ogólnej

organizacji genów w genomach; biologia musi być raczej nauką zajmującą się relacjami pomiędzy obiektami, niż nauką opisującą obiekty (1). Genom to coś więcej niż „worek genów”, w jego organizacji istotna jest liczba genów i gęstość kodowania, liczba i rozkład elementów powtórzonych i insercyjnych, skład nukleotydowy oraz kodonowy. Te czynniki są wzajemnie ze sobą powiązane, a ich udział jest relatywnie różny w kształtowaniu struktury genomów różnych gatunków bakterii. Czynnikiem, który leży u podstaw organizacji chromosomu bakteryjnego i wpływa na wszystkie wymienione czynniki jest replikacja DNA (2).

Większość chromosomów bakterii jest kolistą i posiada tylko jedno miejsce inicjacji replikacji (*ori*), z którego widełki replikacyjne startują w dwóch kierunkach, aby spotkać się na przeciwległym krańcu chromosomu w miejscu terminacji replikacji (*ter*) – rysunek 1. DNA może być replikowany tylko w jednym kierunku, od końca 5' do 3', natomiast nici Watsona i Cricka ułożone są antyrównolegle, dlatego widełki replikacyjne muszą być asymetryczne – nić replikowana zgodnie z kierunkiem ruchu widełek powstaje w sposób ciągły (nić prowadząca), natomiast druga nić syntetyzowana jest w sposób nieciągły w miarę odsłaniania się matrycy, a powstałe fragmenty Okazaki są następnie ze sobą łączone (nić opóźniająca). Każda z nici jest syntetyzowana przez oddzielny kompleks polimerazy DNA, przy czym u niektórych gatunków, jak np. *Escherichia coli*, podjednostki  $\alpha$  są kodowane przez ten sam gen (3), natomiast np. u *Bacillus subtilis* podjednostki  $\alpha$  dla nici prowadzącej i opóźniającej są kodowane przez różne geny (4,5). Niezależnie od tego, widełki replikacyjne muszą być asymetryczne, gdyż kompleks nici prowadzącej musi być bardziej procesywny, aby mógł zostać przyłączony do matrycy przez cały czas replikacji, natomiast kompleks nici opóźniającej musi łatwiej od niej oddysocjowywać (6).

W konsekwencji nić Watsona można podzielić na dwie części – tę replikowaną jako nić prowadząca i jako opóźniająca (rys. 1). Komplementarna do niej nić Cricka ma odpowiednio część opóźniająca i prowadząca. Geny, których nić sensowa (odpowiadająca transkryptowi mRNA) jest nicią prowadzącą, uznawane są za położone na nici prowadzącej, analogicznie jest z genami położonymi na nici opóźniającej. Można także rozpatrywać położenie genów w stosunku odległości od *ori* – te leżące w jego pobliżu określane są jako proksymalne, te leżące w pobliżu *ter* – jako dystalne (rys. 1). W pracy opisano organizację chromosomów bakteryjnych z punktu widzenia procesu replikacji oraz jej wpływ na rozmieszczenie genów na chromosomie, ich ewolucję i rearanżacje.



Rys. 1. Schemat replikacji i organizacji chromosomu bakteryjnego. Linie ciągłe oznaczają nici prowadzące, a przerywane – nici opóźniające. *Ori* – miejsce inicjacji replikacji; *Ter* – miejsce terminacji replikacji.

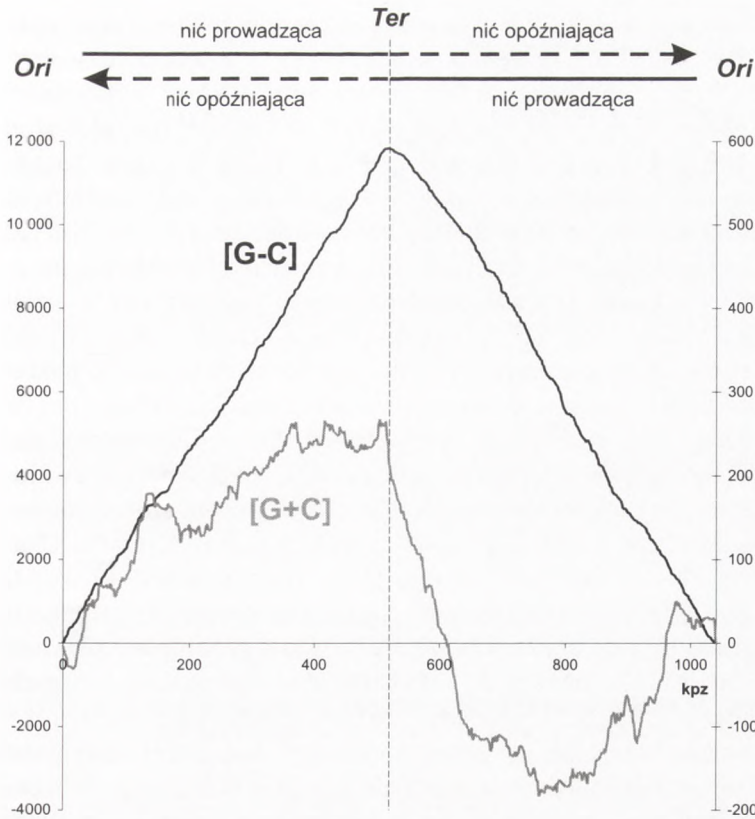
## 2. Asymetria między nicią prowadzącą i opóźniającą

### 2.1. Definicja asymetrii DNA

Gdy w bazach danych pojawiły się sekwencje całych chromosomów bakteryjnych, jedną z pierwszych zaobserwowanych charakterystycznych cech była silna asymetria w składzie nukleotydowym. Czym jest asymetria? W dwuniciowej cząsteczce DNA liczba cytozyn jest dokładnie taka sama jak liczba guanin, natomiast liczba adenin równa jest liczbie tymin – jest to pierwsza reguła parowania sformułowana przez Chargaffa (7). Można wykazać, że przy braku specyficznych czynników mutacyjnych i selekcyjnych ta reguła będzie zachowana również w pojedynczej nici DNA – jest to tzw. druga reguła parowania (8). Jeśli w analizowanej sekwencji występuje statystycznie istotne odchylenie od drugiej reguły parowania, DNA jest określane jako asymetryczny (9).

W pojedynczej nici całego chromosomu nie ma istotnych różnic w ilości komplementarnych nukleotydów na nici Watsona lub Cricka (10,11), natomiast jeśli zliczy się nukleotydy oddzielnie dla nici prowadzącej i opóźniającej, można zaobserwować silną asymetrię (9,12). Nici prowadząca najczęściej charakteryzuje się przewagą guaniny nad cytozyną oraz tyminy nad adeniną, natomiast nici opóźniająca ma więcej cytozyny niż guaniny oraz adeniny niż tyminy. Jest to uniwersalna cecha genomów bakterii (9,13-19). Różnice te są widoczne także, jeśli analizuje się skład sekwencji międzygenowych, jak i kodujących białko. Asymetria odbija się również w składzie kodonowym genów i w składzie aminokwasowym kodowanych przez nie białek (20-22).

Obrazowym sposobem pokazania asymetrii w genomach bakterii są spacery po chromosomie (rys. 2). W tej metodzie zlicza się różnice w ilości komplementarnych nukleotydów, np. [G-C] w kolejnych fragmentach analizowanej nici DNA. Wartości te oblicza się dla kolejnych odcinków nici Watsona lub Cricka, a następnie przedstawia na wykresie w postaci skumulowanych odchyleń od wartości średniej. W ten sposób otrzymuje się wykres pozbawiony trendu charakterystycznego dla całej sekwencji, natomiast uwypuklone zostają trendy lokalne w asymetrii DNA. U większości gatunków bakterii wyraźna zmiana kierunku trendu występuje dokładnie w miejscu *ori* oraz *ter*. Umożliwia to identyfikację tych miejsc wyłącznie na podstawie analizy sekwencji DNA (23-26). Powstawanie asymetrii w genomach bakterii próbowano tłumaczyć na różne sposoby, które można ogólnie podzielić na mutacyjne i selekcyjne (12,16,27).



Rys. 2. Analiza asymetrii [G-C] i składu [G+C] w trzecich pozycjach kodonów w genach *Chlamydia trachomatis* metodą spacerów po DNA. Oś X odzwierciedla położenie na chromosomie, lewa oś Y – skumulowane odchylenia od średniej wartości asymetrii [G-C], a prawa oś Y – skumulowane odchylenia od średniej wartości składu [G+C]. Na nici prowadzącej przeważa guanina nad cytozyną, a w regionie terminacji replikacji maleje udział par GC. Ori – miejsce inicjacji replikacji; Ter – miejsce terminacji replikacji.

## 2.2. Wpływ presji mutacyjnej na asymetrię nici prowadzącej i opóźniającej

Większość badaczy upatruje przyczyn asymetrii w presji mutacyjnej związanej z procesem replikacji. Wskazuje na to zmiana kierunku trendu w miejscach *ori* i *ter*, a także występowanie silnej asymetrii zarówno w trzecich pozycjach kodonów, jak i w sekwencjach międzygenowych (15,17,19,28). Zmiana kierunku trendu w regionach *ori* i *ter* oznacza zmianę sposobu syntezy nici z prowadzącej na opóźniającą i *vice versa*, a zatem zmianę sposobu replikacji nici. Trzecie pozycje kodonów w genach oraz sekwencje międzygenowe podlegają słabszej presji selekcyjnej niż pierwsze i drugie pozycje, toteż powinny lepiej odzwierciedlać presję mutacyjną.

W czasie syntezy nici opóźniającej jej matryca – nić w poprzednim cyklu replikowana jako prowadząca – dłużej pozostaje jednoniciowa, a zatem bardziej podatna na mutacje. Najczęściej obserwowanymi mutacjami w DNA są substytucje. Prawdopodobnie ze zwiększoną częstością dochodzi do deaminacji zmetylowanej cytozyny, co prowadzi do jej zamiany na tyminę (29-31). Frank i Lobry uznali to zjawisko za główną przyczynę powstawania asymetrii i opisali jako teorię deaminacji cytozyny (27). Deaminacja cytozyny wyjaśniałaby obserwację, że tranzycje par GC w AT są dominującymi mutacjami u *E. coli* (32). Zaobserwowano także błędne parowanie tyminy z guaniną w czasie syntezy nici prowadzącej (33). Procesy te sprawiają, że nić prowadząca staje się bardziej bogata w guaninę i tyminę, natomiast nić opóźniająca – w adeninę i cytozynę. Uniwersalność tych odchyień w świecie bakterii wskazuje na to, że za ich powstawaniem stoją procesy replikacji DNA.

Aby dokładniej zbadać, jakie substytucje są odpowiedzialne za powstawanie obserwowanej asymetrii, można porównać geny ortologiczne kilku blisko spokrewnionych gatunków bakterii. Taką analizę wykonali Rocha i Danchin (34), którzy porównywali całkowicie zsekwencjonowane genomy w obrębie rodzaju *Chlamydia*. Okazało się, że najczęściej obserwowanymi substytucjami były przejścia C→T, A→G oraz C→G, co zostało dodatkowo potwierdzone badaniami na dwóch gatunkach z rodzaju *Bacillus*. Autorzy ci zauważyli także, że analizowane genomy nie znajdują się w stanie równowagi z presją mutacyjną. Aby uniknąć problemu wielokrotnych podstawień w tej samej pozycji, badali oni silnie konserwatywne geny o małej dywergencji, które znajdują się pod dużym wpływem selekcji. Tak zatem otrzymane przez nich częstości podstawień nukleotydowych nie stanowią dokładnego odwzorowania presji mutacyjnej. Nieco inne podejście do badania presji mutacyjnej związanej z procesem replikacji zastosowali Kowalczuk i wsp. (35). Porównanie pseudogenów *Borrelia burgdorferi* z oryginalnymi genami pozwoliło na określenie kierunku substytucji, przy założeniu, że pseudogeny kumulują wpadające mutacje, natomiast geny kodujące pozostają pod działaniem selekcji. Dzięki temu udało się znaleźć częstości substytucji nukleotydowych zachodzących w genomie tylko pod wpływem presji mutacyjnej związanej z procesem replikacji. Najczęstszym podstawieniem było C→T, zgodnie z teorią deaminacji cytozyny.

Powstawanie asymetrycznych substytucji próbuje się także tłumaczyć asymetryczną strukturą i funkcją widełek replikacyjnych. Analizy eksperymentalne dają jednak sprzeczne wyniki co do poziomu błędów w czasie syntezy nici prowadzącej i opóźniającej. Trinh i Sinden (36) badali częstość delekcji fragmentu o charakterze palindromu i odkryli, że jest ona większa na nici opóźniającej, prawdopodobnie dlatego, że podczas replikacji odsłonięty jest dłuższy fragment matrycy i łatwiej mogą powstawać w nim pętle. Iwaki i wsp. (37) badali częstość delekcji jednego nukleotydu i substytucji w genach umieszczonych na plazmidzie w różnej orientacji względem kierunku przesuwania się widełek replikacyjnych. Zaobserwowali również większą częstość mutacji na nici opóźniającej, tłumacząc to większą podatnością na mutacje maszynierii replikującej nić opóźniającą. Natomiast Fijałkowska i wsp. (6) zaobser-

wowali mniejszą częstość mutacji na nici opóźniającej, co tłumaczyli większą procesywnością kompleksu polimerazy syntetyzującego nić prowadzącą niż opóźniającą. Prowadzi to do zwiększonej możliwości napraw źle sparowanych zasad na nici opóźniającej, ponieważ kompleks polimerazy tej nici musi często oddysocjowywać. Jednak badania te obejmowały analizę mutacji w operonie *lacZ*, który u *E. coli* jest zlokalizowany na nici opóźniającej, a zatem jego translokacja na nić prowadzącą musi spowodować zwiększenie częstości mutacji (38).

Ponieważ w procesie transkrypcji nić kodująca pozostaje pojedyncza i odsłonięta, a zatem bardziej narażona na mutacje, asymetrię w składzie nukleotydowym genów i chromosomów próbowano także wyjaśnić presją mutacyjną związaną z procesem transkrypcji (13,39,40). Wtedy jednak asymetria byłaby obserwowana jedynie w genomach mających bardzo nierównomierny rozkład genów na nici prowadzącej i opóźniającej, natomiast nie byłaby widoczna w sekwencjach międzygenowych.

### 2.3. Asymetryczne ułożenie genów na różnie replikowanych niciach DNA

Sekwencje kodujące białko mają swoją własną asymetrię pomiędzy nicią sensową a antysensową, wynikającą z funkcji kodowania białka i optymalizacji procesów translacji (41-48). Skład i asymetria są inne dla każdej pozycji w kodonie. Przewaga genów na jednej z nici powinna zatem wprowadzać również asymetrię w skali całego chromosomu.

U niemal wszystkich gatunków bakterii zaobserwowano przewagę liczby genów kodujących białka na nici prowadzącej w stosunku do nici opóźniającej. U niektórych bakterii, np. *Mycoplasma*, różnica ta jest tak znaczna, że to właśnie głównie ona odpowiada za asymetrię obserwowaną pomiędzy odmiennie replikowanymi nićmi (20). Co ciekawe, u gatunków, u których podjednostki  $\alpha$  polimerazy DNA kodowane są przez różne geny (bakterie gramdodatnie, *Mycoplasma*), występuje znacznie większa dysproporcja pomiędzy liczbą genów położonych na nici prowadzącej (średnio 78%), niż u gatunków w których obie podjednostki koduje ten sam gen (średnio 58% genów na nici prowadzącej) (49).

Różnice te tłumaczy się za pomocą różnych mechanizmów działania selekcji. Przede wszystkim rozpatruje się konsekwencje kolizji między polimerazami DNA i RNA. Geny położone na nici prowadzącej ulegają transkrypcji w tym samym kierunku, w jakim porusza się kompleks replikacyjny. Jednak tempo transkrypcji jest dużo wolniejsze od replikacji, jako że musi być dopasowane do szybkości translacji. Gdy kompleks polimerazy DNA napotka kompleks transkrypcyjny na nici prowadzącej, musi zwolnić, najczęściej jednak nie dochodzi do przerwania replikacji ani transkrypcji. Dużo trudniejsza jest sytuacja na nici opóźniającej, ponieważ kompleks polimerazy zderza się czołowo z kompleksem transkrypcyjnym, co powoduje zahamowanie replikacji i przerwanie transkrypcji połączone z oddysocjowaniem częściowo ukończonego transkryptu (50-51). Na podstawie pierwszych obserwacji wska-

zywano, że geny o wysokiej ekspresji położone są preferencyjnie na nici prowadzącej (52), co, jak się wydaje, pozostawało w zgodzie z hipotezą kolizji pomiędzy polimerazami (53). Jednakże w dalszych badaniach wykazano, że różnice w liczbie genów na nici prowadzącej i opóźniającej nie są wcale najwyższe u gatunków szybko rosnących, u których kolizje polimeraz powodowałyby najwięcej szkód (49). Także liczba genów o wysokiej ekspresji jest zbyt mała, by mogły one decydować o przewadze genów kodowanych na nici prowadzącej (2). Zaobserwowano, że w większości genomów bakterii to geny silnie konserwatywne, o funkcjach niezbędnych do życia komórki (*essential genes*) są położone preferencyjnie na nici prowadzącej niezależnie od tego, czy ulegają wysokiej ekspresji (54,55). Wiadomo, że geny kodujące białka rybosomalne są preferencyjnie zlokalizowane na nici prowadzącej (15,16,40). W razie kolizji polimeraz transkrypcja na nici prowadzącej może być ukończona, natomiast na nici opóźniającej zwykle zostaje przerwana i uwolniony niepełny transkrypt może ulec translacji, dając krótszy peptyd, co może prowadzić do zaburzenia istotnej funkcji komórki (2). Jeśli taki peptyd stanowi część białkowego kompleksu, może doprowadzić do jego inaktywacji dając dominujący negatywny fenotyp (56). Szkodliwość kolizji zależy zatem od funkcji genu ulegającego transkrypcji, a nie od częstości kolizji (54).

#### 2.4. Związek ewolucji genów z asymetrią DNA

Występowanie zróżnicowanej presji mutacyjnej i asymetrycznego rozmieszczenia genów na chromosomie nie pozostaje bez wpływu na tempo ewolucji genów w zależności od ich położenia. Porównanie sekwencji genów ortologicznych, tzn. odziedziczonych po wspólnym przodku po etapie specjacji, wykazało dla wielu genomów bakterii, że średnia dywergencja genów położonych na nici prowadzącej jest istotnie statystycznie mniejsza od średniej dywergencji genów położonych na nici opóźniającej (38). Za niższe tempo ewolucji sekwencji z nici prowadzącej może odpowiadać wspomniana wcześniej obecność na tej nici genów niezbędnych do funkcjonowania komórki (*essential genes*), które powinny być konserwatywne.

Niewykluczone, że za różnicę w dywergencji genów z obu nici odpowiedzialne jest odmiennie dopasowanie składu genów do presji mutacyjnej działającej na nią, na której leżą. Zakłada się, że to dopasowanie jest tym większe, im dłużej gen leży na danej nici. Dzięki temu możliwe jest zminimalizowanie liczby mutacji zachodzących w genach i zachowanie kodowanej przez nie funkcji. Ponieważ asymetria w składzie dotyczy nie tylko chromosomów, ale również występuje pomiędzy nicią sensową i antysensową genów, zatem efekt ten może wpływać, jak się wydaje, na różną podatność na mutacje. Stwierdzono, że w niciach sensowych sekwencji kodujących preferencyjnie występuje guanina (np. 15,41,45, 46,48). Dlatego geny położone na nici prowadzącej są w stosunku do genów z nici opóźniającej mniej narażone na mutacje C→T, które zachodzą z większą częstością na nici prowadzącej niż



na nici opóźniającej (27,35). Natomiast geny z nici opóźniającej posiadające nici antysensowe bogate w cytozynę, są bardziej podatne na mutacje. Wobec tego gorsze dopasowanie składu do działającej presji mutacyjnej sekwencji z nici opóźniającej mogłoby odpowiadać za wyższe wartości dywergencji dla genów z tej nici (38).

Dopasowanie genów do działającej na nie presji mutacyjnej sprawia, że gen przeniesiony na odmiennie replikowaną nić zostaje poddany zupełnie innej presji, co powoduje wzrost jego dywergencji. Ortologi, które zmieniły nić w swojej historii ewolucyjnej charakteryzują się najwyższą dywergencją w porównaniu do ortologów z nici opóźniającej i prowadzącej (34,38,57,58). Efekt nowej presji mutacyjnej jest bardzo silny i geny po przeniesieniu na inną nić szybko upodabniają swój skład do składu nowej nici (34,59,57). Należy jednak pamiętać, że wzrost akumulacji mutacji, na jaki narażony jest gen po tego typu inwersji, może przyczyniać się do utraty kodowanej funkcji i jego eliminacji. Efekt ten jest szczególnie niekorzystny dla genów konserwatywnych leżących na nici prowadzącej przeniesionych na nić opóźniającą (60). Możliwe zatem, że inwersje genów związane są z ich duplikacjami, co generuje powstawanie sekwencji paralogicznych, a wraz z nimi prowadzi do nadmiarowości informacji genetycznej (61). Sekwencje zduplikowane mogą podlegać zupełnie innej ewolucji niż geny występujące tylko w jednej kopii. Zwolnione spod selekcji mogą kumulować mutacje (62-64) i być źródłem pseudogenów, które często spotyka się w genomach bakterii (65-70). W związku z tym, że przy odnajdowaniu ortologów często trudne jest całkowite wyeliminowanie tego typu sekwencji, fakt ten może częściowo tłumaczyć dziwnie wysokie wartości dywergencji dla genów, które zmieniły nić (58). Zduplikowane geny mogą również prowadzić do powstawania nowych genów (71,72). Wobec tego mogą stanowić źródło przyspieszania ewolucji bakterii zapewniając im szybkie przystosowanie do zmieniającego się środowiska.

Gen leżący długo na danej nici poddany jest przez dłuższy czas kierunkowej presji mutacyjnej, co dla niektórych genów może być niekorzystne. Aby zachować odpowiedni skład (często pośredni między presjami działającymi na obie nici) korzystne jest dla nich uleganie z pewną częstotliwością inwersjom z jednej nici na drugą i unikanie zbyt skrajnego dla nich składu. Oczywiście wiąże się to ze wzrostem jego dywergencji, jednak powoduje jednocześnie wzrost prawdopodobieństwa przeżycia genu. Symulacje komputerowe potwierdziły możliwość takiej strategii w ewolucji genów (73,74). Wyniki symulacji są zgodne z zaobserwowaną większą dywergencją genów leżących na różnie replikowanych niciach oraz ze stwierdzoną w analizach porównawczych genomów słabą konserwatywnością położenia genów na chromosomie nawet u blisko spokrewnionych organizmów (75-80).

Zjawisko zróżnicowanego tempa ewolucji prowadzi do problemów w oszacowaniu wiarygodnych powiązań filogenetycznych, jeśli są one oparte na analizie genów pochodzących z różnie replikowanych nici (38,58).

### 3. Zróżnicowanie proksymalno-dystalne chromosomów bakterii

Regiony genomu położone na przeciwległych biegunach chromosomu w stosunku do *ori* i *ter* również różnią się składem nukleotydowym. W wielu genomach obszary w okolicy *ori* są bogate w G+C, a w regionie *ter* – bogate w A+T, szczególnie w trzecich pozycjach kodonów (17,81,82) – rysunek 2. Może to wynikać z mutacji, powodowanych różną dostępnością prekursorów nukleotydów potrzebnych do replikacji – na początku tego procesu jest ich pod dostatkiem, natomiast pod koniec może zacząć ich brakować, co może doprowadzić do specyficznych odchyień w składzie nukleotydowym sekwencji położonych w pobliżu *ter* (83,84). Nadmiar A+T może być także związany z obecnością specyficznych miejsc wiążących białka biorące udział w terminacji replikacji oraz z tworzeniem specyficznych struktur trzeciorzędowych ułatwiających rozdział chromosomów po replikacji (82). Również inne procesy naprawcze, które nie są oparte na rekombinacji homologicznej zachodzące w okolicy *ter* mogą preferencyjnie wprowadzać A i T (81,82). Nieuwzględnianie zróżnicowania składu G+C na chromosomie może wpływać na błędne założenie, że wiele sekwencji leżących w rejonie *ter* zostało nabytych w wyniku bocznego transferu (85).

Proksymalno-dystalne zróżnicowanie zaobserwowano także w rozkładzie genów, co wiąże się ze sposobem replikacji chromosomu. Szybko rosnące bakterie, jak np. *E. coli* mogą się dzielić co około 20 minut, natomiast do zreplikowania całego chromosomu potrzeba co najmniej dwa razy tyle czasu. Dlatego kolejna runda replikacyjna rozpoczyna się przed ukończeniem poprzedniej, co prowadzi do obecności 4 lub nawet 8 razy więcej kopii genów położonych w pobliżu inicjacji replikacji dzięki równoczesnej obecności dodatkowych widetek replikacyjnych (86) – rysunek 1. W związku z tym w tym regionie powinny być położone geny, których produkty wymagane są w szczególnie dużej ilości. Rzeczywiście, u szybko rosnących bakterii, np. *E. coli* lub *B. subtilis*, obserwuje się w obszarze chromosomu położonym proksymalnie znaczną przewagę genów o wysokiej ekspresji, np. związanych z translacją – kodujących rRNA lub białka rybosomalne (2,87,88). Położenie genów o wysokiej ekspresji w pobliżu *ori* jest również istotne dla organizacji i podziału całej komórki. Po rozpoczęciu replikacji regiony *ori* szybko oddalają się od siebie do przeciwległych biegunów komórki, natomiast region *ter* lokuje się w jej środku (89). Siła motoryczna polimerazy RNA może odciągać od siebie rejony inicjacji replikacji podczas transkrypcji genów położonych w pobliżu *ori* (90). Ułatwia to później segregację chromosomów i podział, zwłaszcza, że geny o wysokiej ekspresji grupują się wokół *ori* (91).

Natomiast w pobliżu regionu terminacji replikacji preferencyjnie lokują się geny adaptatywne, pochodzące z bocznego transferu lub transpozony (92-94). Może to być związane ze zwiększoną częstością rekombinacji spowodowaną obecnością profagów w tym obszarze, na przykład u *E. coli* (95). Uważa się wręcz, że w tym rejonie występują gorące miejsca (*hot spots*) umożliwiające włączanie obcego DNA do

chromosomu (96). Geny pochodzące z bocznego transferu nie wymagają wysokiej ekspresji i dużej liczby kopii, dlatego mogą się lokować preferencyjnie w pobliżu *ter* (2).

Istnieje również związek między tempem ewolucji genów a ich położeniem względem *ori* i *ter*. Analizując częstość substytucji synonimicznych u *E. coli* i *Salmonella typhimurium* zauważono, że jest ona dwukrotnie wyższa w pobliżu *ter* niż w pobliżu *ori* (83). Jednak po uwzględnieniu różnic w używalności kodonów i w poziomie ekspresji genów okazało się, że zmiany w częstości substytucji związane z odległością od *ori* to jedynie około 5% całkowitej zmienności w poziomie substytucji (97), a geny o najwyższym tempie substytucji grupują się w stosunkowo niewielkim, bogatym w A+T regionie w pobliżu *ter* (82). Efekt ten, jak wcześniej przypuszczano (83,98,99), prawdopodobnie nie wynika z większej możliwości napraw ewentualnych defektów poprzez konwersje i rekombinacje w genach położonych blisko miejsca *ori*. Jest to potwierdzone przez istotny związek odległości względem *ori* a liczbą transwersji, co trudno jest wytłumaczyć efektem procesów naprawczych nie rozróżniających typów podstawień (97). Niewykluczone, że różne tempo mutacji może być związane ze zmianą aktywności enzymatycznej w różnych regionach chromosomu, np. polimeraza DNA wykazuje tendencję do oddysocjowania z syntetyzowanej nici w czasie replikacji, a jej ponowne przyłączenie może generować błędy (100,101). Podwyższone tempo mutacji w pobliżu *ter* może wynikać również z powstawania regionów jednoniciowych charakterystycznych dla zatrzymywanych widełek replikacyjnych, co zwiększa podatność tych regionów na mutacje i rekombinacje (102). Znaczenie mogą również mieć różne procesy zachodzące podczas rozdziału dimerów chromosomów po replikacji oraz procesy naprawcze bardziej podatne na błędy zachodzące w okolicy *ter* (82,83).

#### 4. Związek rearanżacji ze strukturą chromosomu

Wszystkie opisane zjawiska mają związek z rearanżacjami chromosomu kształtującymi jego strukturę i organizację. Tempo rearanżacji uzależnione jest od liczby powtórzeń związanych z rekombinacjami chromosomowymi (103). Genomy nie posiadające tych powtórzeń są bardziej stabilne i charakteryzują się większą asymetrią między nicią prowadzącą i opóźniającą (104,105). Największą asymetrię zaobserwowano u bakterii żyjących obligatoryjnie wewnątrzkomórkowo (2). Genomy te należą także do najbardziej stabilnych, zawierają niewiele elementów powtórzonych ułatwiających rearanżacje, tak zatem ich sekwencje pozostając przez dłuższy czas w takim samym położeniu mogły w czasie ewolucji zgromadzić wiele asymetrycznych substytucji. Rearanżacje nie są obojętne dla genów – mogą zmieniać ich położenie względem miejsca początku replikacji i przemieszczać między różnie replikowanymi niciami.

Porównanie względnego położenia najbliższych ortologów na chromosomie w blisko spokrewnionych genomach pokazało, że wiele genów zmieniło swoje

położenie na chromosomie, zachowując niezmienną odległość od miejsca początku i końca replikacji. Występowanie tego typu symetrycznych translokacji genów wokół osi łączącej *ori* i *ter* było wielokrotnie stwierdzone w genomach bakteryjnych (106-112). Za główną przyczynę tego zjawiska uważa się przebieg procesu replikacji, w którym rozchodzące się symetrycznie widełki replikacyjne zwiększają prawdopodobieństwo zajścia rekombinacji wzajemnej lub transpozycji pomiędzy sekwencjami nowo powstających cząsteczek DNA (108). Możliwość fizycznego kontaktu widełek replikacyjnych w czasie replikacji (113,114), potwierdza, jak się wydaje, tę hipotezę.

Nie jest to jednak jedyne wytłumaczenie przewagi tego typu rearanżacji wokół osi łączącej *ori* i *ter*. Podstawową rolę może odgrywać w tym przypadku selekcja (115). Wspomniano już, że położenie genu w stosunku do miejsca początku i końca replikacji warunkuje jego względną liczbę kopii w komórce. W regionie proksymalnym do *ori* dominują geny o wysokiej ekspresji. Wobec tego dla genów, a w szczególności genów o wysokiej bądź niskiej ekspresji istotne jest optymalne położenie w stosunku do miejsca *ori* i *ter* (87,116-120). Obserwowane rearanżacje właśnie zachowują tego typu położenie. Inną przyczyną częstego występowania tego typu translokacji może być presja selekcyjna utrzymująca replichory (tj. replikowane jednocześnie połówki chromosomu) równej długości (118), gdyż to gwarantuje najkrótszy czas replikacji. W związku z tym, że tylko symetryczne rearanżacje nie zmieniają wielkości replichor, są one preferowane.

Interesujące jest jednak to, że przewagę symetrycznych rearanżacji zaobserwowano dla genów, które nie zmieniły nici ze względu na sposób jej replikacji (115). To sugeruje jej związek z asymetrią pomiędzy nicią prowadzącą a opóźniającą. Podanie obu tych nici różnej presji mutacyjnej związanej z replikacją sprawia, że translokacja genu połączona ze zmianą nici (inwersja) naraża go na zmianę presji mutacyjnej, co zwiększa częstość zachodzących w nim mutacji. Jego przeniesienie może powodować nie tylko substytucje aminokwasowe, ale może również zaburzać specyficzną używalność kodonów, szczególnie istotną dla genów o wysokiej ekspresji (121-123). Efektem tego może być nawet eliminacja genu. Dlatego też akceptowane powinny być, jak się wydaje, takie zmiany położenia genów na chromosomie, które zapewnią im pozostanie na nici replikowanej w ten sam sposób. Natomiast rearanżacje symetryczne, obejmujące *ori* zachowują nie tylko odległość od *ori*, ale również sposób replikacji nici.

Wpływ asymetrii na ograniczenia w przemieszczaniu genów pomiędzy nimi potwierdza otrzymana negatywna korelacja pomiędzy frakcją paralogów położonych na różnie replikowanych niciach w genomie a wielkością asymetrii (124). Chociaż równie dobrze można by przyjąć, że wysoka wartość asymetrii w genomie jest wynikiem zmniejszonej częstości rearanżacji pomiędzy nimi, to jednak to wytłumaczenie, jak się wydaje, jest mało prawdopodobne ze względu na powszechność przemieszczania się genów w genomach bakteryjnych (76,106,108,125,126). Zanik konserwatywnego położenia genów wraz z powiększaniem się dystansu filogenetyczne-

go porównywanych genomów sugeruje, że nie zawsze dochodzi do eliminacji genu po zmianie przez niego nici. Obecność ortologów położonych na różnych niciach może jednak wynikać z inwersji połączonej z duplikacją. Świadczy o tym losowe rozmieszczenie punktów na wykresach typu *dot plot*, kiedy porównywane jest względne położenie na chromosomie wszystkich ortologów leżących na różnych niciach (115,124). Większość z tych ortologów to homologi charakteryzujące się dużą dywergencją. Wobec tego można przypuszczać, że przeniesieniu uległy dodatkowe kopie genów, które jako paralogi mogą być zwolnione spod silnej presji selekcyjnej. W takiej sytuacji wzrost liczby zachodzących mutacji spowodowany zmianą presji mutacyjnej wynikającej ze zmiany nici, nie byłby dla tych genów bardzo szkodliwy.

Na nici prowadzącej położone są preferencyjnie geny istotne dla funkcjonowania komórki ze względu na mniejsze prawdopodobieństwo kolizji między kompleksami enzymów replikacyjnych i transkrypcyjnych, dlatego efekt zmiany nici powinien być bardziej szkodliwy w przypadku translokacji genów z nici prowadzącej na opóźniającą. Rzeczywiście, analizując ortologi, które zmieniły nic w blisko spokrewnionych genomach zaobserwowano, że przeniesienie genów z nici prowadzącej na opóźniającą było akceptowane względnie rzadziej niż w przeciwnym kierunku (60). Również analizy ortologów w wielu genomach wykazały, że wspólne dla tych genomów ortologi położone na nici prowadzącej są obecne nawet przy dodaniu odległego genomu należącego do archebakterii, podczas gdy ortologi położone wyłącznie na nici opóźniającej zanikają bardzo szybko (124).

## 5. Zakończenie

Do niedawna zakładano, że skład nukleotydowy genomów bakterii jest dość jednorodny, natomiast obserwowane odchylenia są zazwyczaj wynikiem horyzontalnego transferu. Obecnie jednak wiadomo, że zróżnicowanie w obrębie jednego genomu może być znaczne. Dotyczy ono różnie replikowanych nici: prowadzącej i opóźniającej, a także u wielu gatunków regionów położonych w pobliżu regionów początku i końca replikacji. Taka organizacja chromosomów ma swoje odbicie również w specyficznej dystrybucji genów na chromosomie – na nici prowadzącej dominują geny istotne dla funkcji komórki, a w okolicy *ori* – geny o wysokiej ekspresji. Za taką polaryzację chromosomów odpowiada wiele zjawisk związanych z replikacją, szczególnie presja mutacyjna, mechanizmy naprawcze i rekombinacje, a także presja selekcyjna. Wszystkie te zjawiska nie są obojętne dla ewolucji genów i ich rearanżacji, które ściśle wiążą się z organizacją chromosomu bakteryjnego.

## Literatura

1. Danchin A., Guerdoux-Jamet P., Moszer I., Nitschké P., (2000), *Phil. Trans. R. Soc. Lond. B*, 355, 179-190.
2. Rocha E. P. C., (2004), *Microbiology*, 150, 1609-1627.
3. Yuzhakov A., Turner J., O'Donnel M., (1996), *Cell*, 86, 877-886.
4. Bruck I., O'Donnel M., (2000), *J. Biol. Chem.*, 275, 28791-28983.
5. Dervyn E., Suski C., Daniel R., Bruand C., Chapuis J., Errington J., Janniére L., Ehrlich S. D., (2001), *Science*, 294, 1716-1719.
6. Fijałkowska I. J., Jonczyk P., Maliszewska-Tkaczyk M., Białoskorska M., Schaaper R. M., (1998), *Proc. Natl. Acad. Sci. USA*, 95, 10020-10025.
7. Chargaff E., (1950), *Experientia*, 6, 201-240.
8. Lobry J. R., (1995), *J. Mol. Evol.*, 40, 326-330; Erratum, 41, 680.
9. Lobry J. R., (1996), *Mol. Biol. Evol.*, 13, 660-665.
10. Lin H. J., Chargaff E., (1967), *Biochim. Biophys. Acta*, 145, 398-409.
11. Rudner R., Karkas J. D., Chargaff E., (1968), *Proc. Natl. Acad. Sci. USA*, 60, 630-635.
12. Kowalczuk M., Mackiewicz P., Mackiewicz D., Nowicka A., Dudkiewicz M., Dudek M. R., Cebrat S., (2001), *J. Appl. Genet.*, 42, 553-577.
13. Freeman J. M., Plasterer T. N., Smith T. F., Mohr S.C., (1998), *Science*, 279, 1827.
14. Grigoriev A., (1998), *Nucleic Acids Res.*, 26, 2286-2290.
15. McLean M. J., Wolfe K. H., Devine K. M., (1998), *J. Mol. Evol.*, 47, 691-696.
16. Mrazek J., Karlin S., (1998), *Proc. Natl. Acad. Sci. USA*, 95, 3720-3725.
17. Mackiewicz P., Gierlik A., Kowalczuk M., Dudek M. R., Cebrat S., (1999), *Genome Res.*, 9, 409-416.
18. Rocha E. P. C., Danchin A., Viari A., (1999), *Mol. Microbiol.*, 32, 11-16.
19. Tillier E. R. M., Collins R. A., (2000), *J. Mol. Evol.*, 50, 249-257.
20. Perrière G., Lobry J. R., Thioulouse J., (1996), *CABIOS*, 12, 519-524.
21. Lafay B., Lloyd A. T., McLean M. J., Devine K. M., Sharp P. M., Wolfe K. H., (1999), *Nucleic Acids Res.*, 27, 1642-1649.
22. Mackiewicz P., Gierlik A., Kowalczuk M., Dudek M. R., Cebrat S., (1999), *J. Appl. Genet.*, 40, 1-14.
23. Lobry J. R., (1996), *Science*, 272, 745-746.
24. Picardeau M., Lobry J.R., Hinnebusch B. J., (1999), *Mol. Microbiol.*, 32, 437-445.
25. Zawilak A., Cebrat S., Mackiewicz P., Król-Hulewicz A., Jakimowicz D., Messer W., Goscinak G., Zakrzewska-Czerwińska J., (2001), *Nucleic Acids Res.*, 29, 2251-2259.
26. Mackiewicz P., Zakrzewska-Czerwińska, J., Zawilak, A., Dudek, M. R., Cebrat, S., (2004), *Nucleic Acids Res.*, 32, 3781-3791.
27. Frank A. C., Lobry J. R., (1999), *Gene*, 238, 65-77.
28. Lobry J. R., Sueoka N., (2002), *Genome Biol.*, 3, RESEARCH0058.
29. Lindahl T., Nyberg B., (1974), *Biochemistry*, 13, 3405-3410.
30. Frederico L. A., Kunkel T. A., Shaw B. R., (1990), *Biochemistry*, 29, 2532-2537.
31. Lutsenko E., Bhagwat A. S., (1999), *Mutat. Res.*, 437, 11-20.
32. Echols H., Goodman M. F., (1991), *Annu. Rev. Biochem.*, 60, 477-511.
33. Gawel D., Maliszewska-Tkaczyk M., Jonczyk P., Schaaper R. M., Fijałkowska I. J., (2002), *Mutat. Res.*, 501, 129-136.
34. Rocha E. P. C., Danchin A., (2001), *Mol. Biol. Evol.*, 18, 1789-1799.
35. Kowalczuk M., Mackiewicz P., Mackiewicz D., Nowicka A., Dudkiewicz M., Dudek M. R., Cebrat S., (2001), *BMC Evol. Biol.*, 1, 13.
36. Trinh T. Q., Sinden R. R., (1991), *Nature*, 352, 544-547.
37. Iwaki T., Kawamura A., Ishino Y., Kohno K., Kano Y., Goshima N., Yara M., Furusawa M., Doi H., Imamoto F., (1996), *Mol. Gen. Genet.*, 251, 657-664.
38. Szczepanik D., Mackiewicz P., Kowalczuk M., Gierlik A., Nowicka A., Dudek M. R., Cebrat S., (2001), *J. Mol. Evol.*, 52, 426-433.
39. Francino M. P., Chao L., Riley M. A., Ochman H., (1996), *Science*, 272, 107-109.

40. Francino M. P., Ochman H., (1997), *Trends Genet.*, 13, 240-245.
41. Shepherd J. C., (1981), *Proc. Natl. Acad. Sci. USA*, 78, 1596-1600.
42. Smithies O., Engels W. R., Devereux J. R., Slightom J. L., Shen S. H., (1981), *Cell*, 26, 345-353.
43. Wong J. T., Cedergren R., (1986), *Eur. J. Biochem.*, 159, 175-180.
44. Lagunez-Otero J., Trifonov E. N., (1992), *J. Biomol. Struct. Dyn.*, 10, 455-464.
45. Karlin S., Burge C., (1995), *Trends Genet.*, 11, 283-290.
46. Gutierrez G., Marquez L., Martin A., (1996), *Nucleic Acids Res.*, 24, 2525-2528.
47. Cebrat S., Dudek M. R., Mackiewicz P., Kowalczyk M., Fita M., (1997), *Microb. & Comp. Genomics*, 2, 259-268.
48. Cebrat S., Dudek M. R., Mackiewicz P., (1998), *Theory Bioscienc.*, 117, 78-89.
49. Rocha E. P. C., (2002), *Trends Microbiol.*, 10, 393-395.
50. French S., (1992), *Science*, 258, 1362-1365.
51. Deshpande A. M., Newlon C. S., (1996), *Science*, 272, 1030-1033.
52. Ellwood M., Nomura M., (1982), *J. Bacteriol.*, 149, 458-468.
53. Brewer B. J., (1988), *Cell*, 53, 679-686.
54. Rocha E. P. C., Danchin A., (2003), *Nat. Genet.*, 34, 377-378.
55. Rocha E. P. C., Danchin A., (2003), *Nucleic Acids Res.*, 31, 6570-6577.
56. Pakula A. A., Sauer R. T., (1989), *Annu. Rev. Genet.*, 23, 289-310.
57. Tillier E. R. M., Collins R. A., (2000), *J. Mol. Evol.*, 51, 459-463.
58. Mackiewicz P., Mackiewicz D., Kowalczyk M., Dudkiewicz M., Dudek M. R., Cebrat S., (2003), *J. Appl. Genet.*, 44, 561-584.
59. Lafay B., Lloyd A. T., McLean M. J., Devine K. M., Sharp P. M., Wolfe K. H., (1999), *Nucleic Acids Res.*, 27, 1642-1649.
60. Mackiewicz P., Szczepanik D., Gierlik A., Kowalczyk M., Nowicka A., Dudkiewicz M., Dudek M. R., Cebrat S., (2001), *J. Mol. Evol.*, 53(6), 615-621.
61. Brookfield J. F., (1997), *Adv. Genet.*, 36, 137-155.
62. Lynch M., Conery J. S., (2000), *Science*, 10, 1151-1155.
63. Conery J. S., Lynch M., (2001), *Pac. Symp. Biocomput.*, 167-178.
64. Kondrashov F. A., Rogozin I. B., Wolf Y. I., Koonin E. V., (2002), *Genome Biol.*, 3, research0008.
65. Andersson S. G., Zomorodipour A., Andersson J. O., Sicheritz-Ponten T., Alsmark U. C., Podowski R. M., Naslund A. K., Eriksson A. S., Winkler, H. H., Kurland C. G., (1998), *Nature*, 396, 133-140.
66. Andersson J. O., Andersson S. G., (2001), *Mol. Biol. Evol.*, 18, 829-839.
67. Cole S. T., Eiglmeier K., Parkhill J., James K. D., Thomson N. R., Wheeler P. R., Honore N., Ganier T., Churcher C., Harris D. i wsp., (2001), *Nature*, 409, 1007-1011.
68. Mira A., Ochman H., Moran N. A., (2001), *Trends Genet.*, 17, 589-596.
69. Homma K., Fukuchi S., Kawabata T., Ota M., Nishikawa K., (2002), *Gene*, 294, 25-33.
70. Liu Y., Harrison P. M., Kunin V., Gerstein M., (2004), *Genome Biol.*, 5, R64.
71. Barker W. C., Dayhoff M. O., (1980), *Bio-Science*, 30, 593-600.
72. Li W. H., (1997), *Molecular Evolution*, Sunderland, MA: Sinauer Associates.
73. Mackiewicz P., Dudkiewicz M., Kowalczyk M., Mackiewicz D., Banaszak J., Polak N., Smolarczyk K., Nowicka A., Dudek M. R., Cebrat S., (2004), *Lecture Notes in Computer Science*, 3039, 687-693.
74. Dudkiewicz M., Mackiewicz P., Mackiewicz D., Kowalczyk M., Nowicka A., Polak N., Smolarczyk K., Banaszak J., Dudek M. R., Cebrat S., (2004), *Biosystems* (in press).
75. Mushegian A. R., Koonin E. V., (1996), *Trends Genet.*, 12, 289-290.
76. Kolsto A. B., (1997), *Mol. Microbiol.*, 24, 241-248.
77. Watanabe H., Mori H., Itoh T., Gojobori T., (1997), *J. Mol. Evol.*, 44(Suppl. 1), S57-S64.
78. Bellgard M. I., Itoh T., Watanabe H., Imanishi T., Gojobori T., (1999), *Ann. N. Y. Acad. Sci.*, 18, 293-300.
79. Itoh T., Takemoto K., Mori H., Gojobori T., (1999), *Mol. Biol. Evol.*, 16, 332-346.
80. Hughes D., (2000), *Genome Biol.*, 1(6), REVIEWS0006.
81. Deschavanne P., Filipski J., (1995), *Nucleic Acids Res.*, 23, 1350-1353.
82. Daubin, V., Perrière, G., (2003), *Mol. Biol. Evol.*, 20, 471-483.

83. Sharp, P.M., Shields D.C., Wolfe K. W., Li W.-H., (1989), *Science*, 246, 808-810.
84. Rocha E. P. C., Danchin A., (2002), *Trends. Genet.*, 18, 291-294.
85. Guindon S., Perrière G., (2001), *Mol. Biol. Evol.*, 18, 1838-1840.
86. Chandler M. G., Pritchard R. H., (1974), *Mol. Gen. Genet.*, 138, 522-529.
87. Louarn J. M., Bouche J. P., Legendre F., Louarn J., Patte J., (1985), *FEMS Microbiol. Rev.*, 26, 533-554.
88. Schmid M. B., Roth J. R., (1987), *J. Bacteriol.*, 169, 2872-2875.
89. Errington J., Daniel R. A., Scheffers D. J., (2003), *Microbiol. Mol. Biol. Rev.* 67, 52-65.
90. Dworkin J., Losick R., (2002), *Proc. Natl. Acad. Sci. USA*, 99, 14089-14094.
91. Rocha E. P. C., Fralick J., Vedyappan G., Danchin A., Norris V., (2003), *Mol. Microbiol.*, 49, 895-903.
92. Corre J., Cornet F., Patte J., Louarn J. M., (1997), *Genetics*, 147, 979-989.
93. Lawrence J. G., Ochman H., (1998), *Proc. Natl. Acad. Sci. USA*, 95, 9413-9417.
94. Kunst F., Ogasawara N., Moszer I., Albertini A. M., Alloni G., Azevedo V., Bertero M. G., Bessieres P., Bolotin A., Borchert S. i wsp., (1997), *Nature*, 390, 249-256.
95. Corre J., Patte J., Louarn J. M., (2000), *Genetics*, 154, 39-48.
96. Danchin A., (2003), *Curr. Issues Mol. Biol.*, 5, 37-42.
97. Mira A., Ochman H., (2002), *Mol. Biol. Evol.*, 19, 1350-1358.
98. Sharp P. M., (1991), *J. Mol. Evol.*, 33, 23-33.
99. Birky C. W. Jr., Walsh J. B., (1992), *Genetics*, 130, 677-783.
100. Goodman M. F., (2000), *Trends Biochem. Sci.*, 25, 189-195.
101. Courcelle J., Hanawalt P. C., (2001), *Proc. Natl. Acad. Sci. USA*, 98, 8196-8202.
102. Bierre H., Ehrlich S. D., Michel B., (1997), *EMBO J.*, 16, 3332-3340.
103. Rocha E. P. C., (2003), *Trends. Genet.*, 19, 600-604.
104. Frank A. C., Amiri H., Andersson S. G., (2002), *Genetica*, 115, 1-12.
105. Achaz G., Coissac E., Netter P., Rocha, E. P. C., (2003). *Genetics*, 164, 1279-1289.
106. Eisen J. A., Heidelberg J. F., White O., Salzberg S. L., (2000), *Genome Biol.*, 1(6), research0011.
107. Read T. D., Brunham R. C., Shen C., Gill S. R., Heidelberg J. F., White O., Hickey E. K., Peterson J., Umayam L. A., Utterback T., et al., (2000), *Nucleic Acids Res.*, 28, 1397-1406.
108. Tillier E. R. M., Collins R. A., (2000), *Nat. Genet.*, 26, 195-197.
109. Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N., Fujii F., Hiramata C., Nakamura Y., Ogasawara i wsp., (2000), *Nucleic Acids Res.*, 28, 4317-4331.
110. Suyama M., Bork P., (2001), *Trends Genet.*, 17, 10-13.
111. Moran N. A., Mira A., (2001), *Genome Biol.*, 2(12), RESEARCH0054.
112. Zivanovic Y., Lopez P., Philippe H., Forterre P., (2002), *Nucleic Acids Res.*, 30, 1902-1910.
113. Lemon K. P., Grossman A. D., (1998), *Science*, 28, 1516-1519.
114. Newport J., Yan H., (1996), *Curr. Opin. Cell Biol.*, 8, 365-368.
115. Mackiewicz P., Mackiewicz D., Kowalczuk M., Cebrat S., (2001), *Genome Biol.*, 2(12), 1004.1-1004.4.
116. Segall A., Mahan M. J., Roth J. R., (1988), *Science*, 241, 1314-1318.
117. Liu S. L., Sanderson K. E., (1995), *Proc. Natl. Acad. Sci. USA*, 92, 1018-1022.
118. Liu S. L., Sanderson K. E., (1996), *Proc. Natl. Acad. Sci. USA*, 93, 10303-10308.
119. Wu L. J., Errington J., (2002), *EMBO J.*, 21, 4001-4011.
120. Campo N., Dias M. J., Daveran-Mingot M. L., Ritzenthaler P., Le Bourgeois P., (2004), *Mol. Microbiol.*, 52, 511-522.
121. Ikemura T., (1981), *J. Mol. Biol.*, 151, 389-409.
122. Gouy M., Gautier C., (1982), *Nucleic Acids Res.*, 10, 7055-7074.
123. Sharp P. M., Li W. H., (1987), *Nucleic Acids Res.*, 15, 1281-1295.
124. Mackiewicz D., Mackiewicz P., Kowalczuk M., Dudkiewicz M., Dudek M. R., Cebrat S., (2003), *Acta Microb. Polon.*, 52(3), 245-261.
125. Koonin E. V., Galperin M. Y., (1997), *Curr. Opin. Genet. Dev.*, 7, 757-763.
126. Tamames J., Ouzounis C., Casari G., Valencia A., (1997), *J. Mol. Evol.*, 44, 66-73.