An improved method to identify BAC clones using pooled overgos

Kavitha Madishetty, Pascal Condamine, Jan T. Svensson, Edmundo Rodriguez and Timothy J. Close*

Department of Botany and Plant Sciences, University of California, Riverside, CA 92521-0124, USA

ABSTRACT

Hybridization using overgo probes is an established approach for screening arrayed bacterial artificial chromosome (BAC) libraries. We have improved the use of overgos by increasing the yield of positive clones using reduced levels of radioisotopes and enzyme. The strategy involves labeling with all four radiolabeled nucleotides in a hot pulse followed by a cold nucleotide chase and then extending the exposure time to compensate for reduced specific activity of the probes. The resulting cost savings and reduced human exposure to radiation make the use of highly pooled overgo probes a more attractive approach for screening of BAC libraries from organisms with large genomes.

INTRODUCTION

Many economically important plants and animals have a genome size that is beyond the current limits of whole-genome sequencing due to cost considerations and difficulties in the assembly of repetitive sequences. For example, hexaploid bread wheat (*Triticum aestivum*) has one of the larger genomes at 17 000 Mb, with diploid barley (*Hordeum vulgare*) at 5300 Mb (1), the genome of each species having at least 80% repetitive DNA (2). Strategies that focus on targeted sequencing of gene-rich regions provide an alternative to whole-genome sequencing. Identification of gene-rich sequences in species with large genomes (3) usually is accomplished by methylation filtration (MF), high Cot selection, hybridization-based approaches (4,5) or PCR. MF and high Cot methods have been applied to maize (6,7) and sorghum (8,9). Hybridization-based screening is performed against gridded macroarrays or combinatorial pools of usually bacterial artificial chromosome (BAC) libraries using radioactively labeled probes either singly or as groups. Several types of probes are used for screening BAC libraries by hybridization (10); the most commonly used are sub-cloned DNA fragments, PCR amplified products (11) or DNA oligonucleotides (12). Probes can be individually radiolabeled, then pooled and hybridized

as a group. Some drawbacks to probe pooling are non-uniform radiolabeling of the probes and the risk of repetitive elements that may confound the results by hybridization to sequences that are abundant in the genome. The use of oligonucleotides can minimize these issues by selecting oligos after considering an entire sequence database (13).

The 'overgo' method of oligonucleotide probes begins by annealing two 24-base or 22-base oligonucleotides with an 8 bp overlapping region and filling in the overhanging bases with Klenow enzyme and radiolabeled nucleotides (14). The resulting probes can be multiplexed and designed to have thermodynamic behavior more uniform than other types of DNA probes. Other advantages of overgos include low background hybridization and simplified tracking because oligos can be synthesized in an orderly and uniform manner. Overgos have been used to screen libraries for physical maps in human (15), mouse (16), rice (17), chicken (18), maize (19), sorghum (20) and peanut (21). Overgos have been used also to screen PAC, fosmid and cosmid libraries, close gaps between sequenced BAC clones in rice (22), anchor BAC contigs (17) and screen rice fosmid libraries to validate genome coverage (23). In animal systems, overgos have been used to identify gene families such as immunoglobulin H chain constant (IGHC) genes in horse (24), for multispecies comparative sequence analysis (25), and as universal probes for the isolation of orthologous regions from multiple vertebrates (26), especially mammalian or avian/reptilian libraries. Overgos have been designed from various sources including expressed sequence tag sequences, RFLP markers, STS markers, BAC-end sequences and telomeric-specific regions (27). The use of overgos as gene-specific hybridization probes in combination with efficient pooling strategies has proven to be a powerful approach in anchoring large numbers of cDNA unigenes to BAC contig maps. Examples are 9371 overgos in maize (19), 56 overgos designed from the RGA sequences of peanut (21) and 648 overgos designed from mapped STS markers in chicken (28). Also, 7573 and 2297 overgos in sorghum and rice, respectively, were used to explore the micro-synteny patterns between the two species (20).

The original overgo protocol described by Ross *et al*. (14) has practical limitations when extrapolated to a large number

*To whom correspondence should be addressed. Tel: +1 951 827 3318; Fax: +1 951 827 4437; E-mail: timothy.close@ucr.edu

of overgos probed in a single hybridization; it becomes very costly and requires a problematic level of human exposure to radioisotopes when used for screening a large genome such as barley. In our effort to identify the majority of gene-rich BAC clones from the barley genome, we modified the pooled overgo strategy to achieve a higher efficiency of clone identification. Here we present the details of these modifications.

MATERIALS AND METHODS

Barley BAC library

A barley BAC library was constructed by Yu *et al.* (29) for cultivar Morex using the restriction endonuclease HindIII. This library contains 313 344 clones (816 384-well plates) and provides 6.3 haploid genome equivalents. The average insert size is 106 kb. The library is arrayed on 17 DNA filters with 18 432 clones represented per filter and is available from Clemson University Genomics Institute (http://www.genome.clemson.edu), which was our source.

Overgo probe design

Ten overgos related to abiotic stress were designed using Overgo Maker (http://genome.wustl.edu/tools/software/overgo.cgi). These sequences are given in Supplementary Table 1. Source sequences for overgo design were unigene consensus sequences, derived from ESTs, in HarvEST:Barley (http://www.harvest.ucr.edu). Each overgo was composed of a pair of oligonucleotides whose ends were complementary by eight bases.

Overgo primer pairs

Oligonucleotides were purchased from Sigma-Aldrich (St Louis, MO). Each oligonucleotide was synthesized as a 24mer at a 25 nmol scale and dissolved in 250 µl of TE buffer. Oligonucleotide pairs were made by combining equal amounts of each 24mer and diluting in TE to a final concentration of 1 µM of each oligonucleotide. These overgo primer pairs were stored at −20°C.

Overgo probe labeling, hybridization, autoradiography

The optimized method was adapted from a protocol obtained from John McPherson (Washington University, St Louis, MO). Ten microliter of each probe pair mix was labeled with 10 µl of freshly prepared master mix composed of 4.0 µl of 5× overgo labeling buffer (OLB; adds 5 pmol of each of four cold nucleotides, see Supplementary Data for complete composition of OLB), 1.0 µl of 2 mg/ml acetylated BSA (Promega, Madison, WI), 0.125 µl of each of four [α-^{32}P]dNTPs (10 µCi/µl, ~3000 Ci/mmol, Perkin Elmer, Wellesley, MA), 3.8 µl of distilled water and 0.2 µl of 2 U/µl Klenow enzyme (New England Biolabs, Ipswich, MA). An overgo primer pair with sequences 5′-AACGGGC-GAGTGATGTAAAATA-3′ and 5′-TGATGGGATCGGGC-TATTTTAC-3′ matching the *Escherichia coli* genome was used in each pool as a background probe to lightly mark all clones for alignment of the image (15). Labeling reactions were carried out in a humidified acrylic box at room temperature for 1 h followed by addition of 5 µl of the a cold chase

solution (cold dNTPs) (15). Later, all the reactions were pooled and probes were denatured at 95°C for 5 min and immediately transferred to hybridization tubes containing prehybridized BAC membranes. Hybridization was performed in 40 ml of Church's buffer (14) at 60°C overnight in a hybridization oven. After hybridization, membranes were washed in solutions with increasing stringency starting with 2 l of 4× SSC, 0.1% SDS followed by 2 l of 1.5× SSC, 0.1% SDS and finally 2 l of 0.75× SSC, 0.1% SDS at 50°C. Membranes were then sealed in plastic wrap and exposed to film (Kodak BIOMAX MS Double Emulsion, 24 cm × 30 cm) with an intensifying screen (BioPlus Screens 10 inch × 12 inch, bioWorld, Dublin, OH) over the film at −80°C for 5–6 days. During exposure at −80°C, lead sheets (0.015 inch thickness) were placed between the cassettes in order to shield the radiation across the filters. The cassettes were clamped between 3/4 inch plywood frames. For more details refer to the Supplementary Data.

RESULTS AND DISCUSSION

Our motivations to optimize the overgo probe method were to minimize the cost of and human exposure to radiation when probing 10 000 or more overgo probes in batches of 192 at a time, and to maximize the discovery rate of gene-positive BAC clones per overgo. The original procedure of Ross *et al.* (14) labeled the overgo primer pairs (20 pmol each) with two radioactive nucleotides [α-^{32}P]dATP and [α-^{32}P]dCTP (10 µCi/µl, 1 µl) in the presence of dGTP and dTTP and 2 U of Klenow enzyme. When these quantities were scaled up to our target of 192 probes the amounts of radioisotopes and Klenow were impractical. Screening the 17-filter barley BAC library with 192 pooled overgos using the standard overgo labeling method of Ross *et al.* (14) would require the handling of nearly four mCi of radioactivity and 400 U of Klenow enzyme for each probing. In order to evaluate the sensitivity and efficiency of various permutations of the labeling and hybridization protocol, we used 10 overgo primer pairs. The variables that we tested included cold chase (15) versus no cold chase, four versus two radionucleotides, hybridization time, presence versus absence of an *E.coli* genome overgo, and concentrations of Klenow enzyme and oligonucleotides.

As noted by Han *et al.* (15) the concentration of each radio-isotope in the standard labeling reaction is only 0.165 µM, which is insufficient to support complete filling of overgos. This was a significant issue for us since any reduction of radionucleotide concentration, as we desired for cost reductions, would further impede the completion of overgo filling. To generate complete overgos, Han *et al.* (15) implemented a cold nucleotide chase. We investigated the effect of the cold chase on the hybridization results using the otherwise standard 2 nt protocol and observed an increase in signal intensity (data not shown). Once this observation was made, the cold chase became an essential part of further modifications. In addition, we noted that 2 U of Klenow enzyme in the standard overgo labeling reaction is in excess. One unit is defined as the amount of enzyme required to catalyze the incorporation of 10 nmol of dNTP in 30 min at 37°C. Complete fill-in of 32 nt (16 per strand) on 0.01 nmol of substrate (10 µl of

1 µM concentration) would be only 3.2% of this load. Even if the enzyme activity at room temperature is only ~1/3 of the activity at 37°C then 2 U still must be an excess. We reduced the amount of enzyme 5-fold (from 2 to 0.4 U per labeling reaction) and observed no effect on the hybridization results (data not shown), so this reduction in Klenow enzyme also became a standard part of our protocol.

We compared hybridization signals detected by labeling with 0.5 µl of each all 4 nt ([α-^{32}P]dATP, [α-^{32}P]dCTP, [α-^{32}P]dGTP, [α-^{32}P]dTTP) keeping the amount of radioactivity the same (20 µCi) as in the standard two-radioisotope protocol, hybridizing to eight library filters containing 147 456 BAC clones. The number of clones detected was much higher in the four-radioisotope procedure (Table 1). There were 8.4 positive BAC clones per filter identified using the four-isotope mixture as compared to 5.8 BAC clones per filter detected by the same set of 10 overgos under the two-isotope labeling procedure, a 45.6% increase. We then examined the consequences of lower amounts of radioactivity in the labeling reactions including 10, 5 and 2 µCi (Table 1). With 10 µCi the two-isotope reaction yielded 4.4 BACs per filter whereas the four-isotope reaction yielded 5.4 clones per filter, a 23% higher yield. With 5 µCi the two-isotope reaction (Figure 1) yielded 6.3 BACs per filter whereas the four-isotope reaction yielded 7.5 clones per filter, a 19% higher yield. With 2 µCi the results of the two-isotope and four-isotope reactions were nearly the same as each other and the required exposure times were too lengthy to be practical. Therefore, reduction to 2 µCi was judged to be unsatisfactory. In addition, higher numbers of positive clones were scored when four radionucleotides were used in the labeling reactions as compared to two radionucleotides in other hybridizations that used lower concentrations of oligos (2 and 5 pmol), but the reduction of oligo concentration to these levels reduced the total number of positive clones and

necessitated longer exposure times (data not shown) and was therefore also judged to be unsatisfactory.

The method that seemed most optimal for our purposes uses one-fourth of the radioactivity (reduced from 20 to 5 µCi per 20 µl reaction; 0.125 µl of each radioactive dNTP at 10 µCi/µl), half the primer concentration (reduced from 20 to 10 pmol) and one-fifth of the Klenow enzyme (reduced from 2 to 0.4 U per reaction) (Table 2) relative to the standard method. The other modifications include a change in the composition of OLB such that it contains a small amount of all four cold dNTPs for the initial labeling reaction (see Materials and Methods, and Supplementary Data), addition of a 1 h cold chase with 0.2 mM of each dNTP to extend the overgos to their full length, an increase in the hybridization time from overnight to 36 h and a longer exposure time from 2–3 to 5–6 days. The increase in hybridization time from overnight to 36 h and longer exposure time of 5–6 days were incorporated only for convenience of the weekly workflow in a large-scale screening project. We did not observe any significant effect of the longer hybridization time. The contrast between positive clones and background was satisfactory with either exposure time. These optimizations all were in the context of using X-ray films for autoradiography, so the use of phosphorimagers, chemiluminescence or other modes of detection may require some other considerations. The detailed modified protocol for pooled overgo hybridization is provided as Supplementary Data.

The rationale behind using all four radionucleotides was that this would help to equalize the intensity of all probes by reducing the influence of nucleotide frequency. This was expected to have two favorable consequences. One was that our overgo design algorithm would not have to take into consideration the frequency and exact position of any particular nucleotide in order to generate an acceptable 36mer, with the consequence being more likelihood that we could choose oligos with comparable thermodynamic behaviour. The other was that the intensity of hybridization among overgos would tend be more equal than would be possible with fewer than four radionucleotide, thus minimizing the number of probes that fail simply due to a relative paucity of label. BACs hybridized with such overgos might then not be observed as positives. We recently applied this modified method to the identification of gene-bearing BAC clones of barley using ~12 600 overgos. In ~70 pools ranging from 96 to 300 simultaneous probes, most often 192 probes per pool, ~65 000 gene-rich BACs were identified. On average 6.7 clones were identified for all the overgos hybridized,

Table 1. Results of four versus two radionucleotide labeling reactions[a]

Radioactivity per reaction (µCi)	# Clones/filter 4 [α-^{32}P]dNTPs	# Clones/filter 2 [α-^{32}P]dNTPs	# Filters[b]	% Increase
20	8.4	5.8	8	45.6
10	5.4	4.4	5	22.7
5	7.5	6.3	4	20.0
2	6.5	7.0	2	−7

[a]The expected number of positives based on library depth would be 3.7 per filter for single-gene-specific probes; not all probes were specific to a single gene.
[b]Different filters were used for each comparison.
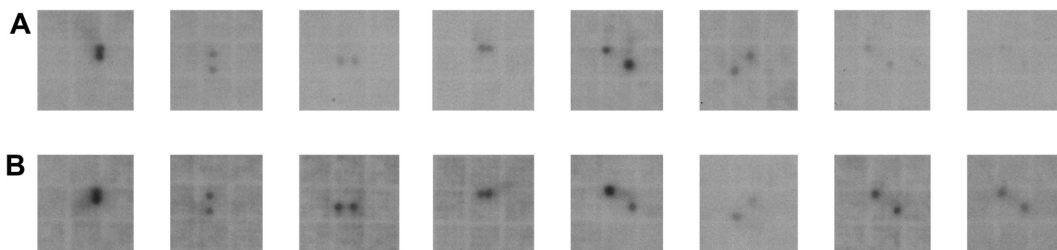
Figure 1. Representative positive clones from autoradiographs of the Morex barley BAC library. (**A**) Two radioactive nucleotides (10 µCi each, exposure time 2 days). (**B**) Same clones with four radioactive nucleotides (1.25 µCi each, 5 µCi total, exposure time 2 days).

Table 2. Differences between the original and modified protocol

Reagent	Original	Modified	Costs for 192 overgos	
			Original	Modified
Primer	20 pmol	10 pmol		
BSA (2 mg/ml) (acetylated)	1.0 µl	1.0 µl	25.00	25.00
5× overgo labeling buffer	4.0 µl	4.0 µl	20.00	20.00
H_2O (add before dNTPs)	2.0 µl	4.0 µl	—	—
[α-^{32}P]dATP (10 µCi/µl)	1.0 µl	0.125 µl	440.00	55.00
[α-^{32}P]dCTP (10 µCi/µl)	1.0 µl	0.125 µl	440.00	55.00
[α-^{32}P]dGTP (10 µCi/µl)	0 µl	0.125 µl	0	55.00
[α-^{32}P]dTTP (10 µCi/µl)	0 µl	0.125 µl	0	55.00
Klenow fragment (2 U/µl)	1.0 µl	0.2 µl	106.00	21.00
Total costs			1031.00	286.00
Exposure time (days)			1–2	5–6

which is in close accordance with the 6.3× genome coverage of this library.

With many plant and animal species in early phases of genome sequencing and physical mapping, the use of methods that are cost-efficient, sensitive and safe are desirable. The procedure described here improves the cost-effectiveness and reduces the radiation exposure hazards of detecting BAC clones en masse using overgo probes. In essence we have made an already-reliable method even better.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Conflict of interest statement. None declared.

REFERENCES

1. Bennett,M.D. and Leitch,I.J. (1995) Nuclear DNA amounts in angiosperms. *Ann. Bot.*, **76**, 113–176.
2. Li,W., Zhang,P., Feller,J.P., Friebe,D. and Gill,B.S. (2004) Sequence composition, organization and evolution of the core Triticeae genome. *Plant J.*, **40**, 500–511.
3. Mayer,K. and Mewes,H.-W. (2002) How can we deliver the large plant genomes? Strategies and perspectives. *Curr. Opin. Plant Biol.*, **5**, 173–177.
4. Gill,B.S., Appels,R., Botha-Oberholster,A.-M., Buell,C.R., Bennetzen,J.L., Chalhoub,B., Chumley,F., Dvorak,J., Iwanaga,M., Keller,B. *et al.* (2004) A workshop report on wheat genome sequencing: international genome research on wheat consortium. *Genetics*, **168**, 1087–1096.
5. Barbazuk,W.B., Bedell,J.A. and Rabinowicz,P.D (2005) Reduced representation sequencing: a success in maize and a promise for other plant genomes. *Bioessays*, **27**, 839–848.
6. Whitelaw,C.A., Barbazuk,W.B., Pertea,G., Chan,A.P., Cheung,F., Lee,Y., Zheng,L., van Heeringen,S., Karamycheva,S., Bennetzen,J.L. *et al.* (2003) Enrichment of gene-coding sequences in maize by genome filtration. *Science*, **302**, 2118–2120.
7. Emberton,J., Ma,J., Yuan,Y., SanMiguel,P. and Bennetzen,J.L. (2005) Gene enrichment in maize with hypomethylated partial restriction (HMPR) libraries. *Genome Res.*, **15**, 1441–1446.
8. Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D., Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K. *et al.* (2005) Sorghum genome sequencing by methylation filtration. *PLoS Biol.*, **3**, e13.
9. Peterson,D.G., Schulze,S.R., Sciara,E.B., Lee,S.A., Bowers,J.E., Nagel,A., Jiang,N., Tibbitts,D.C., Wessler,S.R. and Paterson,A.H. (2002) Integration of cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery. *Genome Res.*, **12**, 795–807.
10. Isidore,E., Scherrer,B., Bellec,A., Budin,K., Faivre-Rampant,P., Waugh,R., Keller,B., Caboche,M., Feuillet,C. and Chalhoub,B. (2005) Direct targeting and rapid isolation of BAC clones spanning a defined chromosome region. *Funct. Integr. Genomics*, **5**, 97–103.
11. Xu,J., Yang,D., Domingo,J., Ni,J. and Huang,N. (1998) Screening for overlapping bacterial artificial chromosome clones by PCR analysis with an arbitrary primer. *PNAS*, **95**, 5661–5666.
12. Klein,P.E., Klein,R.R., Cartinhour,S.W., Ulanch,P.E., Dong,J., Obert,J.A., Morishige,D.T., Schlueter,S.D., Childs,K.L., Ale,M. *et al.* (2000) A high-throughput aflp-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. *Genome Res.*, **10**, 789–807.
13. Zheng,J., Svensson,J., Madishetty,K., Close,T., Jiang,T. and Lonardi,S. (2006) OligoSpawn: a software tool for the design of overgo probes from large unigene datasets. *BMC Bioinformatics*, **7**, 7.
14. Ross,M., LaBrie,T., McPherson,S. and Stanton,V.P. (1999) Screening large-insert libraries by hybridization. In *Current Protocols in Human Genetics*. John Wiley Sons, New York, pp 5.6.1–5.6.32.
15. Han,C.S., Sutherland,R.D., Jewett,P.B., Campbell,M.L., Meincke,L.J., Tesmer,J.G., Mundt,M.O., Fawcett,J.J., Kim,U.-J., Deaven,L.L. *et al.* (2000) Construction of a BAC contig map of chromosome 16q by two-dimensional overgo hybridization. *Genome Res.*, **10**, 714–721.
16. Cai,W.-W., Reneker,J., Chow,C.-W., Vaishnav,M. and Bradley,A. (1998) An anchored framework bac map of mouse chromosome 11 assembled using multiplex oligonucleotide hybridization. *Genomics*, **54**, 387–397.
17. Chen,M., Presting,G., Barbazuk,W.B., Goicoechea,J.L., Blackmon,B., Fang,G., Kim,H., Frisch,D., Yu,Y., Sun,S. *et al.* (2002) An integrated physical and genetic map of the rice genome. *Plant Cell*, **14**, 537–545.
18. Ren,C., Lee,M.-K., Yan,B., Ding,K., Cox,B., Romanov,M.N., Price,J.A., Dodgson,J.B. and Zhang,H.-B. (2003) A BAC-based physical map of the chicken genome. *Genome Res.*, **13**, 2754–2758.
19. Gardiner,J., Schroeder,S., Polacco,M.L., Sanchez-Villeda,H., Fang,Z., Morgante,M., Landewe,T., Fengler,K., Useche,F., Hanafey,M. *et al.* (2004) Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization. *Plant Physiol.*, **134**, 1317–1326.
20. Bowers,J.E., Arias,M.A., Asher,R., Avise,J.A., Ball,R.T., Brewer,G.A., Buss,R.W., Chen,A.H., Edwards,T.M., Estill,J.C. *et al.* (2005) Comparative physical mapping links conservation of microsynteny to chromosome structure and recombination in grasses. *Proc. Natl. Acad. Sci. USA*, **102**, 13206–13211.
21. Yuksel,B., Estill,J., Schulze,S. and Paterson,A. (2005) Organization and evolution of resistance gene analogs in peanut. *Mol. Genet. Genomics*, **274**, 248–263.
22. Yang,T.-J., Yu,Y., Nah,G., Atkins,M., Lee,S., Frisch,D.A. and Wing,R.A. (2003) Construction and utility of 10 kb libraries for efficient clone-gap closure for rice genome sequencing. *Theor. Appl. Genet.*, **107**, 652–660.
23. Ammiraju,J., Yu,Y., Luo,M., Kudrna,D., Kim,H., Goicoechea,J., Katayose,Y., Matsumoto,T., Wu,J., Sasaki,T. *et al.* (2005) Random sheared fosmid library as a new genomic tool to accelerate complete finishing of rice (*Oryza sativa* spp. Nipponbare) genome sequence: sequencing of gap-specific fosmid clones uncovers new euchromatic portions of the genome. *Theor. Appl. Genet.*, **111**, 1596–1607.
24. Wagner,B., Miller,D.C., Lear,T.L. and Antczak,D.F. (2004) The complete map of the ig heavy chain constant gene region reveals evidence for seven igg isotypes and for IgD in the horse. *J. Immunol.*, **173**, 3230–3242.
25. Thomas,J.W., Prasad,A.B., Summers,T.J., Lee-Lin,S.-Q., Maduro,V.V.B., Idol,J.R., Ryan,J.F., Thomas,P.J., McDowell,J.C. and

Green,E.D. (2002) Parallel construction of orthologous sequence-ready clone contig maps in multiple species. *Genome Res.*, **12**, 1277–1285.

26. Kellner,W.A., Sullivan,R.T., Carlson,B.H., Program,N.C.S. and Thomas,J.W. (2005) Uprobe: A genome-wide universal probe resource for comparative physical mapping in vertebrates. *Genome Res.*, **15**, 166–173.

27. Yang,T.-J., Yu,Y., Chang,S.-B., de Jong,H., Oh,C.-S., Ahn,S.-N., Fang,E. and Wing,R. (2005) Toward closing rice telomere gaps: mapping and sequence characterization of rice subtelomere regions. *Theor. Appl. Genet.*, **111**, 467–478.

28. Romanov,M.N., Price,J.A. and Dodgson,J.B. (2003) Integration of animal linkage and BAC contig maps using overgo hybridization. *Cytogenet. Genome Res.*, **102**, 277–281.

29. Yu,Y., Tomkins,J.P., Waugh,R., Frisch,D.A., Kudrna,D., Kleinhofs,A., Brueggeman,R.S., Muehlbauer,G.J., Wise,R.P. and Wing,R.A. (2000) A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. *Theor. Appl. Genet.*, **101**, 1093–1099.