

# Chromosome integration of BAC (bacterial artificial chromosome): evidence of multiple rearrangements

Agnès Le Saux · Louis-Marie Houdebine ·  
Geneviève Jolivet

Received: 14 July 2009 / Accepted: 12 January 2010 / Published online: 27 January 2010  
© Springer Science+Business Media B.V. 2010

**Abstract** This paper reports our attempts to characterize transgene integration sites in transgenic mouse lines generated by the microinjection of large (from 30 to 145 kb) pig DNA fragments encompassing a mammary specific gene, the whey acidic protein gene (WAP). Among the various methods used, the thermal asymmetric interlaced (TAIL-) PCR method allowed us (1) to analyze transgene/genomic borders and internal concatamer junctions for eleven transgenic lines, (2) to obtain sequence information for seven borders, (3) to place three transgenes in the mouse genome, and (4) to obtain sequence data for seven transgene junctions in concatamers. Finally, we characterized various rearrangements in the borders and the inner parts of the transgene. The possibility of such complex rearrangements should be carefully considered when transgenic animals are produced with large genomic DNA fragments.

**Keywords** Transgene · Bacterial artificial chromosome · Integration site · TAIL-PCR · Thermal asymmetric interlaced PCR

## Introduction

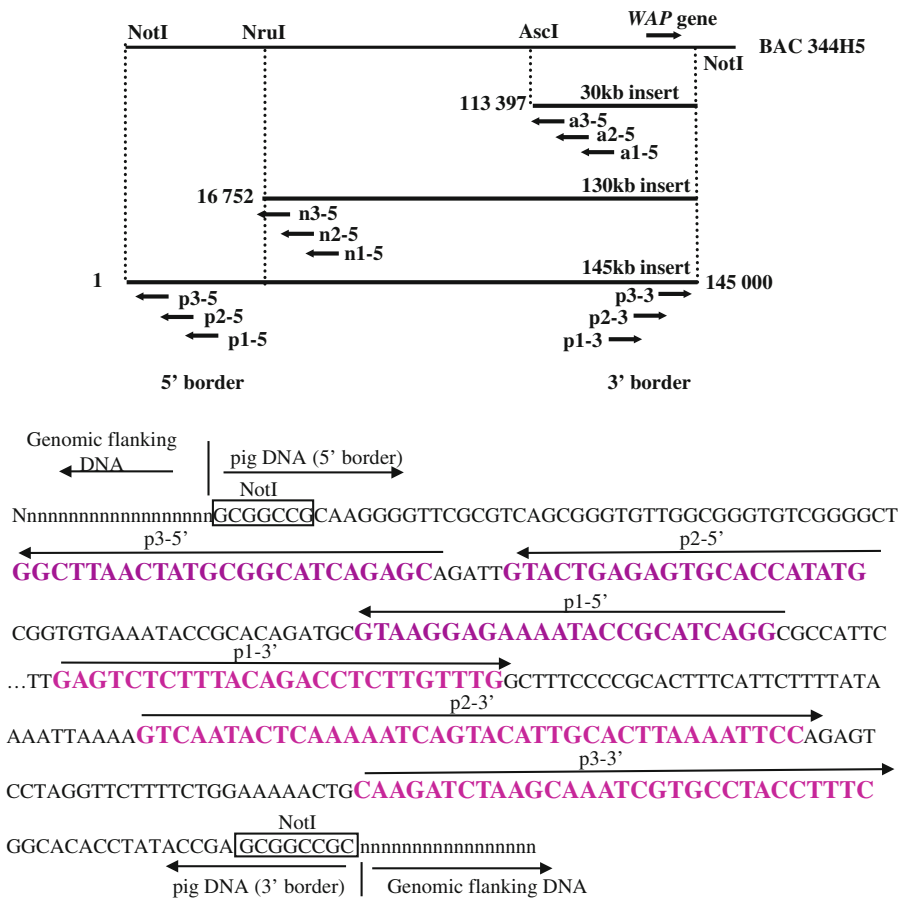
Transgenic animals are frequently generated by random integration of foreign DNA. Except for the constructs based on retrovirus, lentivirus and transposons, the foreign DNA fragment often integrate into the host genome as tandem repeats encompassing several copies of the foreign gene. Transgene expression can be affected by the host genome at the insertion site but also by the number of integrated copies. Numerous studies have focused on the use of insulators and regulatory elements to prevent transgene silencing as well as to obtain a reliable expression independent of the integration site and correlated with the number of integrated copies (Giraldo et al. 2003; Houdebine 2007, 2009; Montoliu et al. 2009).

In our previous experiments (Saidi et al. 2007), we generated several transgenic mouse lines by microinjection of a series of large (30–145 kb, Fig. 1) porcine genomic DNA fragments into fertilized oocytes. By using large DNA fragments, we expected a reliable expression of the pig Whey Acidic Protein (pWAP) gene, one of the genes present in the pig genomic DNA insert. The pWAP gene expression varied more than anticipated within the different transgenic mice lines for each DNA insert. This effect was unrelated to the number of integrated copy. We thus decided to investigate whether this variability could be explained by the structure of the integrated DNA at the insertion site in the host genome.

---

A. Le Saux · L.-M. Houdebine · G. Jolivet (✉)  
Biologie du Développement et Reproduction,  
UMR1198, INRA ENVA, 78352 Jouy-en-Josas,  
France  
e-mail: genevieve.jolivet@jouy.inra.fr

**Fig. 1** location of specific TAIL-PCR primers in the various pig DNA inserts used to generate transgenic mice. Three pig DNA fragments (30, 130, 145 kb) were recovered from the BAC 344H5 by digestion with restriction enzymes (NotI, NruI or AscI) as previously described (Saidi et al. 2007). The position of the 5'P and 3'OH borders of the inserts is given relatively to the transcription start site of the pig WAP gene. All the inserts harboured the same 3'OH border. Thus, the same set of primers (p1-3, p2-3, p3-3) was used for the study of the 3'OH border flanking regions. The lower panel presents the sets of primers used for the 145 kb insert. All other sequences are available upon request. Numbers indicate the position of the nucleotides at the border of the insert. Numbering begins at the 5'P border of the 145 kb fragment



To reach this goal, we attempted to identify the transgene insertion site within the mouse genome. Several protocols, which were not similarly successful, were implemented. The TAIL-PCR method (thermal asymmetric interlaced PCR, Liu et al. 1995; Liu and Chen 2007; Tadege et al. 2008) allowed us (1) to analyze transgene/genomic borders and internal concatamer junctions for eleven transgenic lines, (2) to obtain sequence information for seven borders, (3) to place three transgenes in the mouse genome, and (4) to obtain sequence data for seven transgene junctions in concatamers.

**Materials and methods**

**Production of transgenic mouse lines**

Transgenic mouse lines (C57-Bl6/CBA F1) were obtained by microinjection of long pig DNA fragments

(Fig. 1) as presented in our previous paper (Saidi et al. 2007). Transgenic mice were heterozygous for the integrated transgene. All the procedures for the preparation of large DNA inserts from BACs respected the recommendations described by VanKeuren et al. (2009). Large DNA inserts (from 30 to 145 kb) were prepared by digestion of BAC DNA using specific restriction enzymes, followed by pulse field gel electrophoresis (FIGE) on 1% agarose gel in 0.5× TBE buffer. Recovery of DNA from gel was achieved by gelase digestion (Tebu-bio, Le Perray en Yvelines, France) or electroelution (BioTrap, Whatman Schleicher & Schüell, Versailles, France). The product was submitted to dialysis against microinjection buffer (Tris pH 7.5 10 mM; EDTA 0.1 mM; NaCl 100 mM; spermine 30 µM; spermidine 70 µM) and used for microinjection within 2 weeks. Integrity of the DNA was assessed by FIGE. Microinjection was performed when the expected DNA fragment size was observed in gel without any detectable DNA fragmentation.

### Extraction of genomic DNA from tail tip

Genomic DNA was prepared from tail tips by proteinase K digestion. Biopsies were incubated overnight at 37°C or 2 h at 56°C in 500 µl of proteinase K extraction buffer (Tris pH 8 5 mM; EDTA 5 mM; 1% SDS, sodium acetate 300 mM; 100 µg proteinase K extemporaneously added), then treated with 1 volume of phenol:chloroform:isoamyl alcohol. DNA was recovered after isopropanol precipitation (0.6 volume).

### Thermal asymmetric interlaced PCR (TAIL-PCR)

The protocol was adapted from experiments carried out with transgenic plants (Liu et al. 1995; Tadege et al. 2008) using tail tip DNA as template (Saidi et al. 2007) and sets of primers (Table 1). Primary TAIL-PCR reactions (20 µl) contained 25 ng of genomic DNA, 1× PCR buffer, 200 µM dNTP mix, 300 nM specific primer 1, and an arbitrary degenerated ADn primer (3 µM) with 1.5 units of Taq polymerase. Significant results were obtained with high quality Taq polymerase (Takara ExTaq, Lonza, Courtaboeuf, France) and good quality genomic DNA preparations. The sequences of five distinct degenerate primers (AD1, AD2, AD3, AD5, AD6) were those already designed by other (Liu et al. 1995; Tadege et al. 2008). These sequences are given in Table 1. Specific primers were designed according to the expected position of the borders of the microinjected DNA fragment. For each microinjected DNA fragment, several sets of primers matching more or less internal positions were tested to anticipate the possibility of deletions of the extremities of the transgene (not shown). Sets of primers giving the best results are given in Fig. 1. Sequences of specific primers are available upon request. Each secondary

**Table 1** Sequence of the arbitrary degenerate primers

AD1	5'NTCGA(G/C)T(A/T)T(G/C)G(A/T)GTT 3'
AD2	5'NGTCGA(G/C)(A/T)GANA(A/T)GAA3'
AD3	5'(A/T)GTGNAG(A/T)ANCANAGA3'
AD5	5'(G/C)(G/C)TGG(G/C)STANAT(A/T)AT(A/T)CT3'
AD6	5'CG(G/C)AT(G/C)TC(G/C)AANAA(A/T)AT3'

The sequences are those given elsewhere (Liu et al. 1995; Tadege et al. 2008)

20 µl TAIL-PCR was performed with the specific primer 2, the same degenerate AD primer than in the primary reaction and 2 µl of the 50-fold diluted PCR product of the primary reaction (final dilution = 500). Each tertiary TAIL-PCR was performed with the specific primer 3, the same degenerate AD primer than in the secondary reaction and the 50-fold diluted PCR product of the secondary reaction. The concentrations of all components of the PCR reactions were similar in primary, secondary and tertiary TAIL-PCR. The thermal conditions for TAIL-PCR are given in Table 2. The final products of the third amplification were analyzed using agarose 1% TBE 1× gels. The DNA fragments longer than 200 base pairs were isolated by column purification from agarose gel (GenElute, Sigma) or directly from the tertiary PCR reaction (MSB Spin PCRapace, Invitex, Berlin, Germany). Sequencing was performed (MWG Eurofins service, Ebersberg, Germany) either directly on the product of purification, or after cloning in pGEM-T Easy cloning vector (Promega, Courtaboeuf, France). The specific primer used in the third round of PCR was used as sequencing primer.

To confirm data, PCR amplification was performed using primers specific of the pig DNA insert and of the identified flanking region using the transgenic genomic DNA as template. Confirmation was achieved when only one DNA fragment was amplified with the expected size and sequence.

**Table 2** TAIL-PCR programs

TAIL-PCR1	93°C 1 min, 1 time
	95°C 1 min, 1 time
	94°C 45 s, 62°C 1 min, 72°C 2.5 min, 4 times
	94°C 45 s, 25°C 3 min, ramp to 72°C in 3 min, 72°C 2.5 min, 1 time
	94°C 20 s, 68°C 1 min, 72°C 2.5 min, 94°C 20 s, 68°C 1 min, 72°C 2.5 min, 94°C 20 s, 44°C 1 min, 72°C 2.5 min 14 times
TAIL-PCR2	72°C 5 min, 1 time
	94°C 20 s, 64°C 1 min, 72°C 2.5 min, 94°C 20 s, 64°C 1 min, 72°C 2.5 min
	94°C 20 s, 44°C 1 min, 72°C 2.5 min, 12 times
TAIL-PCR3	72°C 5 min, 1 time
	94°C 20 s, 44°C 1 min, 72°C 2.5 min, 20 times
	72°C, 5 min

These programs are similar to those given by Liu et al. (1995)

## Sequence analysis

Homology search was performed by BLAT analysis using the Ensembl web site (<http://www.ensembl.org/index.html>). The position of integration site within the mouse genome was deduced from the right and left transgene flanking sequences.

## Evaluation of transgene copy number

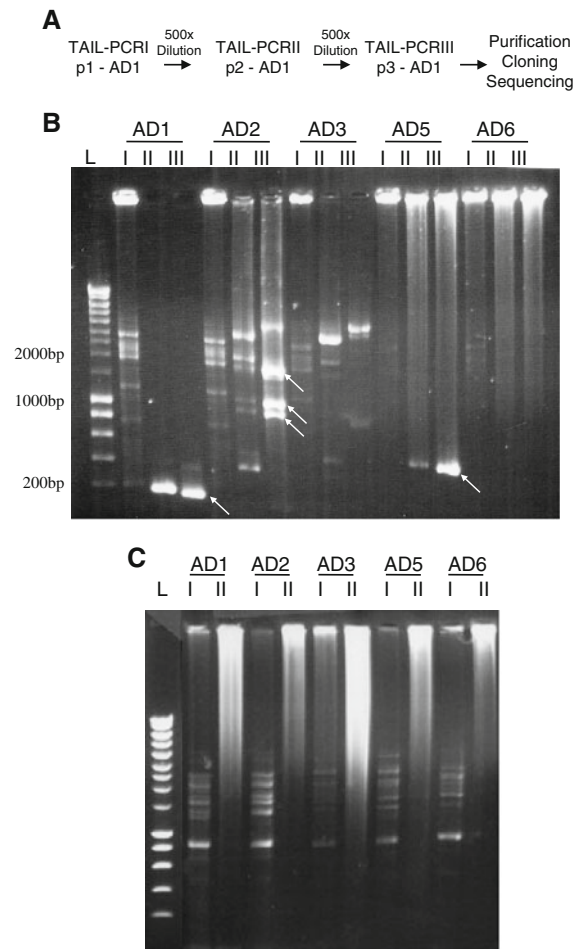
The integrated copy number was evaluated as previously (Saidi et al. 2007) by real time PCR using the  $2^{-\Delta\Delta C_t}$  method with the purified BAC 344H5 as standard. The mouse  $\beta 2$ -microglobulin gene was used as reference gene. For each line, the copy number was determined using different sets of primers specific for the pig DNA injected insert: one set located at the 5'P border, one at the 3'OH border and one on the pig WAP gene. All sequences are available upon request.

## Results

The principle of the TAIL-PCR method consists of three rounds of nested hemispecific PCR amplifications (Fig. 2a). Each reaction includes a long-high  $T_m$  primer complementary to the border of the transgene and a degenerate-low  $T_m$  primer. Non-specific products were eliminated by two successive PCRs using nested primers (Fig. 2b, compare lanes I, II, and III). The TAIL-PCR results in the amplification of fragments overlapping the border of the transgene and the adjacent host genome (Liu and Whittier 1995; Tadege et al. 2008).

To anticipate the possible deletions of the transgene borders, several sets of primers, located more than 150 base-pairs (bp) upstream of the theoretical border were tested. The length of the recovered fragments in the TAIL-PCR experiments ranged between 200 and 2,000 bp as it is commonly reported for this method. However, internal sets of primers gave no better amplification than the sets of primers located near the border.

Eleven transgenic lines were analyzed. A typical analysis of the amplified fragments on gel electrophoresis is given in Fig. 2. Analyses of the sequences are summarized in Table 3 and Figs. 3 and 4. The position of the transgene within the genome was



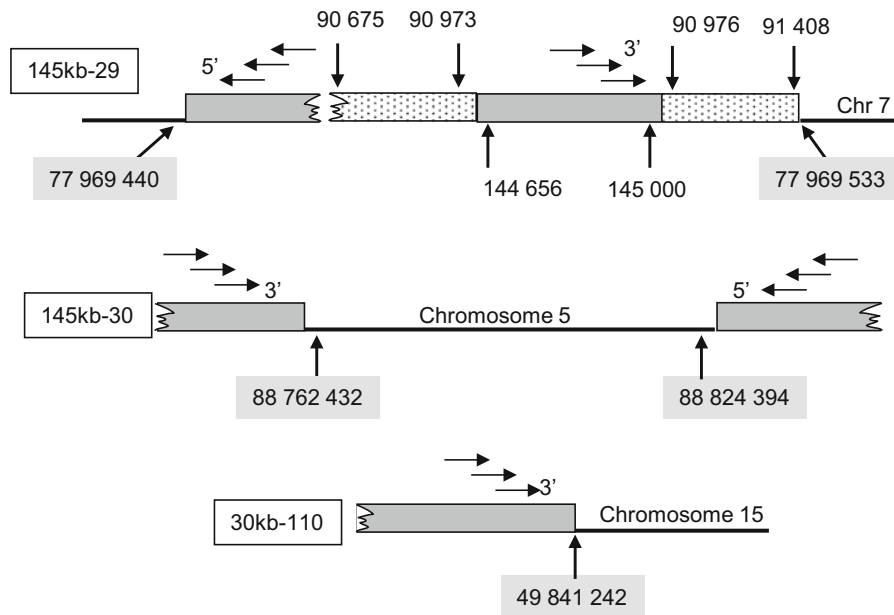
**Fig. 2** Analysis on agarose gel electrophoresis of the products of the TAIL-PCR. **a** Different steps of the TAIL-PCR. **b** Typical electrophoresis analysis of the products amplified after the different rounds of PCR in the line 130 kb-535. Amplified DNA fragments showing a decreasing size in the successive PCR amplifications were supposed to be specific for the border of the transgene. The identity of the best set of primers (specific and degenerated) varied with the transgenic line. A few numbers of specific fragments were observed (labelled by *white arrows*). These DNA fragments were recovered from gel and sequenced. **c** Electrophoresis of amplified products after two rounds of PCR in the line 145 kb-28. Obviously, no specific fragment was amplified after the second PCR. The numbers I, II, and III refer to the number of the PCR.  $AD_n$  arbitrary degenerate primer  $n$ ,  $L$  kb ladder

successfully identified in three lines only (145 kb-29, 145 kb-30, 30 kb-110, Fig. 3). In two of these lines (lines 145 kb-29 and 145 kb-30), the 3'OH and the 5'P borders were flanked by regions that belong both to the same mouse genome locus (respectively in chromosome 7 D1 band, and chromosome 5 E1

**Table 3** Summary of the analysis of genomic DNA sequences flanking pig DNA transgenes

Microinjected insert	Transgenic line	Copy number	5' end of the microinjected insert		Identified 5' border	Identified 3' border	Insert concatenation
			pWAP gene	3' end of the microinjected insert			
145 kb	28	5–10	5–9	7–10	ND	ND	ND
	29	2–3	2–3	2–3	chr7	chr7	Multiple (see Fig. 3)
	30	1	1	1	77 969 440 (-) chr5	77 969 533 (+) chr5	ND
	809	3–4	3–4	2–3	88 844 394 (-) ND	88 762 432 (+) ND	5' border linked to 34862–36017
	815	7–11	7–8	9–10	ND	ND	5' to 3' junction
	820	2	2	2	ND	ND	5' to 3' junction
	863	5–6	4–8	6–8	ND	ND	ND
	19	4–5	4–5	4–6	Unknown sequence	ND	3' border linked to 56666–56856
	23	2–3	2–3	2–3	ND	ND	3' border linked to 41252–41045
	535	13–14	7–12	6–10	ND	ND	5' border linked to 46000–45602
30 kb	110	1–2	1–2	1–2	ND	chr15	ND
						49 841 242 (+)	

Three different pig DNA inserts were microinjected in mouse unicellular embryos (145, 130 and 30 kb) as presented in Fig. 1. The name of each analysed mouse line is given. Three sets of primers were designed at the 5'P end of the microinjected insert, the 3'OH end and within the pig WAP gene. Numbers of integrated copies were obtained for each set of primers using a qPCR assay with a reference mouse gene ( $\beta$ 2-microglobulin gene (Saïdi et al. 2007)). A range of copy number estimated from 2 to 8 mice is given for each mouse line. The “insert concatenation” column summarizes the nature of the flanking regions; numbers indicate the position of the flanking sequence (mouse chromosome or pig DNA) as represented in Fig. 3. ND the determination was attempted but failed. (-) or (+) refers to the orientation of the pig DNA insert relatively to the sense (+) or the antisense (-) strand of the mouse chromosome



**Fig. 3** Detailed description of the transgene borders and their flanking sequences in lines 145 kb-29, 145 kb-30, and 30 kb-110. The 5'P or 3'OH border of the transgene is represented by a grey box, with the set of specific primers that were used for the characterization (horizontal black arrows). Boxes filled with dots represent internal pig DNA fragments. Numbers

indicate their position in the 145 kb pig DNA insert. Number 1 corresponds to the 5'P border, and 145,000 to the 3'OH border. Numbers in grey boxes indicate the position of the mouse chromosome integration site as given by the Ensembl mouse database

Head-to-tail junctions		Border-internal junctions	
	pig DNA 3'OH border    NotI    pig DNA 5'P border Expected ...GCACACCTATACCGA GCGGCCG AAGGGGTTTCGC... 145kb-820 ...GCACACCTATACCGA GCGGCCG AAGGGGTTTCGC... 145kb-815 ...GCACACCTATACCGA GCGGC::GC AAGGGGTTTCGC...		NotI    pig DNA 5'P border Expected ...TGCTGTGGGAACG:::GGTTCGC... 145kb-809 ...TGCTGTGGGAACG:::GGTTCGC... 34 862
	pig DNA 3'OH border    NotI    NruI    pig DNA 5'P border Expected ...GCACACCTATACCGA GCGGCC CGA GATGGATCCTA... 130kb-23 ...GCACACCTATACCGA GCGGCC CGA GATGGATCCTA...		pig DNA 3'OH border    NotI Expected ...GCACACCTATACCGA GCGGC::GTC::AACCCTCGTG... 130kb-19 ...GCACACCTATACCGA GCGGC::GTC::AACCCTCGTG... 56 666 130kb-23 ...GCACACCTATACCGA GCGGC:::CTATTGCTCAGCCT... 41 252
	NruI    pig DNA 5'P border Expected ...TGAACCCGACTAGT:::CGA GATGGATCCTA... 130kb-535 ...TGAACCCGACTAGT:::CGA GATGGATCCTA... 46 000		

**Fig. 4** Sequence of the junctions between tandemly integrated copies of the transgene. The first line represents the expected sequence, reconstituted by the fusion of restricted DNA inserts. Other lines underneath indicate the actually determined sequence. Dotted lines represent missing nucleotides in the

expected sequence. The numbers indicate the position of the first nucleotide of the linked region. Arrows below each sequence indicate the orientation relative to the pig WAP gene in the insert

band), which confirms the transgene position in the mouse genome. In the line 145 kb-29, the integration of the transgene was coincident with a 93 bp deletion in the mouse chromosome 7 harbouring the transgene, without any modification in the second allele (not shown). In the line 145 kb-30, the 3'OH and 5'P borders were located in chromosome 5 in a head-to-tail orientation but spaced out by more than 60 kb (Fig. 3). In this line, only one copy of the transgene

was integrated. We can hypothesize that the transgene was split into two fragments integrated in a reverse orientation. We could not confirm this hypothesis, the 60 kb-long region intercalated between the two hypothetical fragments being too short for the detection of two distinctly labelled points in FISH experiments. Alternatively, in the mouse used for producing this transgenic line, the sequence of the locus in chromosome 5 was in an



inversed orientation comparatively to the published mouse genome. No further identification of such chromosome reorganization was undertaken in this line. In the line 30 kb-110, the 3'OH border of the transgene was flanked by a region that belongs to the mouse chromosome 15 (B3.3 band). We were unable to characterize the region flanking the 5'P border of the transgene, probably because the 5'P border of the 30 kb DNA insert is characterized by a GC rich sequence spanning over more than 1 kb which may reduce the chance of performing specific PCRs successfully.

In the other lines, the position of the transgene in the genome was not identified. In some of the lines (145 kb-28 and 145 kb-863) no specific DNA bands were observed using electrophoresis (Fig. 2c). In the line 130 kb-19, the region flanking the 3'OH border of the transgene was not assigned to any location in the mouse genome. One hypothesis is that the transgene was inserted in repeated-sequences, preventing its positioning in the mouse genome.

Surprisingly, the junctions between the DNA inserts could not be characterized in all cases. Expected 5'P to 3'OH junctions (head-to-tail) were found with almost the exact sequence in three lines only (145 kb-815; 145 kb-820; 130 kb-23, Fig. 4) but not in the other lines harbouring several copies of the transgene. In line 145 kb-820 only, the NotI site was reconstituted. Moreover, we found that internal regions of the foreign DNA insert were linked to the 5'P or the 3'OH borders of the transgene (lines 145 kb-29; 145 kb-809; 130 kb-19; 130 kb-23; 130 kb-535) as indicated in Table 3 and Figs. 3 and 4. Thus, rearrangements of the DNA insert occurred, with multiple recombination events before or at the time of integration into the host genome. Likely, these rearrangements can be an obstacle to the characterization by TAIL-PCR of genome-transgene or copy-copy junctions in most lines.

We then tried to evaluate the integrity of the inner part of the transgene by amplification using specific primers. This led us to identify putative rearrangements in the insert such as those presented in the line 145 kb-29 (Fig. 3). In this line, a 344 bp long DNA fragment encompassing the expected 3'OH border of the transgene was inserted in the 90 973–90 976 region of the pig DNA insert. The actual border of the transgene (chromosome 7-transgene junction) was 432 bp farther, corresponding to position 91 408 of

the pig DNA insert. Rearrangements are therefore not specific of one or the other microinjected DNA fragment. On the contrary, they are observed in several mouse transgenic lines whatever the length of the microinjected pig DNA insert. To summarize, our TAIL-PCR analysis shows evidence of unexpected rearrangements in five transgenic lines (145 kb-29, 145 kb-809, 130 kb-19, 130 kb-23, 130 kb-535), insert concatenation with minor or no deletion in three lines (145 kb-815, 145 kb-820, 130 kb-23) while in two lines no data could be generated (line 145 kb-28, 145 kb-863).

## Discussion

The characterization of genome integration sites is mainly based upon the search for host genome sequences adjacent to each extremity of the transgene. Several commonly employed methods consist of implementing a set of multiple PCR using primers anchored on the theoretical extremities of the transgene. These methods have the advantage to be rapid and easy to perform. However, deletions of the extremities of the microinjected DNA fragment, as already quoted by previous authors (Chen et al. 1995), are an obstacle to the success of these techniques. Nevertheless, among the various methods that we tested (not shown), the TAIL-PCR gave us the best results.

TAIL-PCR allowed us to characterize mouse genome transgene integration sites in several but not all the transgenic lines. Other techniques such as a recently published refined and improved TAIL-PCR technique (Liu and Chen 2007) should help us to identify more easily the integration sites. Besides, in our hands, the inverse PCR based technique (Ochman et al. 1988) was poorly efficient. In order to improve the efficiency of inverse PCR, it has been suggested by others to digest the transgenic mouse DNA with the enzyme that was used for the preparation of the insert to be injected (Liang et al. 2008). Theoretically, this procedure has the advantage to eliminate head-to-tail, head-to-head or tail-to-tail junctions. In lines issuing from the microinjection of the 145 kb fragment, the NotI enzyme was used. In the line 145 kb-820, the NotI site was found between the two copies of the transgene (Fig. 4). However, in the

present work, digestion of the DNA with NotI did not eliminate the tandemly integrated copies (not shown). One hypothesis is that the NotI enzyme action was hampered by DNA methylation that can occur after insertion in the host genome. DNA methylation analysis should be performed to test this hypothesis.

The most surprising finding of our work concerns the complex rearrangements between regions of the pig DNA insert. More, the frequency of these rearrangements was unexpectedly high, since it concerned five lines over the eleven analyzed ones. One hypothesis is that, in the present study, the long DNA insert was fragmented when microinjection was performed, despite the care taken to avoid this artefact. However, fragmentation due to the mechanical shearing of long DNA fragment is probably not the only cause of the rearrangements. It has been postulated that random DNA integration into the host genome occurs through DNA double-strand break repair by illegitimate integration via micro-homologies or non-homologous end joining (NHEJ) (Bishop 1996; Würtele et al. 2003; Kamisugi et al. 2006). Rearrangements were observed as well when a few kilobases long DNA inserts were microinjected in early mouse embryos (Chen et al. 1995; Pawlik et al. 1995). Several works were designed to know whether this phenomenon can be influenced by the sequence of the injected DNA insert. The addition of SINEs elements (short interspersed elements) at both side of foreign DNA increased the efficiency of integration in mouse embryo genome, and this was probably due to the enhancement of homologous recombination (Kang et al. 1999). In the present work, the long pig DNA fragments encompass series of repetitive sequence regions such as SINEs or LINEs. Our hypothesis is that the presence of these elements along the foreign DNA favoured the multiple rearrangements at the integration site or even before integration. The occurrence of such rearrangements can hardly be anticipated. Due to the length of the pig DNA inserts, we can suspect that other internal rearrangements occurred. Thus, the internal structure of the transgene should be investigated through complementary methods. The Southern blot technique could be helpful for this determination. However, this technique is tedious and strenuous especially in the case of hundreds of kilobases-pair long DNA inserts. Moreover, it does not allow the identification of small size rearrangements as those we observed at the 3'OH border of the transgene in the line 145 kb-29.

It is more and more admitted that microinjection of long DNA fragment is a valuable technique to generate transgenic animals expressing in a reliable manner the genes which are included, because it is expected that important regulatory regions are present (Giraldo et al. 2003; VanKeuren et al. 2009). Previous works have shown that deletions occurred frequently in large transgenes (Chandler et al. 2007). We also reported previously this event and we eliminated from our study all the transgenic lines harbouring deleted copies. More complex rearrangements involving large BAC transgenes have already been reported (Abrahams et al. 2003). The TAIL-PCR method allowed us to characterize such rearrangements. Notably, the present work addresses the difficulty of cloning integration sites of BAC transgenes. In the literature, the difficulties and the failures to characterize integration sites are poorly documented. Indeed, previously published papers report successful cloning of integration sites (Liang et al. 2008; Mehta et al. 2009). The difference between our experiments and the previous ones could result from the multiple rearrangements of our transgenes that generate unexpected borders. Moreover, we noticed that the successful reported experiments are frequently based upon transgenesis of transposons, or retrovirus that integrate the host genome with minor rearrangements in opposite of BAC transgenes. Finally, the present paper stresses the frequency of such complex rearrangements. This should be carefully considered when transgenic animals are produced with large genomic DNA fragments.

**Acknowledgments** This work was supported by the ANR program ANR-06-POGM-008\_TRANSINTEX. We thank P. Ratet (Institut des Sciences du Végétal, CNRS, Gif-sur-Yvette, France) for his fruitful advices concerning the use of TAIL-PCR method and the members of the mouse facility (UEAR, Jouy-en-Josas, France) for the maintenance of animals.

## References

- Abrahams BS, Chong ACO, Nisha M, Milette D, Brewster DA, Berry ML, Muratkhodjaev F, Mai S, Rajcan-Separovic E, Simpson EM (2003) Metaphase FISHing of transgenic mice recommended: FISH and SKY define BAC-mediated balanced translocation. *Genesis* 36:134–141
- Bishop J (1996) Chromosomal insertion of foreign DNA. *Reprod Nutr Dev* 36:607–618



- Chandler KJ, Chandler RL, Broeckelmann EM, Hou Y, Sout-hard-Smith EM, Mortlock DP (2007) Relevance of BAC transgene copy number in mice: transgene copy number variation across multiple transgenic lines and correlations with transgene integrity and expression. *Mamm Genome* 18:693–708
- Chen C-M, Choo K-B, Cheng W (1995) Frequent deletions and sequence aberrations at the transgene junctions of transgenic mice carrying the papillomavirus regulatory and the SV40 TAg gene sequences. *Transgenic Res* 4:52–59
- Giraldo P, Rival-Gervier S, Houdebine L-M, Montoliu L (2003) The potential benefits of insulators on heterologous constructs in transgenic animals. *Transgenic Res* 12: 751–755
- Houdebine LM (2007) Transgenic animal models in biomedical research. *Methods Mol Biol* 360:163–202
- Houdebine LM (2009) Design of expression cassettes for the generation of transgenic animals (including insulators). In: I Anegon (ed) *Rat genomics: methods and protocols. Methods in molecular biology*, vol 597. doi: [10.1007/978-1-60327-389-3\\_4](https://doi.org/10.1007/978-1-60327-389-3_4), Humana Press, Totowa, LLC2010
- Kamisugi Y, Schlink K, Rensing SA, Schween G, von Stackelberg M, Cuming AC, Reski R, Cove DJ (2006) The mechanism of gene targeting in *Physcomitrella patens*: homologous recombination, concatenation and multiple integration. *Nucleic Acids Res* 34:6205–6214
- Kang Y-K, Park JS, Lee C-S, Yeom YI, Chung A-S, Lee K-K (1999) Efficient integration of short interspersed element-flanked foreign DNA via homologous recombination. *J Biol Chem* 274:36585–36591
- Liang Z, Breman A, Grimes B, Rosen E (2008) Identifying and genotyping transgene integration loci. *Transgenic Res* 17:979–983
- Liu Y-G, Chen Y (2007) High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *BioTechniques* 43:649–656
- Liu Y, Whittier R (1995) Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 25:674–681
- Liu Y-G, Mitsukawa N, Oosumi T, Whittier R (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J* 8:457–463
- Mehta AK, Majumdar SS, Alam P, Gulati N, Brahmachari V (2009) Epigenetic regulation of cytomegalovirus major immediate-early promoter activity in transgenic mice. *Gene* 428:20–24
- Montoliu L, Roy R, Regales L, García-Díaz N (2009) Design of vectors for transgene expression: The use of genomic comparative approaches. *Comp Immunol Microbiol Infect Dis* 32:81–90
- Ochman H, Gerber AS, Hartl DL (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* 120: 621–623
- Pawlik K, Sun C-W, Higgins P, Townes T (1995) End joining of genomic DNA and transgene DNA in fertilized mouse eggs. *Gene* 165:173–181
- Saidi S, Rival-Gervier S, Daniel-Carlier N, Thépot D, Morgenthaler C, Viglietta C, Prince S, Passet B, Houdebine L, Jolivet G (2007) Distal control of the pig whey acidic protein (WAP) locus in transgenic mice. *Gene* 401: 97–107
- Tadege M, Wen J, He J, Tu H, Kwak Y, Eschstruth A, Cayrel A, Endre G, Zhao P, Chabaud M, Ratet P, Kirankumar S (2008) Large-scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume *Medicago truncatula*. *Plant J* 54:335–347
- VanKeuren M, Gavrilina G, Filipiak W, Zeidler M, Saunders T (2009) Generating transgenic mice from bacterial artificial chromosomes: transgenesis efficiency, integration and expression outcomes *Transgenic Res* (in press)
- Würtele H, Little K, Chartrand P (2003) Illegitimate DNA integration in mammalian cells. *Gene Ther* 10:1791–1799