Wu et al., Clon Transgen 2015, 5:1
DOI: 10.4172/2168-9849.1000146

Cloning & Transgenesis
ISSN: 2168-9849

Review Article Open Access

Achievement, Challenge and Future Perspective of BAC Library Technology with a Focus on its Application to Soybean Genomics Researches

Hongyan Wu[1], Zhengjun Xia[1]* and Kyuya Harada[2]

[1]Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin, China
[2]Graduate School of Engineering,Osaka University,Suita,Japan

Abstract

Large genomic DNA insert-containing libraries e.g. bacterial artificial chromosome (BAC) libraries have made great contributions to advanced genomic studies, such as genome-wide sequencing, syntenic analysis between genomes, physical mapping, and positional cloning of important QTL genes in soybean or other legume. As recent advances in next-generation sequencing (NGS) technology, application of BAC or other large insert-containing library has become less demanding. We briefly review the history and applications of soybean BAC library technology. The challenge and future prospective of soybean BAC library technology in the post genomic-era are also discussed.

Keywords: Bacterial artificial chromosome; Genome; Library; Gene cloning; Soybean

Introduction

In the late 1980s, the discovery of yeast artificial chromosomes (YACs) made cloning of magabase-sized DNA fragments possible [1]. However, several drawbacks for YAC, i.e. almost 50% of YAC clones are chimeric or resulted from insert rearrangements [2-4], lead such clones unsuitable for advanced genomic application, e.g. sequencing and mapping research [3,5]. Also it is rather time consuming to manipulate the YAC clones and deal with chimerase or insert rearrangement [6,7]. Bacterial artificial chromosomes (BACs) library was developed as alternative to YACs [8]. BACs rather are modified bacterial F factors and are not really artificial chromosomes in themselves, which is contrary to their name [9]. BAC had been widely employed as best of choice for advanced genomics research in many plant species including soybean. The advantages of BAC libraries can be summarized as follow:

• BAC clones are almost free of chimerism and insert rearrangements [7,10].

• BAC clone is also very stable since F factor genes (*parA* and *parB*) that prevent more than one BAC from simultaneously inhabiting a bacterium [4,8,11].

• The insert sizes of BAC library are large enough for downstream advanced genomics research e.g. developing physical contig of the region that QTL of interest are located as well as physical map of whole genome. Typically, average insert sizes are with a ranges of 80 and 200 Kb for a good quality BAC library, although some clones can reach 500 kb or more [8,12].

It is also easy to manipulate in terms of preparation, screening and application of BAC libraries [6,13,14]. Most BAC vectors carry an antibiotic resistance gene in order to ensure all acquired colony carrying plasmid. Also polycloning sites within a reporter gene are very suitable for many restriction enzymes as well as colony selection, e.g. the blue/white selection which is based on the principle of α-complementation of the β-galactosidase gene. The most common BAC vector is pBeloBAC11 [9,15].

Application of BAC libraries

Many DNA probes and markers can be used to screen the BAC libraries to localize specific BACs, therefore, we can superimpose genetic maps directly onto BAC-based physical maps [16,17]. Thus map-based cloning of genes responsible for specific phenotypes is greatly facilitated [18,19]. Many important genes were cloned and functionally confirmed [20-22].

High-throughput physical mapping (BAC-based mapping in conjunction with efficient multiplex screening methods) can be applied to the construction of BAC contigs encompassing entire chromosomes and/or complete chromosome sets of a genome [4,17]. Genome sequence information can be achieved through strategy of BAC by BAC i.e. soybean genome [23-25]. BAC-based physical mapping does require DNA polymorphism [26]. Therefore it can replace radiation hybrid mapping in which chromosomes are broken by radiation and propagated in cell cultures [27].

In the post genomic-era, a gap filling is very important and hard task for a perfect genome assembly. BAC libraries, contig assembly, STS-based mapping can facilitate the gap filling in genome assembly using whole-genome shotgun (WGS) sequencing. Fluorescence *in situ* hybridization (FISH) using BACs as probe [28-29] allows molecular and physical maps to be directly superimposed onto the framework of chromosomes, and subsequently the relationship between chromosome structure, DNA sequence, and recombination can be revealed [30].

Application of BAC library technology in soybean genomics research

In 1997, Marek and Shoemaker constructed a soybean BAC 40000 clones (4-5 genome equivalents) in 384-well plate microtiter dishes for studying disease resistance gene [31]. Danesh et al., made approximately 30000 clones with an average insert size of 120 kilobase pairs for their research in cyst nematode resistance [18]. In 1999 Salimath and Bhattacharyya constructed ~ 45 000 clones for cultivar Williams 82, which carries the *Rps1-k* gene to *Phytophthora sojae*

***Corresponding author:** Xia Z, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin 150081, China, Tel: 845187501708, Fax: 845187501708; E-mail: xiazhj@iga.ac.cn

Received November 03, 2015; **Accepted** March 02, 2016; **Published** March 10, 2016

Citation: Wu H, Xia Z, Harada K (2016) Achievement, Challenge and Future Perspective of BAC Library Technology with a Focus on its Application to Soybean Genomics Researches. Clon Transgen 5: 146. doi:10.4172/2168-9849.1000146

Website	URL	Information or resource available	Reference
Soybase	http://soybase.org	BAC based physical map of soybean; a comprehensive repository for professionally curated genetics, genomics and related data resources for soybean	[42]
LIS - Legume Information System	http://legumeinfo.org/	A collaborative, community resource to facilitate crop improvement by integrating genetic, genomic, and trait data across legume species	[43]
The Soybean Genome Database (SoyGD)	http://soybeangenome.siu.edu	(BAC) fingerprint database, genetic map associated genomic data, soybean physical map	[44]
Clemson University	https://www.genome.clemson.edu/services/bacrc/BAC_library	BAC library construction	[45]
Daizubase	http://daizu.dna.affrc.go.jp/	BAC-end sequence of BAC library of a Japanese cultivar Enrei	[46]
The Arizona Genomics Institute (AGI)	http://www.genome.arizona.edu/	Targeted sequencing using BACs and BAC pools; - BAC Library construction - Whole Genome Profile BAC Physical Mapping	[47]
LegumeBase	https://www.legumebase.brc.miyazaki-u.ac.jp/	BAC clones of Lotus japonicus	[48]
Zhang Hongbin's Lab (Texas A&M University)	doi:10.1038/nprot.2011.455	Preparation of megabase-sized DNA from a variety of organisms using the nuclei method	[49]

Table 1: Online resources of BAC library of soybean and related leguminous species.

[32]. Since William 82 was used for genome sequencing, this kind library is very useful for gap-filling and verification of sequencing data. In 1999 Tomkins et al., made a soybean BAC library using the plant introduction (PI) 437654 with 9 haploid genome equivalents for studying soybean cyst nematode (SCN) resistance genes [12]. In 2001 Lewerset et al., built a physical mapping of resistant and susceptible soybean genomes near the soybean cyst nematode resistance gene Rhg_4, which has contribute the successful cloning of this gene [33,34]. Xia et al., constructed 53,760 clones with an average insert size of 116 kb for a Japanese cultivar, Misuzudaizu. This library represents 5.2 genome equivalents [35]. Also three-dimensional pools of the BAC library facilitate rapid and efficient PCR-based screening. This library had facilitated successful cloning of *E1* [22], *E2* [21] and *E3* [20]. In 2004, Wu et al., constructed (BIBAC)-based physical map of the soybean genome based on plant-transformation-competent binary large-insert plasmid clone, which will benefit research in positional cloning of genes and QTLs. The BIBAC library contained 78,001 clones representing 9.6 haploid genomes [36,37]. Together with previous build BIBAC library for soybean cv. Forrest, two (BAC and BIBAC) libraries encompass 13.2 haploid genomes, providing the comprehensive clone resource for a single soybean genotype.

Integrated genetic and physical maps are extremely valuable for genomic studies as well as for guiding assembling whole genome shotgun sequences. Comparative physical mapping can reveal features of microsynteny between *Glycine max*, *Medicago truncatula*, and *Arabidopsis thaliana*. In 2013, Yan et al., used BAC-end cross-hybridization among *G. max* and *M. truncatula contings and uncovered* microsynteny [38]. The result showed 85% of coding and 75% of noncoding sequences between *G. max* homoeologous (within genome duplicate) contigs were conserved in cross-hybridization, which confirmed one recent genome-wide duplication in *G. max* and suggested that *M. truncatula* also experienced ancient large-scale genome duplications.

Wu tried to use genetic marker anchoring by six-dimensional pools for development of a soybean physical map [39]. Six-dimensional soybean BAC pools can be efficiently used to anchor markers to soybean BACs despite the complexity of the soybean genome. In addition to anchoring markers, the 6-D pooling method was also effective for targeting BAC clones for investigating gene families and duplicated regions in the genome, as well as for extending physical map contigs. The usefulness of the BAC library has been constrained since construction, storage and screening are very difficult and laborious. Xia developed a simple and novel hand-made collection device for pooling without the needs for robotic manipulation [40]. In 2014, Xia et al. also developed a non-gridded BAC library for saving storage space as well as procedure for retrieving the target BAC clones from the positive pools [41]. Currently, here we list the websites regarding protocol, resource and application of BAC library of soybean or other leguminous species in Table 1.

Challenge and future perspectives of BAC library

Since the emergence of the next generations of sequencing technology, BAC library technology has faced strong challenge since construction, screening and application of BAC libraries are time and labor intensive [42-49]. Recent third-generation sequencing technology, which can yield longer multi kilo base reads and is getting popular in various genomic researches, further despoils the the applications of BAC library technology. However, the special uniqueness of BAC library still makes a BAC library technology of first choice in modern genomic researches, e.g. refinement of physical map of target region of interest or at the whole genome level, comparative analyses between different species, and preservation of the genetic resources.

Funding

This work was supported by the Programs (31471518, 31271742, and 31301338) from National Natural Science of China; by Program (2011BAD35B06-2) from National Science and Technology Ministry of China.

References

1. Burke DT, Carle G, Olsen MV (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236: 806-812.
2. Burke DT (1990) YAC cloning: options and problems. GATA 7: 94-99.
3. Venter JC, Smith HO, Hood L (1996) A new strategy for genome sequencing. Nature 381: 364-366.
4. Cai WW, Reneker J, Chow CW, Vaishnav M, Bradley A (1998) An anchored framework BAC map of mouse chromosome 11 assembled using multiplex oligonucleotide hybridization. Genomics 54: 387-397.
5. Green ED, Riethman HC, Dutchik JE, Olson MV (1991) Detection and characterization of chimeric yeast artificial-chromosome clones. Genomics 11: 658-669.

6. O'Connor M, Peifer M, Bender W (1989) Construction of large DNA segments in Escherichia coli. Science 244: 1307-1312.

7. Woo SS, Jiang J, Gill BS, Paterson AH, Wing RA (1994) Construction and characterization of a bacterial artificial chromosome library of Sorghum bicolor. Nucleic Acids Res 22: 4922-4931.

8. Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, et al. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. P Natl Acad Sci USA 89: 8794-8797.

9. Peterson DG, Tomkins JP, Frisch DA, Wing RA, Paterson AH (2000) Construction of plant bacterial artificial chromosome (BAC) libraries: An illustrated guide. J Agri Genomics.

10. Venter JC, Adams MD, Sutton GG, Kerlavage AR, Smith RO, et al. (1998) Shotgun sequencing of the human genome. Science 280: 1540-1542

11. Willetts N, Skurray R (1987) Structure and function of the F factor and mechanism of conjugation. American Society for Microbiology, Washington DC 2: 1110-1133.

12. Tomkins JP, Mahalingam R, Smith H, Goicoechea JL, Knap HT, et al. (1999) A bacterial artificial chromosome library for soybean PI 437654 and identification of clones associated with cyst nematode resistance. Plant Mol Bio 41: 25-32.

13. Paterson AH (1996) The DNA revolution. In Paterson AH (ed.), Genome Mapping in Plants. Academic Press, San Diego, CA, pp: 7-21.

14. Marra MA, Kucaba TA, Dietrich NJ, Green ED, Brownstein B, et al. (1997) High throughput fingerprint analysis of large-insert clones. Genome Res 7: 1072-1084.

15. Choi S, Wing RA (1999) The Construction of bacterial artificial chromosome (BAC) libraries. In: Gelvin S and Schilperoort R (eds.) Plant Molecular Biology Manual (2ndedn). Suppl. IV. Kluwer Academic Publishers, The Netherlands pp: 1-32.

16. Yang D, Parco A, Nandi S, Subudhi P, Zhu Y, et al. (1997) Construction of a bacterial artificial chromosome (BAC) library and identification of overlapping BAC clones with chromosome 4-specific RFLP markers in rice. Theor Appl Genet 95: 1147-1154.

17. Mozo T, Dewar K, Dunn P, Ecker JR, Fischer S, et al. (1999) A complete BAC-based physical map of the Arabidopsis thaliana genome. Nat Genet 22: 271-275.

18. Danesh D, Penuela S, Mudge J, Denny RL, Nordstrom H, et al. (1998) A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. Theor Appl Genet 96: 196-202.

19. Sanchez AC, Ilag LL, Yang D, Brar DS, Ausubel F, et al. (1999) Genetic and physical mapping of xa13, a recessive bacterial blight resistance gene in rice. Theor Appl Genet 98: 1022-1028.

20. Watanabe S, Hideshima R, Xia ZJ, Tsubokura Y, Sato S, et al. (2009) Map-based cloning of the gene associated with the soybean maturity locus E3.Genetics 182: 1251-1262.

21. Watanabe S, Xia ZJ, Hideshima R, Tsubokura Y, Sato S, et al. (2011) A map-based cloning strategy employing a residual heterozygous line reveals that the GIGANTEA gene is involved in soybean maturity and flowering. Genetics 188: 395-407.

22. Xia ZJ, Watanabe S, Yamada T, Tsubokura Y, Nakashima H, et al. (2012) Positional cloning and characterization reveal the molecular basis for soybean maturity locus E1 that regulates photoperiodic flowering. P Natl Acad Sci USA 109: E2155-E2164.

23. Zhang HB, Wing RA (1997) Physical mapping of the rice genome with BACs. Plant Mol Biol 35: 115-127.

24. Sasaki T, Burr B (2000) International Rice Genome Sequencing Project: The effort to completely sequence the rice genome. Curr Opin Plant Biol 3: 138-142.

25. International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436: 793-800.

26. Lin YR, Draye X, Qian X, Ren S, Zhu LH, et al. (2000) Locus-specific contig assembly in highly-duplicated genomes using the BAC-RF method. Nucleic Acids Res 28: e23.

27. Deloukas P, Schuler GD, Gyapay G, Beasley EM, Soderlund C, et al. (1998) A physical map of 30,000 human genes. Science 282: 744-746.

28. Cai L, Taylor JF, Wing RA, Gallagher DS, Woo SS, et al. (1995) Construction and characterization of a bovine bacterial artificial chromosome library. Genomics 29: 413-425.

29. Jackson SA, Dong F, Jiang J (1999) Digital mapping of bacterial artificial chromosomes by fluorescence *in situ* hybridization. Plant J 17: 581-587.

30. Peterson DG, Lapitan NLV, Stack SM (1999) Localization of single- and low-copy sequences on tomato synaptonemal complex spreads using fluorescence *in situ* hybridization (FISH). Genetics 152: 427-439.

31. Marek LF, Shoemaker RC (1997) BAC contig development by fingerprint analysis in soybean. Genome 40: 420-427.

32. Salimath SS, Bhattacharyya MK (1999) Generation of a soybean BAC library, and identification of DNA sequences tightly linked to the Rps1-k disease resistance gene. Theor Appl Genet 98: 712-720.

33. Lewers KS, Nilmalgoda SD, Warner AL, Knap HT, Matthews BF (2001) Physical mapping of resistant and susceptible soybean genomes near the soybean cyst nematode resistance gene Rhg_4. Genome 44: 1057-1064.

34. Liu SM, Kandoth PK, Warren SD, Yeckel G, Heinz R, et al. (2012) A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. Nature 492: 256-260.

35. Xia ZJ, Sato H, Watanabe S, Kawasaki S, Harada K (2005) Construction and characterization of a BAC library of soybean. Euphytica 141: 129-137.

36. Wu CC, Nimmakayala P, Santos FA, Springman R, Scheuring C, et al. (2004) Construction and characterization of a soybean bacterial artificial chromosome library and use of multiple complementary libraries for genome physical mapping. Theor Appl Genet 109: 1041-1050.

37. Wu CC, Sun SK, Nimmakayala P, Santos FA, Meksem K, et al. (2004) A BAC and BIBAC-based physical map of the soybean genome. Genome Res 14: 319-326.

38. Yan HH, Mudge J, Kim DJ, Larsen D, Shoemaker RC, et al. (2003) Estimates of conserved microsynteny among the genomes of Glycine max, *Medicago truncatula* and Arabidopsis thaliana. Theor Appl Genet 106: 1256-1265.

39. Wu XL, Zhong GH, Findley SD, Cregan P, Stacey G, et al. (2008) Genetic marker anchoring by six-dimensional pools for development of a soybean physical map. BMC Genomics 9: 28.

40. Xia ZJ, Watanabe S, Chen QS, Sato SS, Harada K (2009) A novel manual pooling system for preparing three-dimensional pools of a deep coverage soybean bacterial artificial chromosome library. Mol Eco Resour 9: 516-524.

41. Xia ZJ, Wu HY, Watanabe S, Harada K (2014) Construction and targeted retrieval of specific clone from a non-gridded soybean bacterial artificial chromosome library. Anal Biochem 444: 38-40.

42. Grant D, Nelson RT, Cannon SB, Shoemaker RC (2009) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucl Acids Res 38: D843-D846.

43. Dash S, Campbell JD, Cannon EK, Cleary AM, Huang W, et al. (2015) Legume information system (LegumeInfo. org): a key component of a set of federated data resources for the legume family. Nucl Acids Res 44: D1181-D1188.

44. Shultz JL, Kurunam D, Shopinski K, Iqbal MJ, Kazi S, et al. (2006) The Soybean Genome Database (SoyGD): a browser for display of duplicated, polyploid, regions and sequence tagged sites on the integrated physical and genetic maps of Glycine max. Nucl. Acids Res 34: D758-D765.

45. Zhu S, Saski CA, Boerma HR, Tomkins JP, All JN, et al. (2009) Construction of a BAC Library for a defoliating insect-resistant soybean and identification of candidate clones using a novel approach. Plant Mol Bio Rep 27: 229-235.

46. Ito K, Katayose Y, Kanamori H, Shimomura M, Ohyanagi H, et al (2012) DaizuBase, an integrated soybean genome database including BAC-based physical maps. Breed Sci 61: 661-664.

47. Ammiraju JSS, Song X, Luo M, Sisneros N, Angelova A, et al. (2010) The Oryza BAC Resource: A genus-wide and genome scale tool for exploring rice genome evolution and leveraging useful genetic diversity from wild relatives. Breed Sci 60: 536-543.

48. Hashiguchi M, Abe J, Aoki T, Anai T, Suzuki A, et al. (2012) The National BioResource Project (NBRP) Lotus and Glycine in Japan. Breed Sci 61: 453-461.

49. Zhang M, Zhang Y, Scheuring CF, Wu CC, Dong JJ, et al. (2012) Preparation of megabase-sized DNA from a variety of organisms using the nuclei method for advanced genomics research. Nat Protoc 7: 467-478.