Original article

A first generation bovine BAC-based physical map

Laurent Schibler[a]*, Anne Roig[a], Marie-Françoise Mahé[a],
Jean-Claude Save[b], Mathieu Gautier[a], Sead Taourit[a],
Didier Boichard[c], André Eggen[a], Edmond P. Cribiu[a]

[a] Laboratoire de génétique biochimique et de cytogénétique, Département de génétique animale, Institut national de la recherche agronomique, Centre de recherche de Jouy, 78352 Jouy-en-Josas Cedex, France

[b] Laboratoire de radiobiologie et d'étude du génome, Département de génétique animale, Institut national de la recherche agronomique, Centre de recherche de Jouy, 78352 Jouy-en-Josas Cedex, France

[c] Station de génétique quantitative appliquée, Département de génétique animale, Institut national de la recherche agronomique, Centre de recherche de Jouy, 78352 Jouy-en-Josas Cedex, France

Abstract – A first generation clone-based physical map for the bovine genome was constructed combining, fluorescent double digestion fingerprinting and sequence tagged site (STS) marker screening. The BAC clones were selected from an Inra BAC library (105 984 clones) and a part of the CHORI-240 BAC library (26 500 clones). The contigs were anchored using the screening information for a total of 1303 markers (451 microsatellites, 471 genes, 127 EST, and 254 BAC ends). The final map, which consists of 6615 contigs assembled from 100 923 clones, will be a valuable tool for genomic research in ruminants, including targeted marker production, positional cloning or targeted sequencing of regions of specific interest.

cattle / physical map / BAC

1. INTRODUCTION

The detection of loci affecting economically important traits represents a major objective in livestock genomics. It should ultimately lead to more efficient breeding schemes (marker-assisted selection or MAS) and improve the accuracy and intensity of selection programs [18, 22]. In this perspective, genetic maps constructed in various livestock species [3, 12, 26, 29, 46] are sufficient to detect regions containing genes and QTL. The identification and

---

* Corresponding author: schibler@jouy.inra.fr

cloning of the corresponding genes may be achieved by standard positional cloning, taking advantage of the existence of large insert libraries and searching for transcribed sequences in these regions. Indeed, large insert libraries are available for livestock species, such as YAC libraries [1, 4, 23, 28, 38] and BAC libraries [5, 6, 14, 37, 41, 45, 48]. In addition, high-resolution comparative mapping projects have been developed in ruminants [2, 40] and pigs [36] in order to benefit from the spectacular progress of human and mouse genome projects, which will facilitate the comparative positional candidate gene approach. Likewise, systematic EST mapping projects have been initiated [42] to efficiently address the crucial step of finding candidate genes for a trait of interest.

To date, genes responsible for economic trait loci have been identified in ruminants [21, 29, 34]. In all cases, the authors used a contig assembly approach based on a walking strategy starting from only two entry points. Comparative mapping could be an efficient way to find new entry points and thus speed up the contig construction [39]. This sequential bi-directional walking strategy is however labor intensive and complicated by repetitive sequences (SINE or LINE) frequently encountered in either BAC end sequences or sequence tagged sites (STS). As in humans [9, 25], the construction of genome wide physical maps in cattle will provide essential tools for efficient positional cloning, target DNA marker development or large-scale sequencing.

Several approaches have been developed to build physical maps, including hybridization-based methods [15, 33], restriction-based fingerprinting methods [11, 35] or sequencing of a large number of BAC ends that serve as sequence-tagged connectors [47]. Since the restriction fingerprinting approach is less hindered by the presence of repeated sequences than the hybridization methods, is faster and more automation friendly, it has been considered a reasonable compromise to detect and measure clone overlaps. Two different fingerprinting approaches were initially developed, using either single [35] or double complete digestion [10] methods. Fingerprinting techniques were later improved in order to increase data quality and to support the generation of fingerprints from a number of clones representative of large genomes [19, 30]. Moreover, data analysis was greatly improved by the development of the FPC software [43], which combines automation and interactive graphics.

The availability of stable large-size clones and FPC software has renewed the fingerprinting method. Currently, BAC libraries have been extensively used to build numerous chromosome specific or whole genome sequence ready maps by fingerprinting [7, 8, 20, 24, 31, 32, 44]. Whole genome maps are being constructed for a number of organisms including the rat, cow, zebrafish, sorghum, maize and tomato (see www.genome.clemson.edu/fpc and www.bcgsc.edu for links to the corresponding web sites).

In this paper, we report the assembly of a first generation clone-based physical map, for the bovine genome, combining fluorescent double digestion fingerprinting and STS marker screening. This map represents the first report of genome wide physical mapping using the fluorescent double digestion fingerprinting method.

2. MATERIALS AND METHODS

2.1. BAC DNA preparation

BAC clones were grown in 2 mL 96-well blocks for 18 h 30 min at 37 °C with shaking, using 900 μL 2YT containing 12.5 μg·mL^{-1} chloramphenicol. Bacteria were pelleted by centrifugation at 3200 $\times g$ for 8 min. The blocks were inverted to discard the supernatant and stored at −20 °C for one week until DNA preparation.

DNA preparation was performed using a modified alkaline lysis procedure. Cell pellets were resuspended by addition of 100 μL of TGE + 30 μg·mL^{-1} RNAse and vigorous vortexing. Lysis was achieved by a 3 min incubation step after addition of 100 μL of NaOH 0.2 M + SDS 1%. One hundred microliters of ice-cold potassium acetate (1.32 M, pH 4.8) were then added to each well, followed by 100 μL of lithium chloride (8 M). After 20 min at −20 °C, the lysis products were purified using Millipore filter plates (Multiscreen MANLY). The samples were then precipitated with 250 μL of isopropanol and centrifuged at 3200 $\times g$ for 30 min with isopropanol. DNA pellets were washed with 250 μL of 70% ethanol per well and centrifuged for 10 min. The blocks were then inverted on paper towel to drain excess ethanol from the pellet and placed in a vacuum hybridization oven for 5 min to dry the DNA. Resuspension was achieved by addition of 10 μL of 10 mM Tris-HCl (pH 8.0) + 10 mg·mL^{-1} RNAse and vortexing after a 30 min incubation step at 37 °C.

The DNA concentration of 24 samples was estimated by fluorimetry for each plate. The plates were then calibrated at a mean concentration of 75 ng·μL^{-1} by addition of 10 mM Tris-HCl.

2.2. Fingerprinting

Restriction enzyme digestions and dye labeling were carried out simultaneously in a 10 μL reaction volume using 300−400 ng of BAC DNA, 2 units *Hin*dIII (Promega), 3 units *Hae*III (Promega), 0.5 units Sequenase II (Amersham Pharmacia Biotech) and 3 pmoles ddA (R110, R6G or TAMRA from NEN) in restriction buffer C (Promega). The reactions were carried out for

90 min at 37 °C. In order to normalize signal intensities, R6G-labeled samples
were diluted twice by adding 10 μL of sterile water.

2.3. Sample pooling and reaction cleanup

Multiscreen filter plates (Millipore, MAHVN45), filled with 360 μL of
preswollen Sephadex G50 (G50 superfine 80 g·L^{-1}, Amersham Pharmacia
Biotech) were used to achieve dye removal and sample desalting. Samples
(10 μL each) from three plates labeled with different dyes were then pooled on
the center of G50 mini-columns using a Hydra 96 dispenser. After a five-
minute incubation step, the samples were recovered by spinning down for
5 min at 910 ×g and stored at −20 °C until injection.

2.4. Sample injection and capillary electrophoresis

Seven microliters of the purified fingerprinting reaction were transferred to
the injection plate and three microliters of a size-standard mix (2.8 μL for-
mamide + 0.2 μL ET 900 ROX standard) were added. The samples were
denatured in a heater block for 3 min at 90 °C prior to injection in a MegaBace
1000 automated 96 capillary DNA sequencer. Injection parameters and run
conditions were respectively 3 kV for 90 s and 10 kV for 100 min using dye
set 2 filters.

2.5. Data analysis and cleanup

The runs were analyzed with the Genetic Profiler software developed to per-
form the genotyping analysis on the MegaBace. After spectral matrix correc-
tion and peak identification, this software creates a "pks" file for each capillary
that contains the scan value, the peak height and the peak width for all peaks
detected in each channel. Since the Genetic Profiler software takes all peaks
into account, it was necessary to eliminate background and artifactual peaks
before exporting data to an FPC formatted file. Therefore, we developed a Vi-
sual Basic software (available upon request to the author) taking a "pks" file as
input and applying three filters. The first one removes all peaks with a signal
height lower than 200 units or not in the range of the analysis (55 to 750 bp);
after an iterative estimation of the mean and standard deviations of the most
likely distribution of peak height, a second filter eliminates external bands,
the height of which is deviated by more than 2.5 standard deviations from the
mean; the third filter removes peaks in the size range (+/–1 bp) of known ar-
tifacts (76 and 129 bp additional bands generated by FAM, 80 bp by TAMRA
and 102 and 160 bp by R6G). This software rewrites "pks" files, so that modifi-
cations can be visualized by the Genetic Profiler. Good results were obtained,
since about 70% of the sample were perfectly cleaned, except for data with

low signals or when multiple clones (cross contaminations) were present in the same well. Therefore, to provide reliable data, all samples were manually checked and edited if necessary before exporting to FPC. As described in [13], we submited fragment sizes as ×10 in order to deal with some limitations of FPC: we could thus set the TOL value at 5 to take into account the fact that the standard deviation of fragment size measurements was less than one base.

2.6. BAC library screening

A PCR-based screening was performed on the Inra bovine BAC library as previously described [14]. Briefly, primer pairs for markers were obtained from the BOVMAP database (http://locus.jouy.inra.fr). PCR reactions were performed on PTC-100 thermocyclers (MJ Research) in a 15 μL reaction volume with 1× standard buffer supplemented with 125 μM dNTP, 1.5 mM MgCl$_2$, 0.5 μM of each primer and 0.035 U$\cdot\mu$L^{-1} Taq polymerase (Promega). The samples were preheated for 5 min at 94 °C, subjected to 35 cycles (94 °C for 20 s, 50−60 °C for 30 s and 72 °C for 30 s), and a final extension step of 5 min at 72 °C.

2.7. BAC end sequencing

The Nucleobond AX100 kit (Macherey-Nagel) was used to prepare a BAC DNA suitable for end sequencing according to the manufacturer's recommendations. Sequencing reactions were performed as previously described [39].

2.8. Contig assembly

The map was constructed with FPC starting from an initial stringent build and using an incremental process, which consisted in joining contigs together based on end-end comparisons. This time consuming step requires generally to manually fuse and reanalyze contigs at higher cutoff values. In order to automate this work and to trace weak joins introduced by manual editing, we developed a new strategy using virtual markers. In fact, FPC allows a less stringent cutoff for clones sharing one or more markers and the "Incremental Build Contig" (IBC) option takes into consideration new marker data to automatically merge and reanalyse contigs. We thus transcribed end-end comparisons results in term of marker data. For example, end-end comparisons showed that clone bI0053B12 matched clone bI0133E09 at a 2×10^{-11} value. We thus defined a Virtual End (VE) Marker (called V_I0053B12_2E-11) hitting these two clones. Marker data were incorporated in FPC through *.ace files using the replace marker command. The same strategy was used to incorporate singleton at higher cutoff.

In practice, the first map was obtained at a 10^{-13} cutoff value using the pure Sulston method, with the Tolerance parameter set to 5 (0.5 base in fact). Contigs with more than 5 Q (questionable clones that FPC failed to place properly in the map) were automatically reanalyzed at a lower cutoff value using the DQER (from 10^{-13} to 10^{-25}) from FPC software. Three large contigs with more than 10 Q remained unchanged after reanalysis. These contigs encompassed 1059 clones, which displayed more than 75% shared bands with an unusually high number of related clones. These clones were discarded as they may represent clones with large repeated or duplicated sequences. Moreover, all contigs encompassing more than 15 clones were manually edited and split when they appeared doubtful. Singletons were then incorporated up to a 10^{-10} cutoff value using virtual markers (called S_*clone name*_*CutOff* in FPC).

The next building step was an end-end comparison at a 10^{-12} cutoff value. We retained as valuable, only reciprocal and unique fusions. Moreover, contigs split by the last DQER step and manual editing, could not be merged. Then, we automatically fused contigs based on end-end comparisons using Virtual Ends (VE) Markers (called *V_Clone Name_CutOff* in FPC) and the IBC option (see above). Fusions, which produced contigs with more than 5 Q, were rejected by deleting the corresponding virtual marker and reanalyzing with IBC. Remaining VE markers are denoted validated VE markers in Table II. All steps required to create or delete VE markers, were done automatically using an Access Database, managing files containing the detailed end-end results (by printing in a file the FPC standard output), the summarized end-end results (by saving the FPC results window), the contig results (by saving the FPC By Contigs window) and the IBC results (by saving the FPC results window). Four additional steps were performed as described, making it possible to merge ends sequentially from 10^{-11} to 10^{-08}. In a last step, we merged contigs based on our framework markers and checked contigs with more than 50 clones to split manually those which seemed to be doubtful.

3. RESULTS

3.1. Fingerprinting of BAC clones

Two libraries were used: the Inra BAC library [14] with 105 984 clones and part of the CHORI-240 (www.chori.org/bacpac) BAC library (26 500 clones). About 9% of the clones were lost during the fingerprinting process, due to poor DNA preparation or low DNA yield (7%), injection failure or damaged capillaries (2%). Cross contamination was identified in about 2% of the wells.

Empty or small clones showing a low number of bands were rejected at the data cleanup step, in order to retain only clones with more than 7 bands for Inra clones and 11 bands for CHORI-240 clones. Moreover, since a clone
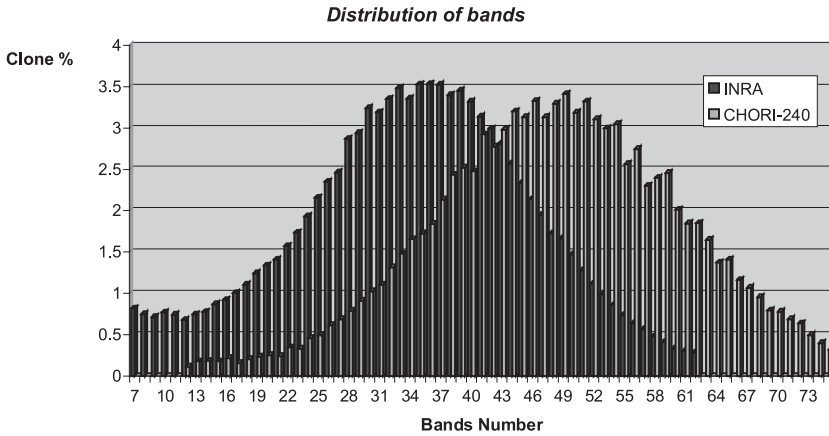
**Figure 1.** Band number distribution after data cleanup.

Table I. Library statistics.

| Library	Number of clones	Number of lost clones	Empty clones	Small clones	Number of cancelled clones	Number of clones in FPC	Mean nb of bands	Mean FPC sizes (bp)	Mean size (bp)
Inra	105 984	7%	5%	6%	1225	82 010	35	10 658	114 000
CHORI-240	26 500	17%	6%	2%	577	18 913	48	15 451	180 000
Total	132 484	9%	5.2%	5.2%	1802	100 923	37	11 609	

with a high number of bands may correspond to two different clones in a same well, INRA clones with more than 63 bands and CHORI-240 clones with more than 75 bands were marked as cancelled in FPC. Finally, 102 725 clones were transferred to FPC, 40 of which were fingerprinted twice and 1802 were cancelled. Table I summarizes all these results and Figure 1 shows the distribution of band numbers after data cleanup.

3.2. Contig assembly

Preliminary maps built using different cutoff values show that cutoff values ranging from 10^{-12} to 10^{-08} produced contigs consistent with either the expected results based on statistical analysis [27] or on our contigs built by chromosome walking (unpublished data). We used two criteria to evaluate the quality of the construct: the number of clones poorly incorporated (Q clones) and the number of contigs exhibiting at least one Q clone (Q contigs). As shown in Figure 2, higher cutoff values (10^{-7}) generate many false positive contigs (an increase in Q clones), with one contig containing up to
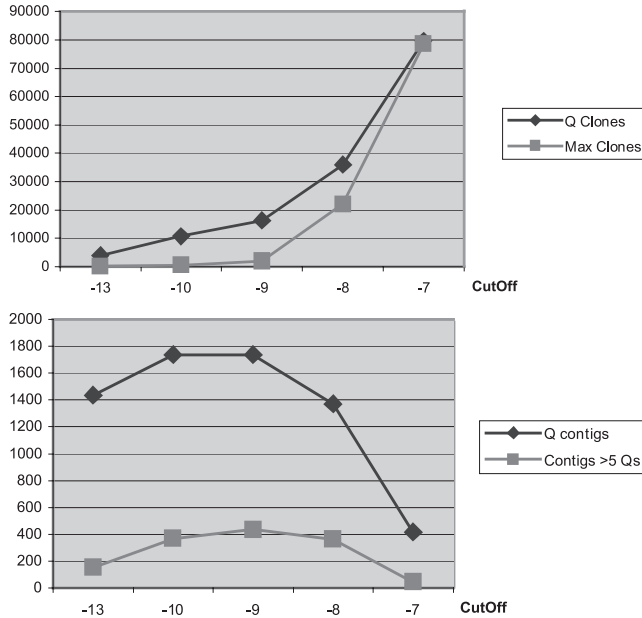
Figure 2. CutOff effect on the reliability of the map.
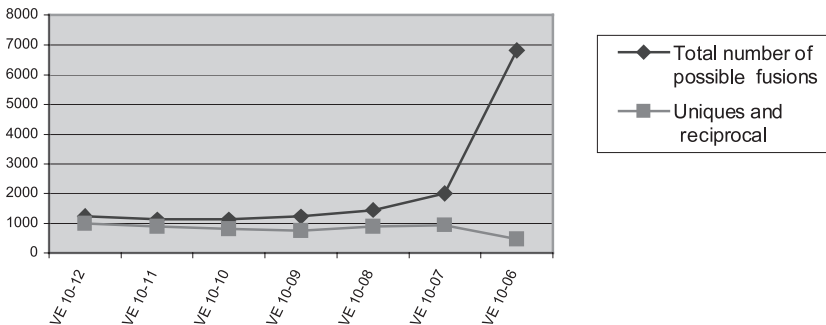
Figure 3. Possible and valuable fusions at each fusion step.

70 000 clones. Moreover, Figure 2 shows clearly that map buildings are more stable in terms of Q clones and Q contig number for cutoff values in the 10^{-12} to 10^{-09} range.

Moreover, we analyzed more precisely the differences between 10^{-10} and 10^{-12} cutoff value maps after multiple DQER steps. We observed that in most cases, contigs built at a cutoff value of 10^{-10} represent end fusions of 2 or more contigs built at 10^{-12}.

Table II. Results of the bovine map iterative automated assembly.

Step	Contigs	Q contigs	Q Contigs >5 Qs	VE markers	Validated VE Markers	Q clones
Build 10^{-13} DQER	10 944	1741	7			2794
Build VE 10^{-12}	9994	1785	12	1167	1064	3058
Build VE 10^{-11}	9099	1835	12	1124	1051	3263
Build VE 10^{-10}	8364	1855	12	949	820	3433
Build VE 10^{-09}	7648	1894	12	877	811	3608
Build VE 10^{-08}	6890	1942	12	1121	858	3848

Q clones: clones poorly incorporated into contigs. Q contigs: contigs containing at least one Q clone. VE markers: virtual end marker obtained by ends-ends comparison in FPC. Validated VE markers: VE markers that do not increase the number of Q after merging the contigs and thus retained in subsequent analysis.

Thus, contig assembly seemed to be reliable in the 10^{-12} to 10^{-09} cutoff range. However, to provide reliable data, we used an incremental process to build our final map, as described in the Materials and methods starting from a stringent build obtained at a 10^{-13} cutoff. We stopped this incremental process at 10^{-08}, since higher cutoff values generated too many non reciprocal and non unique end fusions (Fig. 3).

Table II shows the results obtained at each fusion step, the map encompassing 6890 contigs obtained using 4604 virtual markers and 4157 fusions.

3.3. Map validation and anchoring by PCR screening

A total of 1390 markers was screened on the Inra BAC library, 87 of them being absent from our library. Thus the screening information for a total of 1303 markers (451 microsatellites, 471 genes, 127 EST, and 254 BAC-ends) was used to anchor the contigs and validate our building strategy. These markers were derived from existing bovine genetic and radiation hybrid maps or according to their position on the human genome. In particular, we used 80 type I and II markers placed on a BTA26 comprehensive radiation hybrid map [17].

Fifty-five contigs contained at least one VE marker flanked by screened markers. In all cases, the locations of flanking markers were consistent, suggesting that VE markers introduced no error.

About thirty contigs were built in the course of this work by chromosome walking using framework preliminary maps with VE markers. BAC ends were sequenced and screened on the BAC library, making it possible to confirm 152 VE markers (data not shown). No erroneous fusion could be detected.

Table III. The June 2003 bovine map.

	Number of clones in the contig							Total
	>200	200–100	100–50	50–25	25–10	9–3	2	
Contigs	0	10	196	874	2032	2815	688	6615
Q contigs (>5 Q)	0	8	6	5	0	0	0	19
Q clones		117	909	1752	1140	0	0	3918
Anchored Contigs	0	10	103	223	264	132	15	747

These screening data made it possible to fuse some more contigs, resulting in the first generation bovine physical map.

3.4. The bovine physical map

The current release (June 2003) encompasses 6615 contigs, 747 of them being anchored. The contigs contain 15 clones on average (from two to 164 clones). Table III shows the number of contigs, the Q contigs and number of Q clones for 7 contig size classes.

Our 6615 contigs cover a total of 791 214 bands, each of them representing about 3.5 kb based on 82 000 Inra clones (~120 kb) showing 35 bands and 19 000 CHORI clones (~180 kb) showing 48 bands. Thus, our contigs cover about 400 kb on average and the map covers about 2769 Mb, *i.e.*, more than 90% of the bovine genome. Singletons and cancelled clones containing large repetitive sequences represent about 117 000 additional bands (*i.e.*, ~14%). Therefore, the BAC map covers virtually all the bovine genome. However, this may represent a slight overestimation because of undetected overlaps.

In terms of anchored contigs, the map covers 171 500 bands representing about 600 Mb, *i.e.* 20% of the bovine genome.

The bovine physical map will be continuously updated, based on new marker screening data. It will be publicly available through the BovMap database and webFPC at http://locus.jouy.inra.fr/cgi-bin/lgbc/mapping/common/intro2.pl?BASE=cattle.

Inra BAC clones can be obtained at the Biological Resources Center (http://www-crb.jouy.inra.fr).

3.5. A first draft of a BTA 26 physical map

In order to test the genome coverage of our contigs and the usefulness of our map, we screened a first set of 34 evenly spaced markers from BTA 26. It was thus possible to anchor 23 contigs and to detect two regions not represented in our library (Fig. 4). These 22 contigs covered about 30 Mb, *i.e.* 50% of

the estimated 60 Mb total length of BTA26 [17]. Thus, screening of about 30 evenly spaced markers made it possible to increase from the chromosome coverage with anchored contigs 20% to 50%.

The remaining 44 markers from the BTA26 RH map [17] were then screened: five new contigs were identified and six previously identified contigs were enlarged or fused.

Finally this preliminary BTA26 physical map covers about 35 Mb, which correspond to about 60% of BTA26.

The results are concordant in terms of gene order with the radiation hybrid map data except for five contigs (Ctg520, Ctg146, Ctg525, Ctg530 and Ctg150). These discrepancies concern very closely related markers separated by a distance under the resolution limit of our panel [16].

In addition, these results are consistent in terms of contig length, except for Ctg146 and Ctg525 for which the physical sizes (1700 kb and 2050 kb) represent about 170 and 90 cR_{3000} instead of 60 and 70 cR_{3000} respectively, based on the converting ratio previously defined from human physical data [16].

4. DISCUSSION

We developed a first generation genome-wide BAC-based physical map of the bovine genome. This map consists of 6756 contigs assembled from 100 923 clones selected in two libraries. Restriction profiling was achieved using a fluorescent fingerprinting technique, based on capillary electrophoresis to analyze samples. This strategy has proven to be efficient in terms of speed and data quality. Four persons were sufficient to generate about 2000 fingerprint profiles per day on a MegaBace 1000, using little automation (one Hydra 96). Capillary electrophoresis reduces the variations of fragment sizes to less than 1 bp in a range of 50 to 750 bp, making it possible to more accurately dedect overlaps. However, efficient fragment analysis software is necessary to simplify and speed up the data cleanup process, which required two persons for six months.

Our iterative strategy makes it possible to not only automate the manual merging step but also to precisely monitor the contig assembly process and thus to stop it before data is corrupted. No error could be detected based on 50 contigs containing at least one VE marker flanked by screened markers. Since these 50 contigs represent 118 virtual fusions, we can expect, with a risk lower than 5%, that less than 2.5% of the contigs could be falsely branched by VE markers.

Moreover, the use of VE markers allows any user to trace all 'weak joints' introduced by manual editing. Indeed, the only way to detect a merge at a high cutoff value is to reanalyze this contig. VE markers clearly indicate which
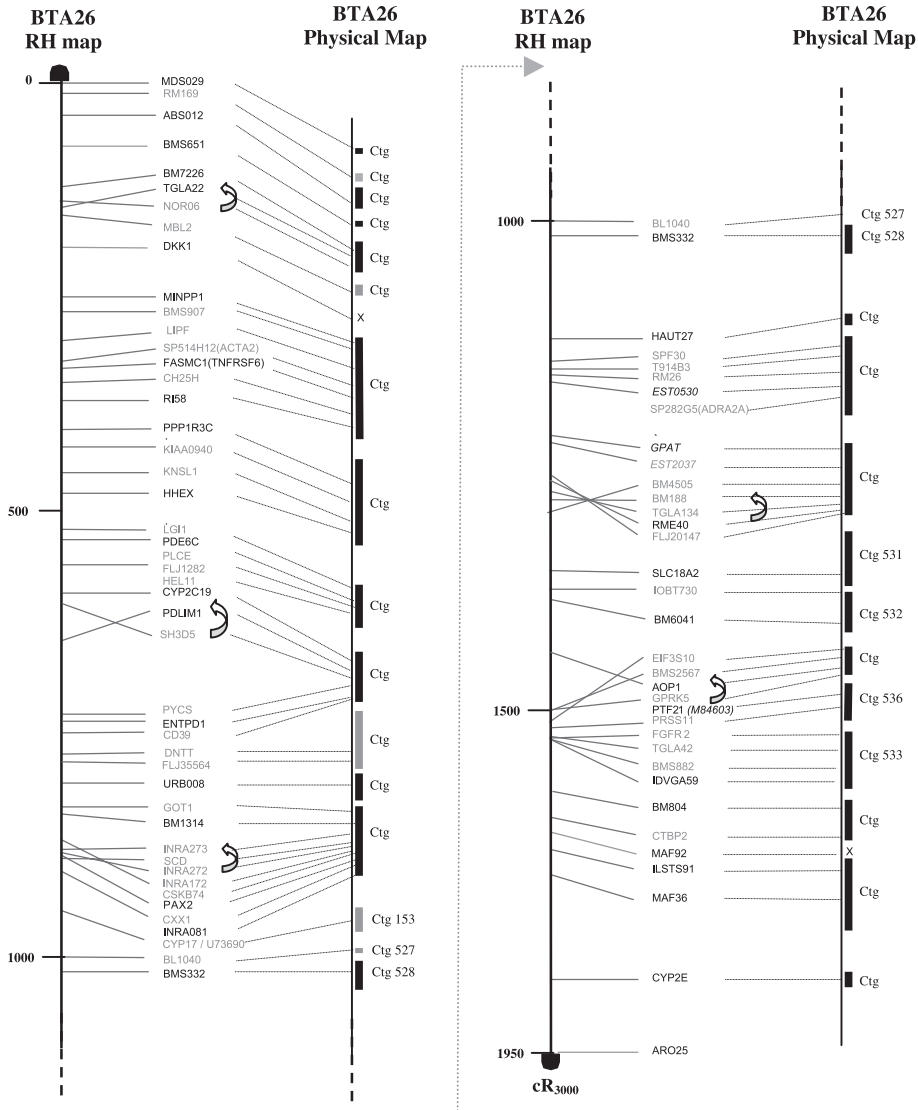
Figure 4. First draft of the BTA26 physical map. Correspondence of the physical map of BTA26 (right) with the radiation hybrid map (left) are indicated by dotted lines. Markers and contigs from the first set are painted in black. The results from the second set are in gray. Contig numbers are indicated as Ctg. X correspond to clones absent in the library for a specific marker and arrows represent discrepancies between the radiation and the physical map.

clones and which cutoff values were used to merge contigs. This could be of great interest for users accessing contigs through webFPC, for example.

The clones analyzed in this study were submitted to a double digestion, which generated about 40 bands on average. It should be possible to achieve a three to fourfold increase of band number since the MegaBace can accurately resolve fragments in the 50−750 bp range. A higher number of bands could be of great interest to detect smaller overlaps. For example, two clones with 30% overlap could share 11 bands out of 35 or 22 bands out of 70. Their coincidence scores would be $2E^{-08}$ and $9E^{-10}$, respectively. Many restriction enzymes could be used simultaneously to generate more bands and thus detect 20% overlaps. This strategy could be an alternative to the strategy of Ding *et al.* [13]. However, these authors suggest that too many bands (~100 bands) could be troublesome for FPC to handle. A three-digestion strategy generating about 70 bands and permitting 25−30% could thus be a good compromise.

Our BAC–based map will be a valuable tool for genomic research in ruminants, including targeted marker production, positional cloning or targeted sequencing of regions of specific interest. Even if our map provides only a four-genome equivalent coverage, it may not be worthwhile to spend more time, fingerprinting additional clones. Three elements support this statement:

– Firstly, STS content mapping may be more effective than fingerprinting for achieving gap closure and contig joining. About 3000 additional markers evenly distributed on the whole genome should be sufficient to achieve a 60% coverage with anchored contigs based on our results on BTA26. About 6000 markers would allow most gap closure. At present, comparative mapping data combined with the human sequence makes it possible to quickly identify these 6000 markers and to develop bovine specific primers from the numerous bovine EST available in databases. This is encouraging since the main applications of this physical map could be dedicated to large contig construction to assist positional cloning in the ETL regions.

– Secondly, our map provides a good framework to initiate a strategy similar to that of Gregory *et al.* [20] and establish high-resolution syntenies among ruminant, human and mouse genomes. About 60 000 bovine BES from the CHORI-240 library have been submitted to GenBank, 10 000 of them corresponding to clones integrated in our contigs. End sequencing of singletons and clones at the end of our contigs will provide about 60 000 additional BES. BLASTN comparison with the human genome should thus provide at least 5000 significant "hits", making it possible to align most bovine contigs along the human genome. The deduced bovine contig juxtaposition could be helpful to identify potential overlapping contigs, which could be easily joined by PCR screening.

– Thirdly, an international physical map is under development (www.bcgsc.bc.ca/projects/bovine_mapping, http://www.livestockgenomics. csiro.au/cattle.shtml) by analyzing single digest fingerprints obtained from 280 000 BAC. Amongst these, 18 913 clones from the CHORI-240 library were also incorporated in our map. Even if two different fingerprinting strategies were used, clones common to both maps could serve as anchors to identify news fusions between contigs from the two maps. Cross validation between these two independently constructed maps should provide a reliable framework to start whole genome sequencing projects.

ACKNOWLEDGEMENTS

REFERENCES

[1] Alexander L.J., Smith T.P., Beattie C.W., Broom M.F., Construction and characterization of a large insert porcine YAC library, Mamm. Genome 8 (1997) 50–51.

[2] Band M.R., Larson J.H., Rebeiz M., Green C.A., Heyen D.W., Donovan J., Windish R., Steining C., Mahyuddin P., Womack J.E., Lewin H.A., An ordered comparative map of the cattle and human genomes, Genome Res. 10 (2000) 1359–1368.

[3] Barendse W., Vaiman D., Kemp S.J., Sugimoto Y., Armitage S.M., Williams J.L., Sun H.S., Eggen A., Agaba M., Aleyasin S.A., Band M., Bishop M.D., Buitkamp J., Byrne K., Collins F., Cooper L., Coppettiers W., Denys B., Drinkwater R.D., Easterday K., Elduque C., Ennis S., Erhardt G., Li L., *et al.*, A medium-density genetic linkage map of the bovine genome, Mamm. Genome 8 (1997) 21–28.

[4] Broom M.F., Hill D.F., Construction of a large-insert yeast artificial chromosome library from sheep DNA, Mamm. Genome 5 (1994) 817–819.

[5] Buitkamp J., Kollers S., Durstewitz G., Fries R., Welzel K., Schafer K., Kellermann A., Lehrach H., Construction and characterization of a gridded cattle BAC library, Anim. Genet. 31 (2000) 347–351.

[6] Cai L., Taylor J.F., Wing R.A., Gallagher D.S., Woo S.S., Davis S.K., Construction and characterization of a bovine bacterial artificial chromosome library, Genomics 29 (1995) 413–425.

[7] Cao Y., Kang H.L., Xu X., Wang M., Dho S.H., Huh J.R., Lee B.J., Kalush F., Bocskai D., Ding Y., Tesmer J.G., Lee J., Moon E., Jurecic V., Baldini A., Weier H.U., Doggett N.A., Simon M.I., Adams M.D., Kim U.J., A 12-Mb complete coverage BAC contig map in human chromosome 16 p13.1- p11.2, Genome Res. 9 (1999) 763–774.

[8] Chen M., Presting G., Barbazuk W.B., Goicoechea J.L., Blackmon B., Fang G., Kim H., Frisch D., Yu Y., Sun S., Higingbottom S., Phimphilai J., Phimphilai D., Thurmond S., Gaudette B., Li P., Liu J., Hatfield J., Main D., Farrar K., Henderson C., Barnett L., Costa R., Williams B., Walser S., Atkins M., Hall C., Budiman M.A., Tomkins J.P., Luo M., Bancroft I., Salse J., Regad F., Mohapatra T., Singh N.K., Tyagi A.K., Soderlund C., Dean R.A., Wing R.A., An integrated physical and genetic map of the rice genome, Plant Cell 14 (2002) 537–545.

[9] Chumakov I.M., Rigault P., Le Gall I., Bellanne-Chantelot C., Billault A., Guillou S., Soularue P., Guasconi G., Poullier E., Gros I., *et al.*, A YAC contig map of the human genome, Nature 377 (1995) 175–297.

[10] Coulson A., Sulston J., Brenner S., Karn J., Toward a physical map of the genome of the nematode *Caenorhabditis elegans*, Proc. Natl. Acad. Sci. USA 83 (1986) 7821–7825.

[11] Coulson A., Waterston R., Kiff J., Sulston J., Kohara Y., Genome linking with yeast artificial chromosomes, Nature 335 (1988) 184–186.

[12] de Gortari M.J., Freking B.A., Cuthbertson R.P., Kappes S.M., Keele J.W., Stone R.T., Leymaster K.A., Dodds K.G., Crawford A.M., Beattie C.W., A second-generation linkage map of the sheep genome, Mamm. Genome 9 (1998) 204–209.

[13] Ding Y., Johnson M.D., Colayco R., Chen Y.J., Melnyk J., Schmitt H., Shizuya H., Contig assembly of bacterial artificial chromosome clones through multiplexed fluorescence-labeled fingerprinting, Genomics 56 (1999) 237–246.

[14] Eggen A., Gautier M., Billaut A., Petit E., Hayes H., Laurent P., Urban C., Pfister-Genskow M., Eilertsen K., Bishop M.D., Construction and characterization of a bovine BAC library with four genome-equivalent coverage, Genet. Sel. Evol. 33 (2001) 543–548.

[15] Evans G.A., Lewis K.A., Physical mapping of complex genomes by cosmid multiplex analysis, Proc. Natl. Acad. Sci. USA 86 (1989) 5030–5034.

[16] Gautier M., Hayes H., Eggen A., An extensive and comprehensive radiation hybrid map of bovine Chromosome 15: comparison with human Chromosome 11, Mamm. Genome 13 (2002) 316–319.

[17] Gautier M., Hayes H., Eggen A., A comprehensive radiation hybrid map of bovine chromosome 26 (BTA26): comparative chromosomal organization between HSA10q and BTA26 and BTA28, Mamm. Genome 14 (2003) 711–721.

[18] Georges M., Andersson L., Livestock genomics comes of age, Genome Res. 6 (1996) 907–921.

[19] Gregory S.G., Howell G.R., Bentley D.R., Genome mapping by fluorescent fingerprinting, Genome Res. 7 (1997) 1162–1168.

[20] Gregory S.G., Sekhon M., Schein J., Zhao S., Osoegawa K., Scott C.E., Evans R.S., Burridge P.W., Cox T.V., Fox C.A., Hutton R.D., Mullenger I.R., Phillips K.J., Smith J., Stalker J., Threadgold G.J., Birney E., Wylie K., Chinwalla A., Wallis J., Hillier L., Carter J., Gaige T., Jaeger S., Kremitzki C., Layman D., Maas J., McGrane R., Mead K., Walker R., Jones S., Smith M., Asano J.,

Bosdet I., Chan S., Chittaranjan S., Chiu R., Fjell C., Fuhrmann D., Girn N., Gray C., Guin R., Hsiao L., Krzywinski M., Kutsche R., Lee S.S., Mathewson C., McLeavy C., Messervier S., Ness S., Pandoh P., Prabhu A.L., Saeedi P., Smailus D., Spence L., Stott J., Taylor S., Terpstra W., Tsai M., Vardy J., Wye N., Yang G., Shatsman S., Ayodeji B., Geer K., Tsegaye G., Shvartsbeyn A., Gebregeorgis E., Krol M., Russell D., Overton L., Malek J.A., Holmes M., Heaney M., Shetty J., Feldblyum T., Nierman W.C., Catanese J.J., Hubbard T., Waterston R.H., Rogers J., de Jong P.J., Fraser C.M., Marra M., McPherson J.D., Bentley D.R., A physical map of the mouse genome, Nature 418 (2002) 743–750.

[21] Grobet L., Martin L.J., Poncelet D., Pirottin D., Brouwers B., Riquet J., Schoeberlein A., Dunner S., Ménissier F., Massabanda J., Fries R., Hanset R., Georges M., A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle, Nat. Genet. 17 (1997) 71–74.

[22] Haley C.S., Livestock QTLs – bringing home the bacon?, Trends Genet. 11 (1995) 488–492.

[23] Hills D., Tracey S., Masabanda J., Fries R., Schalkwyk L.C., Lehrach H., Miller J.R., Williams J.L., A bovine YAC library containing four- to five-fold genome equivalents, Mamm. Genome 10 (1999) 837–838.

[24] Hoskins R.A., Nelson C.R., Berman B.P., Laverty T.R., George R.A., Ciesiolka L., Naeemuddin M., Arenson A.D., Durbin J., David R.G., Tabor P.E., Bailey M.R., DeShazo D.R., Catanese J., Mammoser A., Osoegawa K., de Jong P.J., Celniker S.E., Gibbs R.A., Rubin G.M., Scherer S.E., A BAC-based physical map of the major autosomes of *Drosophila melanogaster*, Science 287 (2000) 2271–2274.

[25] Hudson T.J., Stein L.D., Gerety S.S., Ma J., Castle A.B., Silva J., Slonim D.K., Baptista R., Kruglyak L., Xu S.H., *et al.*, An STS-based map of the human genome [see comments], Science 270 (1995) 1945–1954.

[26] Kappes S.M., Keele J.W., Stone R.T., McGraw R.A., Sonstegard T.S., Smith T.P., Lopez-Corrales N.L., Beattie C.W., A second-generation linkage map of the bovine genome, Genome Res. 7 (1997) 235–249.

[27] Lander E.S., Waterman M.S., Genomic mapping by fingerprinting random clones: a mathematical analysis, Genomics 2 (1988) 231–239.

[28] Libert F., Lefort A., Okimoto R., Womack J., Georges M., Construction of a bovine genomic library of large yeast artificial chromosome clones, Genomics 18 (1993) 270–276.

[29] Maddox J.F., Davies K.P., Crawford A.M., Hulme D.J., Vaiman D., Cribiu E.P., Freking B.A., Beh K.J., Cockett N.E., Kang N., Riffkin C.D., Drinkwater R., Moore S.S., Dodds K.G., Lumsden J.M., van Stijn T.C., Phua S.H., Adelson D.L., Burkin H.R., Broom J.E., Buitkamp J., Cambridge L., Cushwa W.T., Gerard E., Galloway S.M., Harrison B., Hawken R.J., Hiendleder S., Henry H.M., Medrano J.F., Paterson K.A., Schibler L., Stone R.T., van Hest B., An enhanced linkage map of the sheep genome comprising more than 1000 loci, Genome Res. 11 (2001) 1275–1289.

[30] Marra M.A., Kucaba T.A., Dietrich N.L., Green E.D., Brownstein B., Wilson R.K., McDonald K.M., Hillier L.W., McPherson J.D., Waterston R.H., High throughput fingerprint analysis of large-insert clones, Genome Res. 7 (1997) 1072–1084.

[31] McPherson J.D., Marra M., Hillier L., Waterston R.H., Chinwalla A., Wallis J., Sekhon M., Wylie K., Mardis E.R., Wilson R.K., Fulton R., Kucaba T.A., Wagner-McPherson C., Barbazuk W.B., Gregory S.G., Humphray S.J., French L., Evans R.S., Bethel G., Whittaker A., Holden J.L., McCann O.T., Dunham A., Soderlund C., Scott C.E., Bentley D.R., Schuler G., Chen H.C., Jang W., Green E.D., Idol J.R., Maduro V.V., Montgomery K.T., Lee E., Miller A., Emerling S., Kucherlapati R., Gibbs R., Scherer S., Gorrell J.H., Sodergren E., Clerc-Blankenburg K., Tabor P., Naylor S., Garcia D., de Jong P.J., Catanese J.J., Nowak N., Osoegawa K., Qin S., Rowen L., Madan A., Dors M., Hood L., Trask B., Friedman C., Massa H., Cheung V.G., Kirsch I.R., Reid T., Yonescu R., Weissenbach J., Bruls T., Heilig R., Branscomb E., Olsen A., Doggett N., Cheng J.F., Hawkins T., Myers R.M., Shang J., Ramirez L., Schmutz J., Velasquez O., Dixon K., Stone N.E., Cox D.R., Haussler D., Kent W.J., Furey T., Rogic S., Kennedy S., Jones S., Rosenthal A., Wen G., Schilhabel M., Gloeckner G., Nyakatura G., Siebert R., Schlegelberger B., Korenberg J., Chen X.N., Fujiyama A., Hattori M., Toyoda A., Yada T., Park H.S., Sakaki Y., Shimizu N., Asakawa S., *et al.*, A physical map of the human genome, Nature 409 (2001) 934–941.

[32] Mozo T., Dewar K., Dunn P., Ecker J.R., Fischer S., Kloska S., Lehrach H., Marra M., Martienssen R., Meier-Ewert S., Altmann T., A complete BAC-based physical map of the *Arabidopsis thaliana* genome, Nat. Genet. 22 (1999) 271–275.

[33] Mozo T., Fischer S., Meier-Ewert S., Lehrach H., Altmann T., Use of the IGF BAC library for physical mapping of the *Arabidopsis thaliana* genome, Plant J. 16 (1998) 377–384.

[34] Mulsant P., Lecerf F., Fabre S., Schibler L., Monget P., Lanneluc I., Pisselet C., Riquet J., Monniaux D., Callebaut I., Cribiu E.P., Thimonier J., Teyssier J., Bodin L., Cognie Y., Chitour N., Elsen J.M., Mutation in bone morphogenetic protein receptor-IB is associated with increased ovulation rate in Booroola Merino ewes, Proc. Natl. Acad. Sci. USA 98 (2001) 5104-5109.

[35] Olson M.V., Dutchik J.E., Graham M.Y., Brodeur G.M., Helms C., Frank M., MacCollin M., Scheinman R., Frank T., Random-clone strategy for genomic restriction mapping in yeast, Proc. Natl. Acad. Sci. USA 83 (1986) 7826–7830.

[36] Pinton P., Schibler L., Cribiu E.P., Gellin J., Yerle M., Localization of 113 anchor loci in pigs: improvement of the comparative map for humans, pigs, and goats, Mamm. Genome 11 (2000) 306–315.

[37] Rogel-Gaillard C., Bourgeaux N., Billault A., Vaiman M., Chardon P., Construction of a swine BAC library: application to the characterization and mapping of porcine type C endoviral elements, Cytogenet. Cell Genet. 85 (1999) 205–211.

[38] Rogel-Gaillard C., Bourgeaux N., Save J.C., Renard C., Coullin P., Pinton P., Yerle M., Vaiman M., Chardon P., Construction of a swine YAC library allowing an efficient recovery of unique and centromeric repeated sequences, Mamm. Genome 8 (1997) 186–192.

[39] Schibler L., Cribiu E.P., Oustry-Vaiman A., Furet J.P., Vaiman D., Fine mapping suggests that the goat Polled Intersex Syndrome and the human Blepharophimosis Ptosis Epicanthus Syndrome map to a 100-kb homologous region, Genome Res. 10 (2000) 311–318.

[40] Schibler L., Vaiman D., Oustry A., Giraud-Delville C., Cribiu E.P., Comparative gene mapping: A fine-scale survey of chromosome rearrangements between ruminants and humans, Genome Res. 8 (1998) 901–915.

[41] Schibler L., Vaiman D., Oustry A., Guinec N., Dangy-Caye A.L., Billault A., Cribiu E.P., Construction and extensive characterization of a goat bacterial artificial chromosome library with threefold genome coverage, Mamm. Genome 9 (1998) 119–124.

[42] Smith T.P., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T., Casas E., Wray J.E., White J., Cho J., Fahrenkrug S.C., Bennett G.L., Heaton M.P., Laegreid W.W., Rohrer G.A., Chitko-McKown C.G., Pertea G., Holt I., Karamycheva S., Liang F., Quackenbush J., Keele J.W., Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle, Genome Res. 11 (2001) 626–630.

[43] Soderlund C., Humphray S., Dunham A., French L., Contigs built with fingerprints, markers, and FPC V4.7, Genome Res. 10 (2000) 1772–1787.

[44] Tao Q., Chang Y.L., Wang J., Chen H., Islam-Faridi M.N., Scheuring C., Wang B., Stelly D.M., Zhang H.B., Bacterial artificial chromosome-based physical map of the rice genome constructed by restriction fingerprint analysis, Genetics 158 (2001) 1711–1724.

[45] Vaiman D., Billault A., Tabet-Aoul K., Schibler L., Vilette D., Oustry-Vaiman A., Soravito C., Cribiu E.P., Construction and characterization of a sheep BAC library of three genome equivalents, Mamm. Genome 10 (1999) 585–587.

[46] Vaiman D., Schibler L., Bourgeois F., Oustry A., Amigues Y., Cribiu E.P., A genetic linkage map of the male goat genome, Genetics 144 (1996) 279–305.

[47] Venter J.C., Smith H.O., Hood L., A new strategy for genome sequencing, Nature 381 (1996) 364–366.

[48] Zhu B., Smith J.A., Tracey S.M., Konfortov B.A., Welzel K., Schalkwyk L.C., Lehrach H., Kollers S., Masabanda J., Buitkamp J., Fries R., Williams J.L., Miller J.R., A 5× genome coverage bovine BAC library: production, characterization, and distribution, Mamm. Genome 10 (1999) 706–709.