



Biblioteki DNA – droga do poznania organizacji genomów

Łukasz Wolko, Ryszard Słomski

Katedra Biochemii i Biotechnologii, Akademia Rolnicza
im. Augusta Cieszkowskiego, Poznań

DNA libraries – a way to understand genome structure

Summary

Wide genome physical mapping is becoming a main goal of current genomic research on many plant and animal species. It provides essential basis for large-scale genome sequencing, but requires well-saturated large-insert DNA genomic libraries. This review presents a progress in the construction and application of genomic libraries. New trends towards the improvement of cloning vectors and adjustment to new tasks of molecular genetics are also presented.

Key words:

DNA library, cloning, vector, plasmid, bacteriophage lambda, cosmid, bacteriophage P1, PAC, YAC, BAC, TAR, TAC, BIBAC.

1. Wstęp

Molekularna analiza genetycznych podstaw funkcjonowania systemów biologicznych jest jednym z podstawowych narzędzi pozwalających na poznawanie mechanizmów rządzących światem ożywionym. Techniki używane do tego rodzaju analiz są zazwyczaj opracowywane na organizmach modelowych, których (małe zazwyczaj) genomy stają się obiektem intensywnych badań genetycznych. U wyższych organizmów eukariotycznych, ssaków i roślin, badania napotykały dotychczas techniczne trudności związane z wielkością i złożonością ich genomów. W ciągu ostatnich lat postęp w opracowywaniu nowych technik w zakresie mapowania fizycznego i klonowania pozwolił na wypełnienie luki pomiędzy klasyczną analizą genetyczną a badaniami mole-

Adres do korespondencji

Łukasz Wolko,
Katedra Biochemii
i Biotechnologii,
Akademia Rolnicza
im. Augusta
Cieszkowskiego,
ul. Wołyńska 35,
60-637 Poznań;
e-mail:
lukwolko@au.poznan.pl

kularnymi organizmów o dużych genomach. W badaniach złożonych genomów bardzo pożyteczne okazują się biblioteki DNA. Biblioteki tworzone są przez fragmentację genomu i klonowanie uzyskanych fragmentów w różnego rodzaju wektorach (np. plazmidy, fagi lambda i P1, kosmidy, YAC, BAC) w taki sposób, aby pokrywały cały badany genom.

Liczba klonów niezbędnych do dobrego pokrycia genomu przez bibliotekę zależy od wielkości sklonowanych fragmentów DNA. Liczbę tę można obliczyć ze wzoru Clarke'a i Carbona (1), $N = \ln(1-p)/\ln(1-x/y)$, gdzie liczba klonów (N) daje prawdopodobieństwo (p), że jakakolwiek sekwencja będzie reprezentowana w bibliotece o średniej długości insertu (x) i wielkości genomu (y). Przy średniej wielkości insertu 20 kpz, prawdopodobieństwo znalezienia zatem sekwencji w bibliotece genomu o wielkości 400 Mpz i liczbie klonów 1×10^5 jest większe niż 0,99. Jest to obliczenie dla biblioteki o pięciokrotnym pokryciu i przypadkowej dystrybucji fragmentów. W praktyce należy uwzględnić preferencje hydrolizy niektórych enzymów restrykcyjnych i dlatego dobrze reprezentatywne biblioteki osiągają stan dziesięciokrotnego pokrycia (2). Z analizy wynika, że średnia wielkość fragmentów DNA ma bezpośredni wpływ na liczbę klonów biblioteki niezbędnych do pokrycia całego genomu. Większa liczba klonów powoduje utrudnienia w operowaniu i przeglądaniu biblioteki dlatego w badaniach nad strukturą dużych genomów eukariotycznych preferowane są biblioteki o średniej wielkości insertów przekraczającej 1 Mpz (tab. 1).

Tabela 1

Liczba klonów niezbędna do całkowitego pokrycia haploidalnego genomu przy średniej wielkości fragmentów w bibliotece 40 kpz, 150 kpz i 500 kpz, na przykładzie genomów wybranych roślin

Organizm	Haploidalny genom Mpz	Kosmidy 40 kpz	BAC 150 kpz	YAC 500 kpz
<i>Allium cepa</i> (cebula)	15 290	$1,8 \times 10^6$	$4,7 \times 10^5$	$1,4 \times 10^5$
<i>Arabidopsis thaliana</i> (rzodkiewnik)	145	$1,7 \times 10^4$	$4,5 \times 10^3$	$1,3 \times 10^3$
<i>Avena sativa</i> (dąb)	11 315	$1,3 \times 10^6$	$3,5 \times 10^5$	$1,0 \times 10^5$
<i>Glycine max</i> (soja)	1 115	$1,3 \times 10^5$	$3,4 \times 10^4$	$1,0 \times 10^4$
<i>Hordeum vulgare</i> (jęczmień)	4 873	$5,6 \times 10^5$	$1,5 \times 10^5$	$4,5 \times 10^4$
<i>Lycopersicon esculentum</i> (pomidor)	953	$1,1 \times 10^5$	$2,9 \times 10^4$	$8,8 \times 10^3$
<i>Nicotiana tabacum</i> (tytoń)	4 434	$5,1 \times 10^5$	$1,4 \times 10^5$	$4,1 \times 10^4$
<i>Oryza sativa</i> (ryż)	431	$5,0 \times 10^4$	$1,3 \times 10^4$	$4,0 \times 10^3$
<i>Triticum aestivum</i> (pszenica)	15 966	$1,8 \times 10^6$	$4,9 \times 10^5$	$1,5 \times 10^5$
<i>Zea mays</i> (kukurydza)	2 504	$2,9 \times 10^5$	$7,7 \times 10^4$	$2,3 \times 10^4$

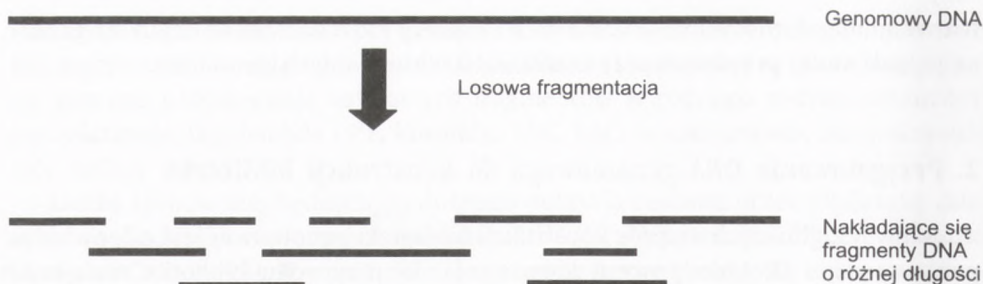
Należy również wspomnieć o bibliotekach cDNA, które w odróżnieniu od bibliotek genomowego DNA, zawierają zestaw sekwencji mRNA danego organizmu. Są one konstruowane przez enzymatyczną syntezę cząsteczek DNA na matrycy całkowitego RNA wyizolowanego z komórek i zawierają fragmenty DNA będące kopiami genów podlegających ekspresji w danych warunkach (tzw. cDNA). Tego typu biblio-

teki znajdują zastosowanie w badaniach ekspresji i mechanizmów działania genów, są jednak mniej przydatne przy ustalaniu struktury całych genomów.

2. Przygotowanie DNA genomowego do konstrukcji biblioteki

Jednym z głównych etapów konstrukcji biblioteki genomowej jest odpowiednie przygotowanie DNA do procesu klonowania. W przypadku bibliotek mniejszych fragmentów (do 40 kpz) chodzi przede wszystkim o izolację dobrze oczyszczonego DNA. Problemy pojawiają się przy konstrukcji bibliotek dużych fragmentów, ponieważ podczas standardowych izolacji duże cząsteczki genomowego DNA ulegają fizycznej degradacji do fragmentów o długości 20-50 kpz. Izolacja DNA do konstrukcji bibliotek o wielkości insertów powyżej 1 Mbp odbywa się poprzez unieruchomienie komórek, protoplastów lub jąder komórkowych w małych bloczkach żelu agarozowego w postaci korków lub mikropaciorków. Agaroz tworzy sieć przestrzenną, zabezpieczającą DNA przed fizycznym uszkodzeniem, a jednocześnie umożliwia dyfuzję odczynników. Unieruchomione w żelu komórki, są poddawane hydrolizie przez trawienie proteinazą K w obecności sarkozyłu. Uwolniony w ten sposób DNA jest poddawany częściowej hydrolizie enzymami restrykcyjnymi. Fragmenty DNA, przed wprowadzeniem do wektorów selekcjonuje się pod względem wielkości przez rozdział elektroforetyczny w pulsacyjnym polu elektrycznym. W przypadku roślin dodatkowym utrudnieniem jest konieczność uzyskania protoplastów lub jąder komórkowych, ponieważ ściana komórkowa uniemożliwia izolację wysokocząsteczkowego DNA (3,4).

Jeśli tworzona biblioteka ma służyć do całkowitego sekwencjonowania danego genomu to niezbędne stanie się uporządkowanie klonów biblioteki w celu rekonstrukcji naturalnego układu DNA organizmu. Możliwe jest to jedynie w przypadku, gdy fragmenty DNA w bibliotece częściowo, wzajemnie nakładają się (rys. 1). Taka budowa biblioteki osiągnana jest przez losową fragmentację DNA. Fragmentacja przez sonifikację genomowego DNA, spełnia ten warunek, jednak fragmenty uzyskane w ten sposób sprawiają trudności w klonowaniu (2). Lepszym sposobem, jak się wydaje, jest tzw. niepełna hydroliza enzymatyczna DNA, osiągnana przez wcześniejsze przerwanie reakcji hydrolizy restrykcyjnej, wiążące się z niedopuszczeniem do przecięcia wszystkich sekwencji rozpoznawanych przez enzym w obrębie genomu. Innym rozwiązaniem jest stosowanie metody polegającej na współzawodnictwie podczas jednoczesnego działania na DNA enzymu restrykcyjnego rozpoznającego metylowane nukleotydy i metylazy (np. enzym *MboI* rozpoznający GATC i metylaza *dam* metylująca adeninę) (5). W bibliotekach DNA zbudowanych ze wzajemnie pokrywających się fragmentów możliwe jest zastosowanie techniki „spaceru po chromosomie” (*chromosome walking*) w celu uporządkowania klonów o poznanej sekwencji.



Rys. 1. Otrzymywanie nakładających się fragmentów DNA przez losową fragmentację genomowego DNA.

3. Rodzaje wektorów

Najistotniejszą sprawą przy tworzeniu biblioteki DNA jest dobór odpowiedniego wektora do klonowania. Pojawienie się w badaniach nowych wektorów umożliwia klonowanie coraz większych fragmentów DNA, pozwala stosować prostsze procedury wprowadzania rekombinowanego wektora do komórek gospodarza, efektywniej namnażać i przeglądać klony biblioteki. Postęp w dziedzinie konstrukcji coraz doskonalszych bibliotek DNA można najłatwiej przedstawić śledząc udoskonalania wektorów stosowanych do tworzenia bibliotek DNA.

3.1. Plazmidy

Większość stosowanych obecnie plazmidów jest pochodnymi pBR322, skonstruowanego przez Bolivara i wsp. (6). Ten macierzysty wektor był w kolejnych latach poddawany licznym ulepszeniom, dając początek wielu różnym typom plazmidów oferowanych obecnie komercyjnie, dopasowanych do różnych specyficznych zastosowań, takich jak sekwencjonowanie (np. pPCR, pBluescript) czy transkrypcja i ekspresja (pET20, pQE70).

Wektory plazmidowe są stosunkowo małe i zwykle pozwalają na klonowanie fragmentów DNA mniejszych od własnej długości. Zazwyczaj wydajność transformacji jest ujemnie skorelowana z wielkością insertu, a długość plazmidu wraz z włączonym fragmentem rzadko może przekroczyć 15 kpz. Z powodu małej pojemności plazmidy nie znajdują szerszego zastosowania w tworzeniu w pełni reprezentatywnych bibliotek genomowych. Dlatego termin „biblioteka plazmidowa”, choć powszechnie stosowany jest błędny. Bardziej adekwatne jest, jak się wydaje, określenie „plazmidowy bank klonów” (7).

Plazmidy stosuje się do tworzenia banków klonów fragmentów, których długość nie przekracza 1 kpz. Są one stosowane do:



Rys. 2. Plazmid pBluescript II KS+ (Stratagen).

- Pozyskiwania sond RFLP, gdzie biblioteka tworzona jest przez hydrolizę genomu często tnącą endonukleazą restrykcyjną (8-10).
- Izolacji sekwencji specyficznych gatunkowo do analiz produktów fuzji somatycznych, hybrydów międzygatunkowych (11), amfiploidów i ich pochodnych (12).
- Poznawania sekwencji powtórzonych do identyfikacji markerów (13,14).
- Pozyskiwania chromosomospecyficznych sond molekularnych (15).
- Subklonowania dużych insertów z innych wektorów w celu sekwencjonowania.
- Klonowania produktów PCR.
- Tworzenia bibliotek ze śladowych ilości wyjściowego DNA przez DOP-PCR (*degenerate oligonucleotide primed PCR*) (16-21) lub LAM-PCR (*linker adaptor-mediated PCR*) (22,23).

Do najczęściej stosowanych należą wektory plazmidowe serii pBluescript (Stratagen) (rys. 2). Plazmidy te należą do wektorów bardzo uniwersalnych; są małe (2,96 kpz), występują w dużej liczbie kopii w komórce, zawierają gen oporności na ampicylinę (do selekcji transformowanych komórek *E. coli*) i miejsce wielokrotnego klonowania (*multiple cloning site* lub *polylinker*) rozpoznawane przez 21 enzymów restrykcyjnych. Miejsce to usytuowane jest w N-końcowym fragmencie genu *lacZ*, co umożliwi selekcję transformantów testem α -komplementacji. W razie potrzeby plazmidy pBluescript, w obecności faga pomocniczego M13, mogą być izolowane w postaci jednoniciowej w celu sekwencjonowania lub ukierunkowanej mutagenyzy (plazmidy posiadające tę właściwość nazywane są również fagemidami). Miejsce wielokrotnego klonowania jest oflankowane przez promotory polimeraz RNA T7 i T3, które mogą służyć do syntezy RNA *in vitro*. Usytuowane po obu stronach miejsca wielokrotnego klonowania sekwencji wiązania się trzech par różnych starterów: M13F/M13R, T3/T7 oraz KS/SK, umożliwiają amplifikację insertu metodą PCR oraz sekwencjonowanie.

3.2. Bakteriofagi lambda

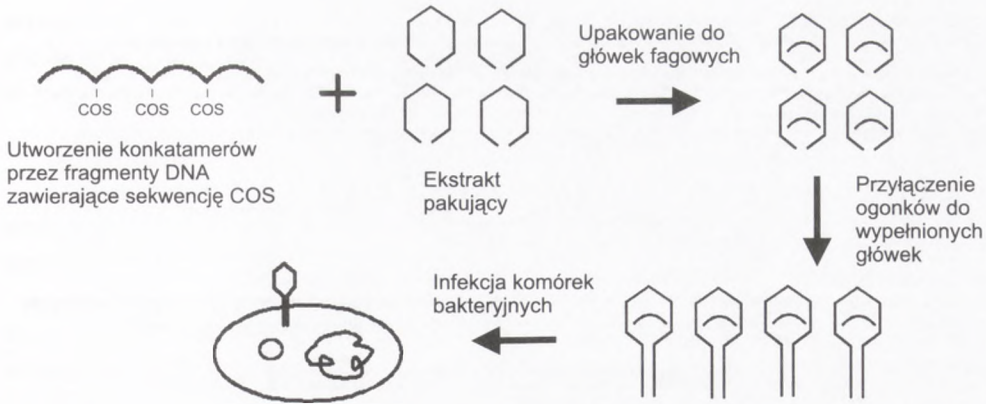
Wektory typu lambda, pomimo niewielkiej pojemności, pozwalają w prosty i wydajny sposób tworzyć reprezentatywne biblioteki, o niskiej liczbie klonów chimerycznych i nie rekombinowanych (2).

Wektor zbudowany jest z dwóch ramion (9 i 19 kpz) flankujących właściwy genom faga. Przy konstrukcji biblioteki wewnętrzny fragment DNA jest wymieniany na insert genomowego DNA. Ramiona zawierają sekwencje niezbędne do namnażania faga. Miejsca wielokrotnego klonowania usytuowane są zazwyczaj na wewnętrznej granicy obu ramion. Aby konstrukcja mogła być upakowana do kapsydu, całkowita wielkość genomu faga, wliczając w to długości obu ramion (ok. 29 kpz) i insertu nie może być mniejsza niż 39 i większa niż 52 kpz. Dlatego do wektorów lambda, możliwe jest wprowadzenie fragmentów o długości 10-23 kpz. Konstrukcje zawierające fragment krótszy niż 10 kpz nie podlegają pakowaniu do kapsydu, co powoduje naturalną eliminację krótszych insertów.

Upakowanie faga do kapsydu przy użyciu komercyjnie dostępnych ekstraktów pakujących (mieszanina białek kapsydu faga oraz enzymów biorących udział w procesie pakowania) umożliwia bardzo wydajne wprowadzenie DNA do komórek bakteryjnych. Proces pakowania fagów zapoczątkowywany jest przez utworzenie konkatamerów przy udziale końców kohezyjnych (COS), położonych na zewnątrz ramion. DNA jest pakowany do kapsydów tylko, wtedy gdy jest zachowany odpowiedni układ fragmentów (lewe ramię – insert – prawe ramię) i występuje odpowiednia odległość pomiędzy końcami COS (39-52 kpz) (rys. 3).

Przeglądanie biblioteki fagów lambda jest bardzo wydajne i powtarzalne. DNA faga przenosi się z łysinek przez wykonanie repliki na filtr nylonowy. Hybrydyzacja utrwalonego na filtrze DNA z odpowiednią sondą pozwala na identyfikację łysinek zawierających rekombinanty poszukiwanego fragmentu DNA. Powyżej 50 000 łysinek można wydajnie przeglądać na pojedynczej, dużej płytce Petriego, co stanowi reprezentację ok. 1000 Mpz DNA insertów.

Nowsze warianty wektorów lambda umożliwiają sprawniejsze operowanie biblioteką i zapewniają dodatkowe możliwości zastosowań. Jednym z przykładów może być wektor Lambda ZAP Express (Stratagene), pozwalający na wycięcie *in vivo* klonowanego fragmentu z wektora lambda i zamknięcie go w postaci jednoniciowej cząsteczki fagemidu pBK-CMV. Wycięcie to następuje dzięki obecności specjalnej sekwencji wektora lambda i szeregu białek kodowanych przez tzw. faga pomocniczego M13. Genomowy DNA jest klonowany do wektora lambda, a następnie przez transfekcję wprowadzany do *E. coli*. Jednocześnie komórki bakteryjne są infekowane fagiem pomocniczym M13. Wewnątrz komórki *E. coli* białka kodowane przez faga pomocniczego rozpoznają sekwencję replikonu faga f1 usytuowaną na jednym z ramion wektora lambda, inicjując replikację jednoniciowego DNA. Proces inicjacji rozpoczyna się nacięciem jednej nici DNA i utworzony w ten sposób koniec 3' służy jako starter do replikacji całego regionu faga lambda, aż do sygnału terminacji repli-



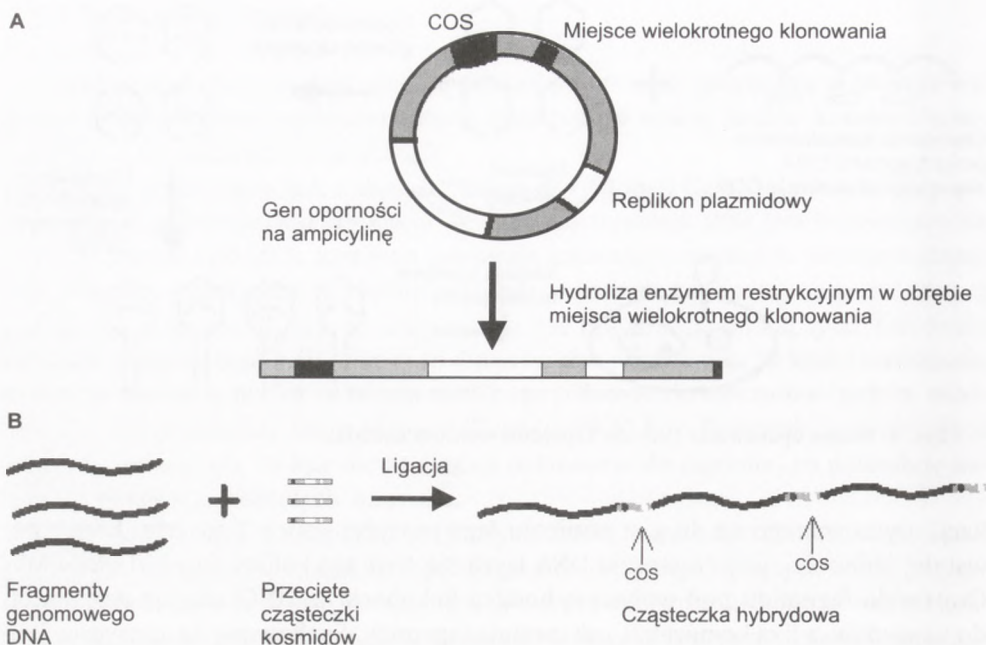
Rys. 3. Proces upakowania DNA do kapsydów wektora lambda.

kacji usytuowanego na drugim ramieniu faga powyżej końca 3' insertu. Końce powstałej jednoniciowej cząsteczki DNA łączą się tworząc kolisty fagemid pBK-CMV. Cząsteczki fagemidu pod wpływem bodźca (inkubacja w 70°C) ulegają pakowaniu do kapsydów, a liza komórek *E. coli* uwalnia fagemidy. Spakowane do kapsydów jednoniciowe cząsteczki mogą być wygodnie przechowywane i w miarę potrzeby użyte do transformacji bakterii. Po wnikięciu do komórki nowego gospodarza następuje synteza drugiej nici i cząsteczka zachowując się jak normalny plazmid może być łatwo namnażana i izolowana, a sekwencja insertu może ulegać ekspresji pod kontrolą promotorów T7 i T3.

3.3. Kosmidy

Kosmidy to wektory o stosunkowo dużej pojemności (35-45 kpz). Typowe kosmidy zawierają sekwencje COS wektora lambda, miejsce inicjacji replikacji plazmidu (często jest to ColE1) i gen oporności na antybiotyk. Sekwencja COS umożliwia pakowanie wektora do kapsydu fagowego i transformację bakterii. Plazmidowe miejsce inicjacji replikacji i gen markerowy umożliwiają namnażanie wektora jak plazmidu. Połączenie cech plazmidu i bakteriofaga lambda pozwoliło na uzyskanie dużej pojemności i wydajnej transformacji.

Wektory kosmidowe (np. SuperCos I) są dostępne w postaci kolistej (7,6 kpz). W takiej formie wektor może być namnażany i izolowany podobnie jak plazmid. Przed klonowaniem, hydroliza wektora enzymem restrykcyjnym w obrębie miejsca wielokrotnego klonowania, uwalnia sekwencję COS. Fragmenty genomowego DNA są łączone z liniowymi cząstkami kosmidu, tworząc długie cząsteczki hybrydowe, podlegające następnie hydrolizie przez specyficzne nukleazy w procesie pakowania do kapsydów faga lambda (rys. 4). Upakowany DNA wprowadzany jest do komórek



Rys. 4. Przygotowanie wektora kosmidowego do klonowania. A – Hydroliza kolistej cząsteczki wektora powodująca uwolnienie ramienia COS; B – Ligacja fragmentów DNA genomowego z wektorem i utworzenie długich cząsteczek hybrydowych.

E. coli przez infekcję, gdzie formuje koliste cząsteczki podlegające replikacji pod kontrolą plazmidowego miejsca inicjacji replikacji. Przeglądanie biblioteki odbywa się przez hybrydyzację do kolonii, a izolacja wektorów ze zidentyfikowanych rekombinantów jest zbliżona do izolacji plazmidowego DNA.

Wraz ze wzrostem pojemności wektorów coraz istotniejszy staje się problem charakterystyki insertów. Odbywa się ona przez tworzenie dokładnych map restrykcyjnych insertu, klonowanie jego fragmentów w plazmidy, sekwencjonowanie i w końcu rekonstrukcję sekwencji całego, dużego fragmentu bazując na opracowanej wcześniej mapie restrykcyjnej. Wielkość insertów w przypadku wektorów kosmidowych osiągnęła poziom, przy którym opracowanie mapy za pomocą kilku hydroliz endonukleazami restrykcyjnymi jest utrudnione z powodu powstawania zbyt wielu fragmentów. Dlatego w tworzeniu map insertów w kosmidach wykorzystuje się dodatkowo metody oparte na hybrydyzacji z sekwencją COS lub z promotorami T7 i T3 na końcu wektora (mapowanie „cos”), lub posługuje się transkryptami z tych promotorów (mapowanie transkrypcyjne) (24).

Połączenie cech plazmidu i bakteriofaga lambda pozwoliło, w przypadku kosmidów, na uzyskanie dużej pojemności i wydajnej transformacji. Należy jednak zwrócić uwagę na utrudnione, w porównaniu do fagów lambda, przeglądanie biblioteki i jej przecho-

wywanie w postaci wyjściowych hodowli bakteryjnych. Podstawowym technicznym problemem pojawiającym się przy konstrukcji i utrzymywaniu biblioteki genomowej w wektorach kosmidowych jest niska stabilność insertów powodowana przez:

- Tworzenie się konkatermerów cząstek wektora powodujące utratę insertu.
- Wprowadzenie więcej niż jednego fragmentu obcego DNA do jednego kosmidu.
- Niestabilność insertów, wynikająca z rekombinacji pomiędzy elementami powtórzonymi insertu, mogąca powodować rearanżacje lub utratę fragmentów klonowanego DNA.

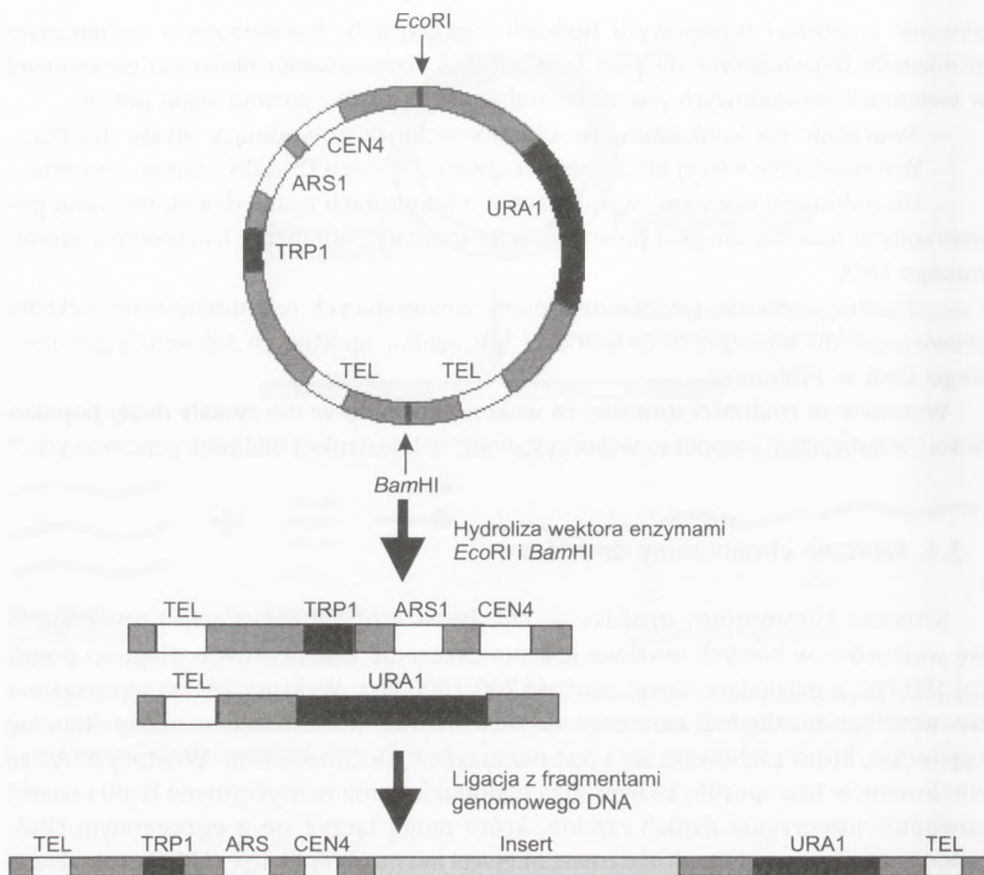
- Różnice szybkości podziałów klonów zawierających rekombinowane wektory prowadzące do nadreprezentatywności lub zaniku niektórych sekwencji genomowego DNA w bibliotece.

Wszystkie te trudności sprawiły, że wektory kosmidowe nie zyskały dużej popularności i w mniejszym stopniu są wykorzystywane w konstrukcji bibliotek genomowych.

3.4. Sztuczne chromosomy drożdżowe

Sztuczne chromosomy drożdży – YAC (*yeast artificial chromosome*) otwierają listę wektorów, w których możliwe jest umieszczenie fragmentów o długości powyżej 100 kpz, a osiągające nawet wartość 700-1800 kpz. Wektory YAC są wyposażone we wszelkie niezbędne elementy do utworzenia w komórce drożdży liniowej cząsteczki, która zachowuje się i jest namnażana jak chromosom. Wektory pYAC są zbudowane w taki sposób, że hydroliza endonukleazami restrykcyjnymi *EcoRI* i *BamHI* powoduje utworzenie dwóch ramion, które mogą łączyć się z egzogennym DNA. Jedno z ramion zawiera drożdżowe miejsce inicjacji replikacji (ARS1), sekwencję centromerową (CEN4), gen markerowy syntezy tryptofanu (TRP1) i sekwencję telomerową (TEL). Drugie ramię zawiera gen markerowy syntezy uracylu (URA1, kodujący dehydrogenazę dwuhydroorotonianu) i sekwencję telomerową (TEL) (rys. 5). Pierwotny wektor przed hydrolizą może być namnażany w *E. coli* ponieważ zawiera sekwencję inicjacji replikacji plazmidu pBR322 i gen selekcyjny oporności na ampicylinę. W dalszych wariantach wektora ramiona zostały zaopatrzone w dodatkowe miejsca rozpoznawane przez restryktazy, promotory T7 i T3 oraz dodatkowe geny selekcyjne. Pewnym ułatwieniem klonowania było wprowadzenie systemu opartego na dwóch plazmidach pJS97-pJS98, z których każdy zawiera jedno z ramion właściwego wektora i może być niezależnie namnażany (25).

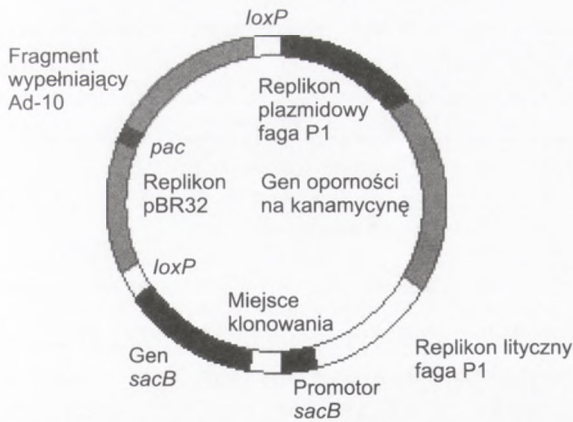
YAC w komórce drożdżowej jest stabilnie dziedziczony w kolejnych pokoleniach jako pojedyncza kopia cząsteczki. Ta cecha powoduje problemy z interpretacją słabych sygnałów w czasie przeglądania bibliotek. Problem ten rozwiązano przez konstrukcję tzw. „warunkowego centromeru”. Sekwencja centromerowa usytuowana jest poniżej regulowanego promotora drożdżowego. Podczas hodowli w warunkach indukcyjnych funkcje centromeru są zakłócone przez wysoki poziom transkrypcji z promotora i liczba kopii cząsteczki YAC wzrasta. Jeśli promotor zostaje



Rys. 5. Proces klonowania do wektora YAC.

zablokowany przez zmianę warunków hodowli, przywrócona zostaje funkcja centromeru i liczba kopii spada w ciągu kilku pokoleń (26,27).

Biblioteki genomowe w wektorach YAC były konstruowane dla szeregu gatunków roślin (28-32). Biblioteki tego typu były też stosowane w programie poznania genomu człowieka (33). Stosowanie bibliotek opartych na wektorach YAC wiązało się jednak z szeregiem problemów technicznych. Jednym z nich był wysoki odsetek klonów chimerowych, powstałych w czasie klonowania, ale także przez rekombinację w komórce drożdży. Klony wykazywały również tendencje do wewnętrznych delecji (34). Dużą trudność stanowiło również oddzielenie wektorów YAC od 15 Mbp chromosomu drożdży. Jest to w zasadzie możliwe tylko przy użyciu elektroforezy w polu pulsacyjnym (PFGE). Mało wydajna jest również transformacja komórek drożdżowych za pośrednictwem sferoplastów. Wszystkie te trudności, a także szereg niedogodności wiążących się z hodowlą drożdży, miały wpływ na niemal całkowitą rezygnację z ich stosowania na rzecz bardziej nowoczesnych wektorów.



Rys. 6. Budowa wektora P1 pAd10-SacBII (35).

3.5. Wektory P1

Jedną z alternatyw dla wektorów YAC okazało się zastosowanie jako wektorów fagów P1. Wektor ten ma zdolność przenoszenia fragmentów o długości do 100 kpz. Jego pojemność jest zatem około dwukrotnie większa niż kosmidów, ale znacznie mniejsza niż wektorów YAC. Wektory P1 posiadają jednak szereg korzystnych cech rekompensujących braki pojemności. Dzięki temu, że DNA jest pakowane do kapsydu fagowego i wprowadzane do *E. coli* na zasadzie infekcji, uzyskuje się bardzo wysoką wydajność transformacji. O wiele wyższa jest również stabilność insertów.

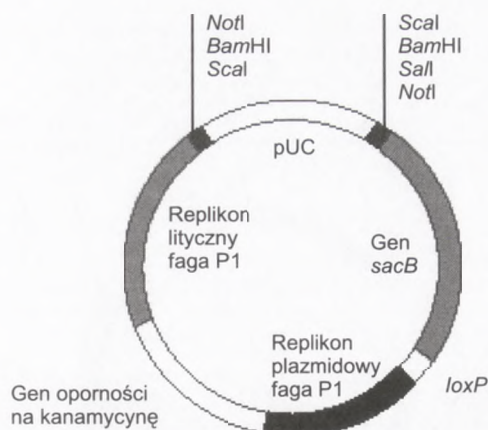
Wektory P1 zawierają sekwencję *pac*, która pozwala na pakowanie *in vitro* klonowanych fragmentów do otoczek faga. Posiadają one także dwa miejsca *loxP* rozpoznawane przez rekombinazę P1 *cre*, replikony P1 (plazmidowy i lityczny), a także gen selekcyjny oporności na ampicylinę. Selekcja pozytywna zachodzi ze względu na obecność genu *sacB* kodującego enzym transferazę sacharylo-6-fruktozową, którego produkty są silnie toksyczne dla komórki *E. coli*. Promotor genu *sacB* znajduje się powyżej miejsca klonowania obcego DNA, dzięki czemu wprowadzenie dużego insertu DNA do wektora powoduje unieczynnienie genu (35) (rys. 6). Klonowanie do wektora pAd10-SacBII (30 kpz) rozpoczyna się uwolnieniem dwóch ramion. Wyselekcjonowane pod względem wielkości fragmenty DNA (85-100 kpz) przyłączone do dwóch ramion są gotowe do dwustopniowego procesu pakowania (35). W pierwszym etapie wektor jest hydrolizowany w miejscu *pac* przez specyficzną endonukleazę faga. Od przeciętej sekwencji *pac* rozpoczyna się pakowanie fragmentu do główki faga. Po wypełnieniu główki odcinkiem długości 115 kpz następuje niespecyficzne odcięcie reszty DNA, a następnie główki są łączone z ogonkami faga. Fagi służą do transformacji *E. coli* szczepu *cre*⁺. Rekombinaza *cre* działając na sekwencje *loxP* powoduje utworzenie kolistych cząstek zawierających inserty oraz gen oporności na ampicylinę. Takie plazmidopodobne cząstki utrzymują się w komórce w licz-

bie jednej kopii, dzięki replikacji regulowanej przez replikon plazmidowy P1. W tym czasie replikon lityczny faga, odpowiedzialny za wydajną replikację wielu kopii wektora jest zablokowany przez represor *lac I*. W razie potrzeby izolacji cząsteczek z komórek bakteryjnych, indukcja za pomocą IPTG powoduje odblokowanie replikonu litycznego i 25-krotne zwiększenie liczby kopii (35).

3.6. Wektory PAC

Sztuczne chromosomy wprowadzone z faga P1-PAC (*P1-derived artificial chromosomes*) są kolejnym krokiem na drodze udoskonalania wektorów stosowanych do tworzenia bibliotek genomowych. Wektory tego typu mogą służyć do klonowania fragmentów o długości powyżej 140 kpz. Do konstrukcji wektorów pCYPAC posłużono się wcześniej opracowanym przez Pierce'a i wsp. (33) wektorem pAd10-SacBII. Dokonując dosyć dużych przegrupowań w pierwotnym wektorze P1 doprowadzono do znacznego zmniejszenia wielkości cząsteczki (pAd10-SacBII – 30 kpz, pCTPACII – 18,7 kpz) (rys. 7). Dla poprawienia wydajności i precyzji procesu klonowania obcego DNA usunięto fragment wypełniający wektora P1, a w jego miejsce, w pozycji rozpoznawanej przez endonukleazę restrykcyjną *Bam*HI, wprowadzono sekwencję plazmidu pUC (36). W procesie przygotowania wektora do klonowania, fragment pUC zostaje wycięty przez podwójne trawienie przy użyciu enzymów *Bam*HI i *Scal*. Aby zapobiec powrotnej ligacji fragmentu pUC do wektora, proces ten prowadzony jest w kilku etapach. Po wycięciu sekwencji pUC przez endonukleazę *Bam*HI fragment ten ponownie poddawany jest trawieniu, tym razem przez enzym *Scal*, powodując utworzenie końców niekompatybilnych z *Bam*HI. Następnie łącznikowe oligonukleotydy *Bam*HI/*Scal* są fizycznie oddzielane od wektorów PAC (np. na złożu magnetycznym). Uzyskane końce są poddawane działaniu fosfatazy alkalicznej w celu zapobiegnięcia ligacji wektor-vektor. W kolejnych wariantach tego wektora miejsce rozpoznawane przez enzym *Scal* wymieniono na sekwencje rozpoznawane przez rzadko tnący enzym *Not*I, używany zwykle do fragmentacji genomów. Wektory te są zazwyczaj wprowadzane do komórek *E. coli* poprzez elektroporację.

Podniesienie pojemności i poprawa stabilności insertów sprawiły, że wektory PAC są obecnie, obok wektorów BAC, szeroko stosowane do konstrukcji bibliotek genomowego DNA. Jedyną niedogodnością stosowania wektorów PAC jest ich stosunkowo duży rozmiar (19,5 kpz). Przy przeglądaniu biblioteki przez subklonowanie i losowe sekwencjonowanie (*shotgun sequencing*) powoduje to powiększenie frakcji klonów zawierających sekwencje wektora, a co za tym idzie, podniesienie kosztów badań.

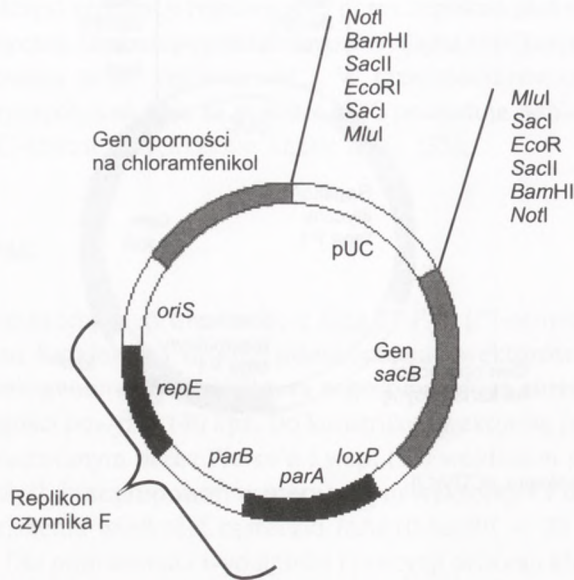


Rys. 7. Budowa wektora pCTPACII.

3.7. Sztuczne chromosomy bakteryjne (BAC)

Ważnym krokiem na drodze doskonalenia wektorów do tworzenia pełnych bibliotek do mapowania i analiz złożonych genomów było skonstruowanie wektorów BAC (*bacterial artificial chromosome*) (37). Wektor ten oparty jest funkcjonalnie na dobrze opisanym czynniku płci *E. coli*. Jest to duża cząsteczka utrzymująca się w kolejnych pokoleniach bakterii w postaci jednej lub dwóch kopii. Cecha ta regulowana jest przez produkty własnych genów: *oriS*, *repE*, *parA*, i *parB*. Geny *oriS* i *repE* uczestniczą w jednokierunkowej replikacji czynnika F, *parA* i *parB*, natomiast utrzymują liczbę kopii na poziomie jednej do dwóch na komórkę *E. coli* (rys. 8). Cecha ta zabezpiecza wektor przed wzajemnymi rekombinacjami. Wektory BAC zawierają również gen oporności na chloramfenikol (Cm^r), sekwencję COS z faga lambda, sekwencję *loxP* z bakteriofaga P1, dwa miejsca klonowania (*HindIII* i *BamHI*) oraz kilka bogatych w pary G/C miejsc rozpoznawanych przez enzymy restrykcyjne, które mogą służyć do wycięcia insertu. Miejsce klonowania jest oskrzydłone przez promotory T7 i SP6 wykorzystywane do uzyskiwania sond RNA przydatnych w mapowaniu fragmentów. Sekwencje COS i *loxP* mogą być hydrolizowane przez specyficzne enzymy – terminazę faga lambda i rekombinazę *cre* faga P1. Cięcie w tych miejscach może generować fragmenty przydatne w mapowaniu restrykcyjnym. W dalszych modyfikacjach wektory BAC zostały uzupełnione o łącznik pUC i gen selekcyjny *sacB*.

Wektory BAC zawierające inserty długości powyżej 300 kbp są w wydajny sposób wprowadzane do bakterii przez elektroporację. Fakt, że sztuczne chromosomy bakteryjne są cząsteczkami kolistymi ułatwia ich izolację i zwiększa ich stabilność. Potwierdzony jest niski poziom rekombinacji i pojawiania się klonów chimerowych przez więcej niż 100 pokoleń (37).

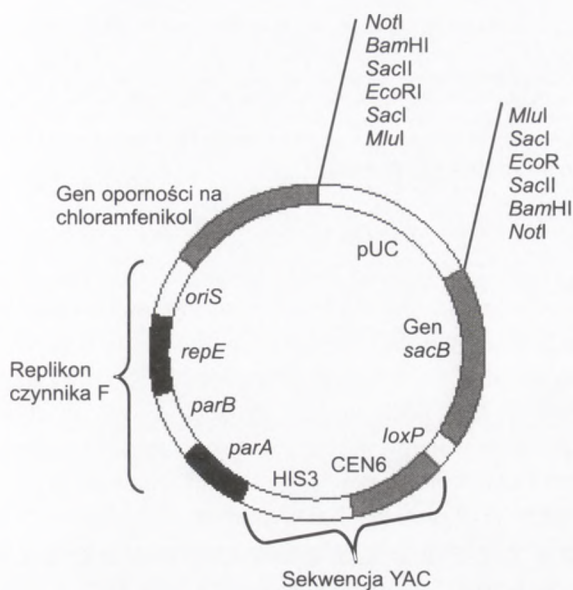


Rys. 8. Budowa wektora BAC.

Biblioteki genomowe w wektorach BAC były konstruowane dla szeregu roślin (38-40). Wektory tego typu zostały także użyte do konstrukcji bibliotek wykorzystywanych w projekcie sekwencjonowania genomu człowieka (41).

3.8. Konstrukcja bibliotek YAC/BAC przy udziale rekombinacji zasocjowanej z transformacją TAR

Techniką, która w znaczący sposób może usprawnić skomplikowane procedury laboratoryjne związane z konstrukcją bibliotek dużych fragmentów jest opracowana przez Larionova i wsp. rekombinacja sprzężona z transformacją (TAR, *transformation-associated recombination*) (42). Metoda opiera się na wykorzystaniu *Saccharomyces cerevisiae* jako organizmu charakteryzującego się dużą zdolnością do homologicznych rekombinacji DNA. Wektor pTARBAC1 posiada sekwencje pochodzące z wektorów YAC (sekwencja centromerowa i telomerowa drożdży CEN3 i TEL3, markery selekcyjne HIS3 i URA3) oraz sekwencje wektora BAC (bakteryjne miejsce rozpoczęcia replikacji czynnika F, łącznik pUC i sekwencje *loxP*) (rys. 9). Wektor jest linearyzowany przed transformacją i wyposażony na końcach w krótkie specyficzne sekwencje tzw. „haki” (*hooks*), homologiczne do regionów w genomie gospodarza. Włączenie fragmentów obcego DNA do wektora poprzedzone jest przez kotransformację sferoplastów drożdżowych fragmentami genomowego DNA i wektorem TAR



Rys. 9. Budowa wektora pTARBAC1 (52).

w formie liniowej. Wewnątrz komórek drożdżowych dochodzi do homologicznej rekombinacji wektora i wprowadzonych fragmentów DNA z utworzeniem kolistych cząsteczek. Jako „haki” można wykorzystać sekwencje często występujące w genomie (elementy repetytywne *Alu*, lub długie elementy repetytywne LINE dla genomu człowieka, B1 – dla genomu myszy). Sekwencja ARS (drożdżowa sekwencja odpowiedzialna za autonomiczną replikację) nie występuje w wektorze. Brak tej sekwencji powoduje, że niezrekombinowane cząsteczki wektora nie podlegają replikacji i są eliminowane z biblioteki w następnych pokoleniach. Sekwencje typu ARS są elementem występującym w dużej liczbie w genomach eukariotycznych (jedna sekwencja typu ARS na 20-40 kbp genomowego DNA eukariota) (43), a zatem wprowadzenie sekwencji genomowego DNA do wektora rekompensuje jej brak i pozwala na replikację wektora w kolejnych pokoleniach. Z uwagi na stabilność insertów, jak i wygodę prowadzenia dalszych eksperymentów, zrekombinowane wektory są następnie izolowane z drożdży i wprowadzane do *E. coli*.

Technika ta pozwala także na ukierunkowane wprowadzenie do wektora sekwencji pochodzących ze zdefiniowanego regionu genomu lub określonych genów (42-49). W tym przypadku jako „haki” stosuje się zdefiniowane sekwencje flankujące poszukiwany region. Na tej zasadzie możliwe jest bezpośrednie klonowanie wszystkich funkcjonalnych kopii danego genu z bardzo nawet złożonych genomów (50,51). Innym podejściem jest wykorzystanie zdegradowanych klonów już istniejącej biblioteki (utworzonych przez trawienie enzymem restrykcyjnym nie rozpoznającym miejsc w obrębie wektora, ale tworzącym „haki” na jego końcach) w celu

izolacji klonów zawierających homologiczny fragment pochodzący z genomu innego organizmu (52).

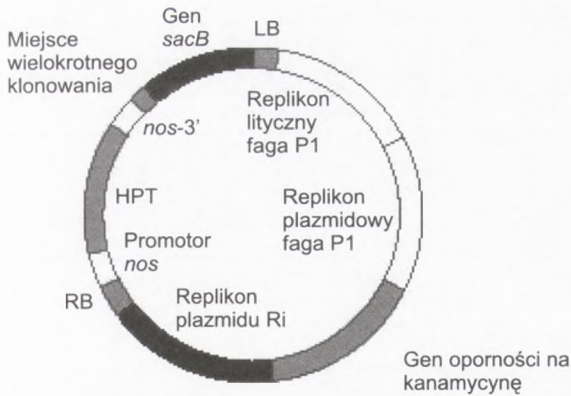
3.9. Wektory do transformacji roślin

Jednym z podstawowych zastosowań bibliotek genomowych jest poszukiwanie i analiza nowych genów. Odbywa się to zazwyczaj przez klonowanie pozycyjne, polegające na oznaczeniu markerów DNA blisko sprzężonych z badanym genem, a następnie izolacji z biblioteki genomowej dużego fragmentu pokrywającego się z markerami. Ostateczne poznanie sekwencji genu odbywa się przy użyciu techniki „spaceru po chromosomie”. Jednak dowodem na prawidłową identyfikację poznanego genu jest dopiero test komplementacji, w którym mutanta pozbawionego aktywności badanego genu transformuje się jego dzikim allelem. Wprowadzenie dużego genu eukariotycznego do komórek bakteryjnych, drożdżowych i zwierzęcych nie stanowi w chwili obecnej większego problemu technicznego (53,54). Dużo większe trudności sprawia wydajna transformacja komórek roślinnych. Skonstruowane systemy opierające się na wektorach kosmidowych (55) i typu lambda (56) pozwalały na konstrukcję bibliotek o wielkości 5-25 kpz. Taka pojemność wektorów ograniczała jednak możliwości izolacji genów przez mapowanie pozycyjne.

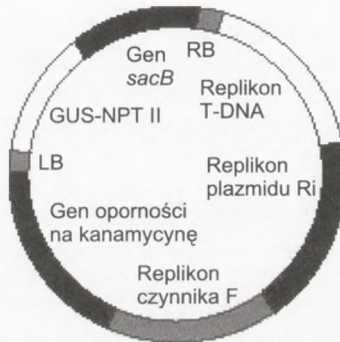
Rozwój nowych typów wektorów charakteryzujących się dużą pojemnością doprowadził do opracowania systemów pozwalających na transformację roślin fragmentami DNA przekraczającymi 100 kpz. Jeden z systemów opiera się na użyciu tzw. sztucznego chromosomu zdolnego do transformacji (TAC, *transformation-competent artificial chromosome*), który może być replikowany zarówno w komórkach *E. coli* jak i w *Agrobacterium tumefaciens*. Wektor TAC został skonstruowany na bazie wektora P1 (zawiera replikony P1, plazmidowy i lityczny oraz gen *sacB* kodujący enzym transferazę sacharylo-6-fruktozową). Posiada także replikon plazmidu Ri - pRiA4, 5' i 3' sekwencję graniczną T-DNA (RB i LB) ze znajdującymi się pomiędzy nimi promotorem syntetazy nopolinowej *Pnos*, region kodujący genu fosfotransferazy higromycynowej HPT i *nos* 3'-sygnał poliadenylacji genu syntetazy nopolinowej (57) (rys. 10).

Drugi system opracowany przez Hamiltona i wsp. to tzw. binarny system BAC (BIBAC) (56). Wykorzystuje on wektory BAC, do których wprowadzono 5' i 3' sekwencję graniczną T-DNA (RB i LB) oraz geny markerowe β -glukuronidazy GUS, fosfotransferazy neomycynowej NPTII i oporności na higromycynę oraz miejsce inicjacji replikacji plazmidu Ri (rys. 11). Dodatkowo w systemie tym stosuje się plazmid pomocniczy pCH zawierający geny wirulencji *A. tumefaciens*: *virG*, *virE1* i *E2*.

Oba systemy znajdują szerokie zastosowanie w pracach nad identyfikacją nowych genów roślinnych (57-64).



Rys. 10. Budowa wektora pYLAC7 (57).



Rys. 11. Budowa wektora BIBAC1 (58).

4. Podsumowanie

Przed podjęciem prac nad konstrukcją biblioteki genomowej podstawowe znaczenie ma dobór odpowiedniego wektora. Przy stale powiększającej się ofercie różnego rodzaju wektorów, ważne jest dobranie wektora do zadań, jakie w przyszłości spełniać ma biblioteka. Przy mapowaniu dużych obszarów genomu, ogólnym mapowaniu fizycznym czy w pracach nad rekonstrukcją pełnej sekwencji genomu preferowane są biblioteki dużych fragmentów. Należy jednak pamiętać, że przy charakteryzacji dużych insertów, zwykle niezbędne staje się ich klonowanie w postaci krótszych fragmentów, a zatem przejście na poziom wektorów plazmidowych. Dlatego w pracach nad pozyskiwaniem nowych markerów EST, STS, SSR czy mapowaniem indywidualnych genów bardziej przydatne są biblioteki mniejszych insertów. Rozpatrzenie zalet i wad różnych systemów klonowania pozwala na dostosowanie biblioteki do potrzeb, a także może usprawnić i obniżyć koszty uzyskania wyniku.

Tabela 2

Podstawowe parametry wektorów stosowanych do konstrukcji bibliotek genomowego DNA

Rodzaj wektora	Wielkość wektora	Pojemność	Gospodarz	Liczba kopii	Stabilność	Zastosowanie	Cechy charakterystyczne	
							wady	zalety
plazmidy	ok. 3 kpz	do 15 kpz	<i>E. coli</i>	kilka do kilku- set	duża	— banki klonów służące do izolacji markerów (np. RFLP) i specyficznych sekwencji — sekwencjonowanie — subklonowanie	— mała pojemność	— wydajne klonowanie, selekcja i izolacja — dodatkowe zastosowania np. nadekspresja
bakteriofagi lambda	2 ramiona: 9 i 19 kpz	10-23 kpz	<i>E. coli</i>	faza — lizogenna — pojedyncze — lityczna — liczne	duża	— biblioteki cDNA — w mniejszym stopniu biblioteki DNA	— mała pojemność	— wydajna transformacja przez infekcję — wstępna selekcja wielkości insertów
kosmidy	7,6 kpz	35-45 kpz	<i>E. coli</i>	od kilku do kilkunastu	niska	— biblioteki genomowe	— niska stabilność insertów — utrudnione przeglądanie biblioteki	— wydajna transformacja przez infekcję — stosunkowo duża pojemność
YAC	11 kpz	700-1800 kpz	<i>S. cerevisiae</i>	pojedyncza	niska	— biblioteki genomowe	— niska stabilność insertów — mało wydajna transformacja i izolacja	— bardzo duża pojemność
bakteriofagi P1	30 kpz	do 100 kpz	<i>E. coli</i>	faza — lizogenna — pojedyncze — lityczna — kilkadziesiąt	wysoka	— biblioteki genomowe	— skomplikowane i mniej wydajne procedury klonowania	— wydajna transformacja — duża stabilność insertów
PAC	19,5 kpz	do 150 kpz	<i>E. coli</i>	kilka	wysoka	— biblioteki genomowe	— duże rozmiary wektora utrudniające przeglądanie biblioteki przez <i>SbaI</i> <i>sequencing</i>	— duża pojemność i stabilność insertów
BAC	7,4 kpz	powyżej 150 kpz	<i>E. coli</i>	1-2	bardzo wysoka	— biblioteki genomowe	pojemność mniejsza od YAC	— duża pojemność i stabilność insertów

Szerokie programy mapowania fizycznego całych genomów przy użyciu bibliotek dużych fragmentów DNA stały się w ostatnim czasie głównym nurtem badań z zakresu genomiki wielu gatunków roślin i zwierząt. Mapy fizyczne tworzą bowiem niezbędną podstawę do sekwencjonowania całych genomów i efektywnego klonowania pozycyjnego genów. Stąd bierze się, przedstawiony w zarysie, szybki postęp w dziedzinie udoskonalania wektorów i procedur operowania bibliotekami genomowymi.

Literatura

1. Clarke L., Carbon J., (1976), *Cell*, 9, 91-99.
2. Elgar G., (1997), *Plant Molecular Biology – A Laboratory Manual*, Ed. M. S. Clark, 94-116, Springer-Verlag, Berlin, Heidelberg.
3. Wing R. A., Rastoloi V., Zhang H., Paterson A., Tanksley S. D., (1993), *Plant J.*, 4, 893-898.
4. Zhang H-B., Zhao X. P., Ding X. D., Paterson A. H., Wing R. A., (1995), *Plant J.*, 7, 175-184.
5. Nizetic D., Zechetner G., Monaco A. P., Gellen L., Young B. D., Lechrach H., (1991), *Proc. Natl. Acad. Sci. USA*, 88, 3233-3237.
6. Bolivar F., Rodrigez R. L., Greene P. J., Betlach M. C., Heyneker H. L., Boyer H. W., (1997), *Gene*, 2, 95-113.
7. Clark M. S., (1997), *Plant Molecular Biology – A Laboratory Manual*, Ed. M. S. Clark, 75-93, Springer-Verlag, Berlin, Heidelberg.
8. Kochert G., Halward T., Branch W. D., Simpson C. E., (1991), *Theor. Appl. Genet.*, 81, 565-570.
9. Nodari R. O., Koinange E. M. K., Kelly J. D., Gepts P., (1992), *Theor. Appl. Genet.*, 84, 186-192.
10. Keim P., Beavis W., Schupp J., Freestone R., (1992), *Theor. Appl. Genet.*, 85, 205-212.
11. Itoh K., Iwabuchi M., Shimamoto K., (1991), *Theor. Appl. Genet.*, 81, 356-362.
12. Hugros G., Monte J. V., Ferrer E., (1990), *Theor. Appl. Genet.*, 80, 24-32.
13. Schweizer G., Borisjuk N., Borisjuk L., Stadler M., Stelzer T., Schilde L., Hemleben V., (1993), *Theor. Appl. Genet.*, 85, 801-808.
14. Winter P., Pfaff T., Udupa S. M., Huttel B., Sharma P. C., Sahi S., Arreguin-Espinoza R., Weigard F., Muehlbauer F. J., Kachl G., (1999), *Mol. Gen. Genet.*, 262, 90-101.
15. Devos K. M., Atkinson M. D., Chinoy C. N., Lui C. J., Gale M. D., (1992), *Theor. Appl. Genet.*, 83, 931-939.
16. Arumuganathan K., Martin G. B., Telenius H., Tanksley S. D., Erle E. D., (1994), *Mol. Gen. Genet.*, 242, 551-558.
17. Vega J. M., Abbo Sh., Feldman M., Levy A. A., (1994), *Proc. Natl. Acad. Sci. USA*, 91, 12041-12045.
18. Macas J., Gualberti G., Nouzova P., Lucretti S., Dolezel J., (1996), *Chromosome Res.*, 4, 531-539.
19. Liu B., Segal G., Vega J. M., Feldman M., Abbo Sh., (1997), *Plant J.*, 11(5), 959-956.
20. Zhou Y., Wang H., Wei J., Cui L., Deng X., Wang X., Chen Z., (2000), *BioTechniques*, 28, 766-774.
21. Kittler R., Stoneking M., Kayser M., (2002), *Anal. Biochem.*, 300, 237-244.
22. Albani D., Cote M-J., Armstrong K. C., Chen Q., Segal A., Robert L. S., (1993), *Plant J.*, 4(5), 899-903.
23. Stein N., Ponielies N., Musket T., McMullen M., Weber G., (1998), *Plant J.*, 13(2), 281-289.
24. Liu C-N., Rubenstein I., (1992), *Mol. Gen. Genet.*, 231, 304-312.
25. Matallana E., Simpson J., Guzman P. A., (1997), *Plant Molecular Biology – A Laboratory Manual*, Ed. M. S. Clark, 131-153, Springer-Verlag, Berlin, Heidelberg.
26. Smith D. R., Smith A. P., Moir D. T., (1990), *Proc. Natl. Acad. Sci. USA*, 87, 8242-8246.
27. Smith D. R., Smith A. P., Strauss W. M., Moir D. T., (1992), *Mamm. Genome*, 4, 141-147.
28. Ecer J. R., (1990), *Methods*, 1, 186-194.
29. Ward E. R., Jen G. C., (1990), *Plant Mol. Biol.*, 14, 561-568.
30. Gril E., Samerville C., (1991), *Mol. Gen. Genet.*, 226, 484-490.

31. Edwards K. J., Thompson H., Edwards D., de Saizieu A., Sparks C., Thompson J. A., Greenland A. J., Evers M., Schuch W., (1992), *Plant Mol. Biol.*, 19, 299-308.
32. Martin G. B., Ganai M. W., Tanksley S. D., (1992), *Mol. Gen. Genet.*, 233, 25-32.
33. Bellann-Chantelot C., Lacroix B., Ougen P., Billault A., Beaufile S., Bertrand S., Georges I., Gilbert F., Gros I., Lucotte G., Susini L., Codani J. J., Gesnouin P., Pook S., Cohen D., (1992), *Cell*, 70, 1059-1068.
34. Larionov V., Kouprina N., Nikolaishvili N., Resnick M. A., (1994), *Nucleic Acids Res.*, 22(20), 4154-4162.
35. Pierce J. C., Sauer B., Sternberg N., (1992), *Proc. Natl. Acad. Sci. USA*, 89(6), 2056-2060.
36. Ioannou P., Amemiya C. T., Garnes J., Kroisel P. M., Shizuya H., Chen C., Batzer M. A., de Jong P. J., (1994), *Nature Genetics*, 6, 84-89.
37. Shizuya H., Birren B., Kim U. J., Mancino V., Slepak T., Tachiiri Y., Simon M., (1992), *Proc. Natl. Acad. Sci. USA*, 89(18), 8794-8797.
38. Woo S. S., Jiang J., Gill B. S., Paterson A. H., Wing R. A., (1994), *Nucleic Acids Res.*, 22(23), 4922-4931.
39. Coyne C. J., Meksem K., Lightfoot D. A., Keller K. E., Martin R. R., McClendon M. T., Inglis D. A., Storline E. W., McPhee K. E., (2000), *Pisum Genetics*, 32, 23-26.
40. Tao Q., Chang Y.-L., Wang J., Chen H., Islam-Faridi M. N., Scheuring Ch., Wang B., Stelly D. M., Zhang H.-B., (2001), *Genetics*, 158, 1711-1724.
41. Osoegawa K., Mammoser A. G., Wu C., Frengen E., Zeng C., Catanese J. J., de Jong P. J., (2001), *Genome Res.*, 11(3), 483-496.
42. Larionov V., Kouprina N., Graves J., Chen X. N., Korenberg J. R., Resnick M. A., (1996), *Proc. Natl. Acad. Sci. USA*, 93(1), 491-496.
43. Stinchomb D. T., Thomas M., Kelly I., Selker E., Davis R. W., (1980), *Proc. Natl. Acad. Sci. USA*, 77, 4559-4563.
44. Larionov V., (1999), *Genet. Eng.*, 21, 37-55.
45. Cancilla M. R., Tainton K. M., Barry A. E., Larionov V., Kouprina N., Resnick M. A., Sart D. D., Choo K. H., (1998), *Genomics*, 47(3), 399-404.
46. Kouprina N., Graves J., Cancilla M. R., Resnick M. A., Larionov V., (1997), *Gene*, 97(1-2), 269-276.
47. Humble M. C., Kouprina N., Noskov V. N., Graves J., Garner E., Tennant R. W., Resnick M. A., Larionov V., Cannon R. E., (2000), *Genomics*, 70(3), 292-299.
48. Noskov V., Kouprina N., Leem S. H., Koriabine M., Barrett J. C., Larionov V., (2002), *Nucleic Acids Res.*, 30(2), E8.
49. Noskov V. N., Koriabine M., Solomon G., Randolph M., Barrett J. C., Leem S. H., Stubbs L., Kouprina N., Larionov V., (2001), *Nucleic Acids Res.*, 29(6), E32.
50. Kouprina N., Annab L., Graves J., Afshari C., Barrett J. C., Resnick M. A., Larionov V., (1998), *Proc. Natl. Acad. Sci. USA*, 95(8), 4469-4474.
51. Annab L. A., Kouprina N., Solomon G., Cable P. L., Hill D. E., Barrett J. C., Larionov V., Afshari C. A., (2000), *Gene*, 250(1-2), 201-208.
52. Zeng C., Kouprina N., Zhu B., Cairo A., Hoek M., Cross G., Osoegawa K., Larionov V., de Jong P., (2001), *Genomics*, 77(1-2), 27-34.
53. Hejna J. A., Johnstone P. L., Kohler S. L., Bruun D. A., Reifsteck C. A., Olson S. B., (1998), *Nucleic Acids Res.*, 26(4), 1124-1125.
54. Wang Z., Engler P., Longacre A., Storb U., (2001), *Genome Res.*, 11(1), 137-142.
55. Olszewski N. E., Martin F. B., Ausubel F. M., (1988), *Nucleic Acids Res.*, 16(22), 10765-10782.
56. Fuse T., Kodama H., Hayashida N., Shinozaki K., Nishimura M., Iba K., (1995), *Plant J.*, 7(5), 849-856.
57. Liu Y. G., Shirano Y., Fukaki H., Yanai Y., Tasaka M., Tabata S., Shibata D., (1999), *Proc. Natl. Acad. Sci. USA*, 96(11), 6535-6540.
58. Hamilton C. M., Frary A., Lewis C., Tanksley S. D., (1996), *Proc. Natl. Acad. Sci. USA*, 93(18), 9975-9979.
59. Hamilton C. M., (1997), *Gene*, 200(1-2), 107-116.
60. Liu Y. G., Nagaki K., Fujita M., Kawaura K., Uozumi M., Ogihara Y., (2000), *Plant J.*, 23(5), 687-695.
61. Liu Y. G., Liu H., Chen L., Qiu W., Zhang Q., Wu H., Yang C., Su J., Wang Z., Tian D., Mei M., (2002), *Gene*, 282(1-2), 247-255.

62. Bentolila S., Hanson M. R., (2001), *Mol. Genet. Genomics*, 266(2), 223-230.
63. Farary A., Hamilton C. M., (2001), *Transgenic Res.*, 10(2), 121-132.
64. Nakamura Y., Kaneko T., Asamizu E., Kato T., Sato S., Tabata S., (2002), *DNA Res.*, 9(2), 63-70.