1296–1302 *Nucleic Acids Research, 1997, Vol. 25, No. 6*

Evaluation of β-globin gene therapy constructs in single copy transgenic mice

James Ellis*, Peter Pasceri, Kian C. Tan-Un[1,+], Xiumei Wu, Alex Harper[1,§], Peter Fraser[1,¶] and Frank Grosveld[1,¶]

Department of Genetics, Blood Gene Therapy Program and Program in Developmental Biology, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G1X8, Canada and Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, Canada and [1]Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, London NW7 1AA, UK

Received August 30, 1996; Revised and Accepted January 16, 1997

ABSTRACT

Effective gene therapy constructs based on retrovirus or adeno-associated virus vectors will require regulatory elements that direct expression of genes transduced at single copy. Most β-globin constructs designed for therapy of β-thalassemias are regulated by the 5′HS2 component of the locus control region (LCR). Here we show that a human β-globin gene flanked by two small 5′HS2 core elements or flanked by a 5′HS3 (footprints 1–3) core and a 5′HS2 core are not reproducibly expressed in single copy transgenic mice. In addition, low copy transgene concatamers that contain only dimer 5′HS2 cores fail to express, whereas those that contain monomer 5′HS2 cores express at 14% per copy. These data suggest that spacing between HS cores is crucial for LCR activity. We therefore constructed a novel 3.0 kb LCR cassette in which the 5′HS2, 5′HS3 and 5′HS4 cores are each separated by ~700 bp. When linked to the 815 bp β-globin promoter this LCR directs 45% levels of expression from four independent single copy transgenes. However, the 3.0 kb LCR linked to the 265 bp promoter expresses variable levels, averaging 18%, from three single copy transgenes. Our findings suggest that sequences in the distal promoter play a role in single copy transgene activation and that larger LCR and promoter elements are most suitable for gene therapy applications.

INTRODUCTION

One difficult aspect of gene therapy is in reproducibly obtaining high level, tissue-specific and long term expression from single copy genes transferred into stem cells (1,2). The locus control region (LCR) confers such position-independent activation on single copy β-globin transgenes and also directs transgene concatamers to express in a copy number-dependent manner (3–6). The complete LCR is composed of four DNase I hypersensitive sites on a 20 kb fragment (7,8) that is too large to be incorporated into retrovirus or adeno-associated virus (AAV) vectors designed for therapy of β-thalassemias. Therefore, most groups have concentrated on the 200–300 bp core elements of the individual hypersensitive sites and in particular the 5′HS2 core element (9–11) for regulating transduced globin genes (12–15). Although it has proved difficult to obtain stable high titer viruses bearing these sequences (16–20), high titer virus can be produced using the related α-globin HS-40 element (21). Nevertheless, it has yet to be demonstrated that single copy integration of these vectors in human hematopoietic stem cells will lead to consistently high levels of globin gene expression in differentiated erythroid cells.

In order to test expression levels from a potential gene therapy construct, the construct must be inserted into a viral vector, transferred into a helper cell line to assemble stable high titer virus and then successfully transduced into mouse hematopoietic stem cells (22). Long term repopulation of mice by the infected stem cells finally produces mature differentiated cells that can be assessed for expression of the transduced gene. Ideally, β-globin viral vectors can be tested in NOD/SCID mice that have been reconstituted with transduced bone marrow cells from β-thalassemia patients (23). An alternative method is to test expression from a potential gene therapy construct present in transgenic mice at a single copy before the difficulties of assembling a high titer virus are confronted. In the case of β-globin vectors, expression can be evaluated in erythroid cells derived from stem cells that contain the construct as a single copy transgene and only those β-globin constructs that express appropriately and reproducibly need be packaged into virus for further studies.

Our previous work demonstrated that a 215 bp synthetic 5′HS2 core fails to activate expression from a linked β-globin gene in single copy founder transgenic mice and therefore is not ideal for gene therapy constructs (24,25). However, these studies also showed that 5′HS2 elements in multicopy head-to-tail transgene concatamers directed significant β-globin expression, suggesting that two or more 5′HS2 cores interact and cooperate with each other to open chromatin and enhance transcription. It may

*To whom correspondence should be addressed at: Department of Genetics, Hospital for Sick Children, 555 University Avenue, Torronto, Ontario M5G1X8, Canada. Tel: +1 416 813 7295; Fax: +1 416 813 4989; Email: jellis@sickkids.on.ca

Present addresses: [+]School of Professional and Continuing Education, Knowles Building, University of Hong Kong, Hong Kong, [§]Transgenic Research Unit, The Rayne Institute, St Thomas's Hospital, London SE1 7EH, UK and [¶]Department of Cell Biology and Genetics, Erasmus University, Rotterdam, 3000 DR Rotterdam, The Netherlands

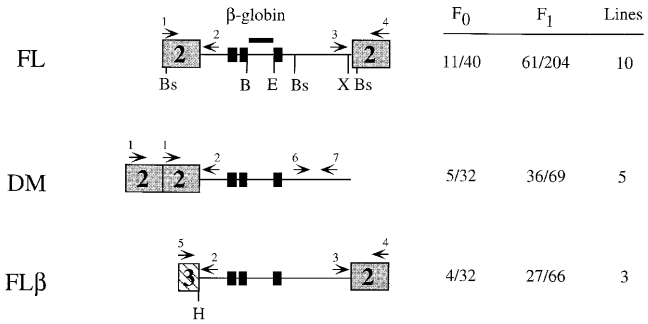

Figure 1. Map of the transgene constructs (not drawn to scale) and ratios of transgenic/total numbers of mice generated by microinjection (F_0) or breeding (F_1). Constructs FL, DM and FLβ contain synthetic 5′HS2 cores cloned from the 4m.1 construct (24), shown as grey boxes. FLβ also contains the 5′HS3 (footprints 1–3) core from FP1-3 (26), shown as a striped box. In all cases the hypersensitive sites were cloned into the polylinker 5′ of the 815 bp human β-globin promoter or into the *Eco*RV site 3′ of the human β-globin gene (exons shown in thick black boxes) in GSE 1758. The narrow black box is the βivs2 probe. Numbered arrows correspond to PCR primers (see Materials and Methods). Bs, *Bsp*HI; B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; X, *Xba*I.

therefore be possible to mimic this activity by linking multiple 5′HS2 cores to a β-globin gene to make a construct that would be functional at single copy. Here, we evaluate expression of various β-globin constructs containing multiple HS elements present in mice as single copy transgenes.

MATERIALS AND METHODS

Plasmid construction

The 3.0 kb LCR in pBGT14 was constructed by sequentially inserting the various HS into the polylinker of pGSE1758 which contains the 4.2 kb human β-globin gene regulated by the 815 bp promoter. The 1.15 kb *Stu*I–*Spe*I fragment of 5′HS4 was cloned into the *Cla*I site, the 0.85 kb *Sac*I–*Pvu*II fragment of 5′HS3 was cloned into the *Xho*I site and the 0.95 kb *Sma*I–*Stu*I fragment of 5′HS2 was cloned into the *Spe*I site. All new restriction sites were created by blunting overhanging ends with T4 DNA polymerase and ligation of appropriate linkers (New England Biolabs). The orientation of each HS with respect to the β-globin transgene is the same as in the wild-type locus. The 265 bp promoter in pBGT15 was cloned by digestion at the *Sna*bI site and addition of *Sal*I linkers.

Transgenic mice

Transgenic mice (FVB or CBA/C57BL strains) were generated by microinjection of 0.25–0.50 ng/µl purified DNA as described (24) and screened by Southern blot hybridization and PCR on tail or fetal head DNA by standard procedures. The sequences of PCR primers (see Fig. 1) were as follows: (1) 5′-AAGCACAGCAA-TGCTGAGTCATG-3′; (2) 5′-TCAATGGGGTAATCAGTGG-TGTC-3′; (3) 5′-GGGTGGGAGAATCAGGAAACTAT-3′; (4) 5′-TGTCACATTCTGTCTCAGGCATC-3′; (5) 5′-TGCCAGA-TGTGTCTATCAGAGGT-3′; (6) 5′-CATGGTTTGACTGTCC-TGTGAGC-3′; (7) 5′-GGTGGTTGATGGTAACACTATGC-3′.

DNA analysis

Southern transfer and hybridization were by standard procedures. Copy number determination was performed using a Molecular Dynamics PhosphorImager and adjusted for loading differences or the presence of non-intact transgenes (observed after digestion with *Bsp*HI or other diagnostic restriction enzymes). Mosaicism in the fetal liver of founder 15.5 day transgenic mice was calculated by quantifying the intensity of bands on a Molecular Dynamics PhosphorImager and using the formula (Tg Hβ/Tg mThy-1)/(B26 Hβ × Tg copy number/B26 mThy-1), where Tg is transgenic, Hβ is human β-globin, mThy-1 is mouse Thy-1 and B26 is one copy bred line B26.

RNA analysis

Fetal liver RNA (13.5–15.5 day) was extracted, 1 µg was hybridized to kinased double-stranded DNA probes, digested with 75 U S1 nuclease and run on a 6% sequencing gel as described (24). Probe excess was demonstrated by including a sample containing 3 µg fetal liver RNA. Specific activities of human β-globin (Hβ) relative to the mouse β major (βmaj) probe was 2:1 unless otherwise noted. The protected 160 nt Hβ and 95 nt βmaj bands were quantified on a Molecular Dynamics PhosphorImager and percentage expression levels calculated according to the formula (Hβ/2βmaj) × 100 to account for specific activity differences. Percentage expression per copy was calculated as (2 βmaj genes/number Hβ transgenes) × % expression.

RESULTS

We built several new constructs to determine whether a β-globin transgene can be activated by multiple hypersensitive site cores in various spatial configurations when integrated at single copy (Fig. 1). Two synthetic 5′HS2 cores (24) were assembled to flank the human β-globin gene in the FL construct and were dimerized in the DM construct. The FLβ construct resembled FL but the 5′HS2 core positioned upstream of the gene was replaced by footprints 1–3 of the 5′HS3 core (26), leaving the 5′HS3 and 5′HS2 cores in their natural configuration with respect to each other.

Founder adult transgenic mice were generated with each of the constructs and identified by Southern blotting on tail DNA (Fig. 1). However, founder transgenic mice are occasionally a mosaic of transgenic and non-transgenic cells. Therefore, we bred the founders to non-transgenic animals to obtain non-mosaic F_1 fetuses representing 18 different integration events (Fig. 1). The copy number of these lines was unambiguously determined by Southern blot analysis on fetal head DNA digested with *Eco*RI and hybridized with a probe specific for human β-globin (Fig. 2A). *Eco*RI cleaves the transgene downstream of the probe and therefore single copy transgenes will be visualized as an end fragment of random size and higher copy numbers will usually contain one single copy end fragment and a multicopy head-to-tail transgene concatamer. FL transgenes are an exception and generate two multicopy transgene concatamer bands dependent on whether a monomer or a dimer of the 5′HS2 core is formed at the junction between two transgenes (Fig. 2B).

By this process, we identified a range of copy numbers of intact transgenes, including a number of single copy lines for the FL and FLβ constructs. The copy number was verified for all lines by Southern blot analyses of *Bam*HI digested DNA and transgene intactness was confirmed by PCR analyses of the 5′- and 3′-ends

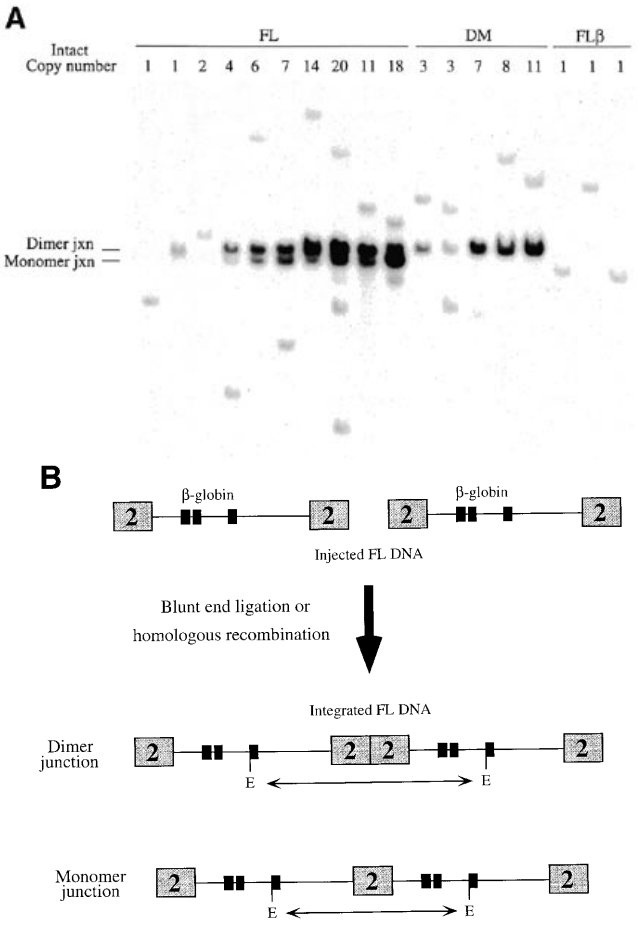

Figure 2. Generation of transgenic mice containing human β-globin transgenes regulated by multiple HS core elements. (**A**) Copy numbers of FL, DM and FLβ transgenic mouse lines determined by Southern blot analysis of F_1 fetal head DNA digested with *Eco*RI and hybridized to the βivs2 probe. The two copy line contains two end fragments in Southern blots of *Bam*HI digests (data not shown). (**B**) Map of two different head-to-tail FL transgene concatamer junctions. Blunt end ligation of two FL transgenes generates a dimer junction containing two adjacent 5′HS2 cores. Homologous recombination of the 5′HS2 core at the 3′-end of one FL transgene with the 5′HS2 core at the 5′-end of another FL transgene generates a monomer junction containing a single 5′HS2 core. These junctions can form in the same multicopy transgene concatamer and are identified by diagnostic *Eco*RI fragments (see A).

(primers shown in Fig. 1; data not shown). Transgene intactness was further confirmed by additional Southern blots of *Bsp*HI and *Bam*HI/*Xba*I digested DNA for the FL and DM or *Hin*dIII/*Eco*RI and *Bam*HI/*Xba*I digested DNA for the FLβ lines (data not shown).

S1 nuclease analysis was performed on RNA extracted from 13.5 day transgenic F_1 fetal livers to examine the expression status of the transgenes relative to the endogenous mouse β major genes. As a positive control we used fetal liver RNA from line 72, which contains a single copy of the entire human β-globin locus and expresses human β-globin mRNA at ~50% of the level produced by the two mouse β major genes, or 100% per copy (27). In the S1 analysis shown in Figure 3 human β-globin expression by line 72 was calculated to be 90% per copy. Globin genes that are not linked to LCR sequences express at <1% per copy (6,26).

In contrast to the positive control line 72, the FL construct was expressed at <1% per copy in both single copy lines and also in the two copy line (Fig. 3). The latter contains two transgenes that are not in a head-to-tail concatamer but must be present on the same chromosome (Southern data not shown). However, human β-globin expression was detected to a mean average of 14% per copy in the lines containing four or more FL transgenes, all of which contain both monomer and dimer junctions. The DM construct was expressed at <1% per copy in both three copy lines and significant human β-globin levels of 6% per copy were only obtained from the 11 copy line (Fig. 3). It is notable that the low copy DM transgenes (three to seven copies) fail to express to significant levels, in contrast to the equivalent copy numbers of the FL construct. The sole difference between the two constructs at low copy number is that monomer junctions are only found in FL transgene concatamers. These data demonstrate that dimers of 5′HS2 cores actually inhibit LCR activity in low copy animals and suggest that HS core elements must be spaced apart to function.

To test whether the smallest functional core of 5′HS3 (footprints 1–3) is able to cooperate with the 5′HS2 core in a flanking configuration to reproducibly activate single copy gene expression, we examined the three FLβ lines. FLβ transgenes were expressed to a maximum of 10% per copy in one single copy line, but activation was not reproducible, as shown by the detection of a single copy line that expresses at <1% per copy (Fig. 3). These data demonstrate that footprints 1–3 of the 5′HS3 core are unable to reproducibly activate single copy transgene expression.

We previously showed that the 10.5 kb microlocus construct composed of all four HS linked to the 1555 bp β-globin promoter (28) fully activates single copy β-globin transgene expression in a reproducible manner (29). However, this construct is too large to be packaged into a retrovirus vector for gene therapy. As it is important to find an LCR construct that functions at single copy and is of a reduced size, we created a novel LCR cassette which contains 5′HS2, 5′HS3 and 5′HS4 cloned in such a way that each core is separated by ~700 bp (Fig. 4A). This 3.0 kb LCR cassette was cloned 5′ of the 815 bp β-globin promoter to make the 7.2 kb BGT14 construct.

Rather than continue with the time consuming process of establishing transgenic lines as described above, we wished to take advantage of a more rapid assay to detect transgene expression in founder transgenic mice at the fetal liver stage. Previous reports have demonstrated that passage through the germline does not affect human β-globin transgene expression levels in mice. For example, the microlocus LCR construct expresses at ~100% levels in single copy founder fetuses (10) and to similar levels in single copy transgenic lines (29). In addition, 5′HS2 fragments fail to express reproducibly in both single copy founders (24) and lines (29). These findings validate the founder approach for determining whether a transgene construct reproducibly expresses at all single copy integration sites. However, in this context it is crucial to demonstrate that founder mice are transgenic in the fetal liver.

Therefore, we microinjected the BGT14 construct into fertilized mouse eggs and dissected potential transgenic F_0 fetuses at day 15.5 to obtain slightly larger fetal livers, which were divided into two and frozen. Genomic DNA was extracted from the corresponding fetal head tissue and screened by slot blot hybridization for the transgene. The copy number of positive fetuses was determined

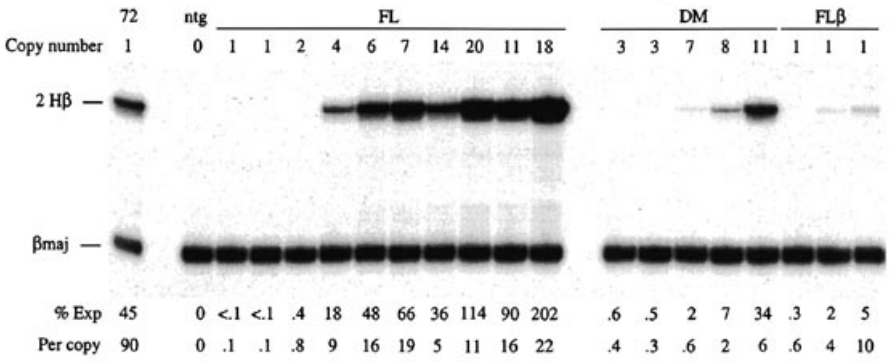

Figure 3. Expression of globin mRNA in transgenic mice containing multiple HS core elements. S1 nuclease analysis of globin expression in RNA of 13.5 day fetal livers from FL, DM and FLβ transgenic lines showing that single copy mice do not reproducibly express. Transgenic samples loaded in the same order as in Figure 2A. Transgene copy number is shown at the the top of the figure and transgene expression level at the bottom of the figure. Hβ, human β-globin protected probe fragment; βmaj, mouse β major protected probe fragment; Ntg, non-transgenic.

by Southern blot analysis of *Eco*RI (Fig. 4B, upper panel) and *Bam*HI (data not shown) digested DNA and transgene intactness confirmed by Southern blot analysis of *Sac*I/*Nde*I digested DNA (data not shown). A total of 14 positive fetuses with intact transgenes were identified, four of which were single copy.

To demonstrate the presence of BGT14 transgenes in the fetal liver and to assess the degree of mosaicism in that tissue, DNA was extracted from half the frozen fetal liver and digested with *Acc*I, which liberates a 1.8 kb β-globin fragment regardless of the integration site (Fig. 4B, lower panel). The intensity of the B26 single copy bred line (29) corresponds to 100% transgenic tissue. By comparison of the intensity of the *Acc*I digested single copy BGT14 transgene bands with the B26 standard, it was possible to calculate the percentage of transgenic cells in each fetal liver. Three of the single copy BGT14 fetuses were fully transgenic in the fetal liver and only one was a mosaic composed of 29% transgenic cells. The mosaic levels deduced for multicopy animals may be less accurate as their copy number (Fig. 4B, upper panel) is more difficult to assign with certainty. Nevertheless, all the multicopy BGT14 animals were transgenic in the fetal liver.

S1 nuclease analysis was then performed on RNA extracted from the remaining half of the frozen BGT14 fetal livers (Fig. 5A). As a positive control we used fetal liver RNA from the μD14 line, which contains a single copy of the microlocus LCR that expresses at ~100% per copy (29). The four single copy BGT14 transgenes all expressed between 6 and 34% of the two mouse βmaj genes, or 12–68% per copy. Taking into account the level of mosaicism, the per copy expression ranges from 30 to 71% for the single copy transgenes, with a mean of 45%. The multicopy transgenes express in a similar range. These data demonstrate that the BGT14 construct reproducibly expresses to high levels at all integration sites and at single copy.

Finally, we assessed whether the novel LCR cassette functioned at single copy when linked to the 265 bp β-globin promoter commonly used in gene therapy vectors (17,19). This BGT15 construct (Fig. 4A) was microinjected and three single copy 15.5 day transgenic F_0 fetuses were obtained in addition to six multicopy animals, as deduced by Southern blot analysis on *Eco*RI (Fig. 4C, upper panel) and *Bam*HI (data not shown) digested fetal head DNA. Transgene intactness was confirmed by Southern blot analysis of *Sac*I/*Nde*I digested DNA (data not shown). Mosaic analysis of fetal liver DNA was performed (Fig. 4C, lower panel) and showed that the intact single copy animals were 66 and 71% transgenic in the fetal liver, but insufficient DNA was obtained for accurate quantification of the third single copy mouse. This mouse is assumed to be 67% transgenic, as deduced from fetal head DNA (data not shown).

S1 nuclease analysis (Fig. 5B) of fetal liver RNA demonstrated that two of the single copy BGT15 animals express low β-globin levels of 3 and 11% per copy, but 39% expression levels from the third single copy animal increases the mean average to 18% per copy. The multicopy BGT15 transgenes express high levels of β-globin mRNA to a mean of 65% per copy. These data show that expression from single copy BGT15 transgenes containing the 265 bp promoter is low in comparison with the BGT14 construct. This effect could be due to: (i) deletion of crucial *trans*-acting factor sites in the 265 bp promoter; (ii) reduced spacing between the LCR and the basal promoter elements. We conclude that expression directed by the 815 bp promoter in the BGT14 construct is more suitable for gene therapy applications.

DISCUSSION

The LCR is currently defined as an activity that confers copy number-dependent, position-independent expression on linked transgenes in mice (3). In practice, this assay is conducted by generating several transgenic mice with a range of copy numbers and the emphasis is on demonstrating expression from multicopy transgenes. A consequence is that the behavior of single copy transgenes is generally considered to be of lesser importance, when in fact single copy transgenes are a more accurate reflection of the *in vivo* context of a gene and its regulatory elements. In order for a putative LCR element to reproducibly direct significant expression from single copy transgenes at independent integration sites, it must possess both transcriptional enhancement and chromatin opening activities (29). Both these activities are implicated in LCR activity *in vivo* and clearly would be of benefit in gene therapy vectors that integrate at single copy.

We previously showed that the small 5′HS2 core element directed copy number-dependent expression of multicopy β-globin transgene concatamers but failed to direct expression of single copy transgenes (24,25). More recently it has been shown that

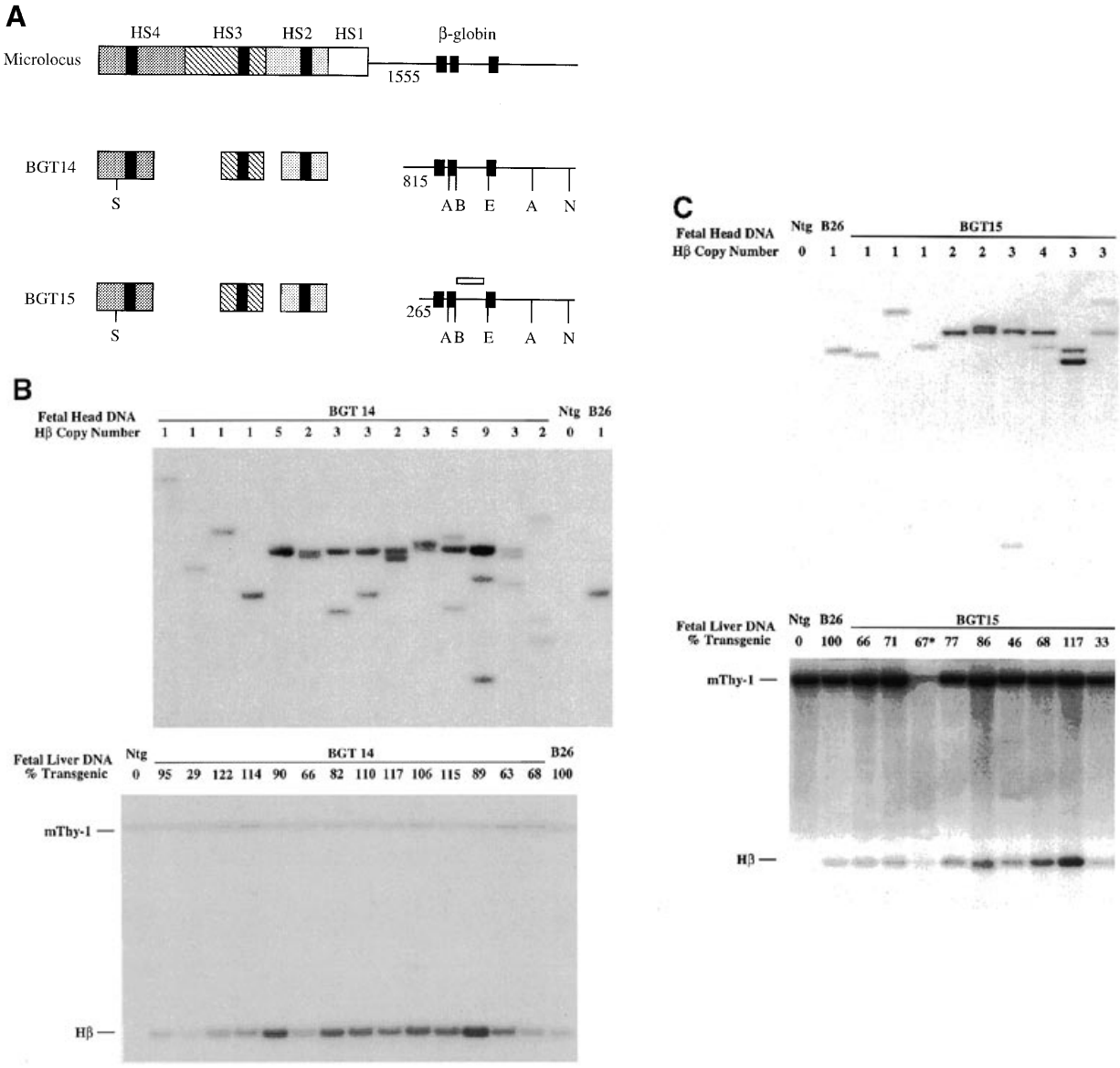

Figure 4. Generation of transgenic mice containing the novel 3.0 kb LCR cassette. (**A**) Map of the 10.5 kb microlocus (28), 7.2 kb BGT14 and 6.7 kb BGT15 constructs. The 3.0 kb LCR cassette in BGT14 and BGT15 contains the 1.15 kb *Stu*I–*Spe*II fragment of 5′HS4, the 0.85 kb *Sac*I–*Pvu*II fragment of 5′HS3 and the 0.95 kb *Sma*I–*Stu*I fragment of 5′HS2 (see Materials and Methods for cloning strategy). The length of the promoter is shown under each construct. S,*Sac*I; A, *Acc*I; B, *Bam*HI; E, *Eco*RI; N, *Nde*I. (**B**) (Upper panel) Copy numbers of BGT14 transgenic mice determined by Southern blot analysis of F_0 fetal head DNA digested with *Eco*RI and hybridized to the βivs2 probe. (Lower panel) Southern blot of *Acc*I digested DNA extracted from BGT14 fetal livers and hybridized with the βivs2 (Hβ) and mouse Thy-1 (mThy-1) probes. The percentage of transgenic cells was calculated by comparison of the Hβ band intensity of each BGT14 transgenic mouse with the one copy bred line (B26), taking into account copy number and the mouse Thy-1 loading control (see Materials and Methods). (**C**) (Upper panel) Copy numbers of BGT15 transgenic mice determined as in (B). (Lower panel) Mosaic analysis of BGT15 transgenic fetal livers determined as in (B). *Percentage transgenic cells in this mouse was calculated using fetal head DNA (see text).

5′HS2 is not essential for high level globin expression in mice (30) and is unable to activate β-globin gene expression from retroviral vectors microinjected into transgenic mice (31). Here, we evaluate whether multiple hypersensitive sites present as small core elements can reproducibly activate single copy transgene expression. Our results demonstrate that flanking configurations of two 5′HS2 cores or of one 5′HS3 (footprints 1–3) and one 5′HS2 core cannot direct reproducible single copy transgene expression. These data indicate that 5′HS2 and 5′HS3 (footprints 1–3) cores are not ideally suited for gene therapy. The importance of 5′HS3 in control of the β-globin locus in mouse and man is controversial (32).

We also demonstrate that pairs of 5′HS2 cores in a dimer (DM) configuration fail to direct expression from low copy number transgene concatamers. In contrast, expression at low copy numbers is reproducibly detectable when the 5′HS2 cores flank (FL) the β-globin gene. The only difference between the DM and FL constructs when integrated as low copy head-to-tail concatamers is at the junctions between adjacent transgenes. FL transgene concatamers have both dimer and monomer 5′HS2 cores at their

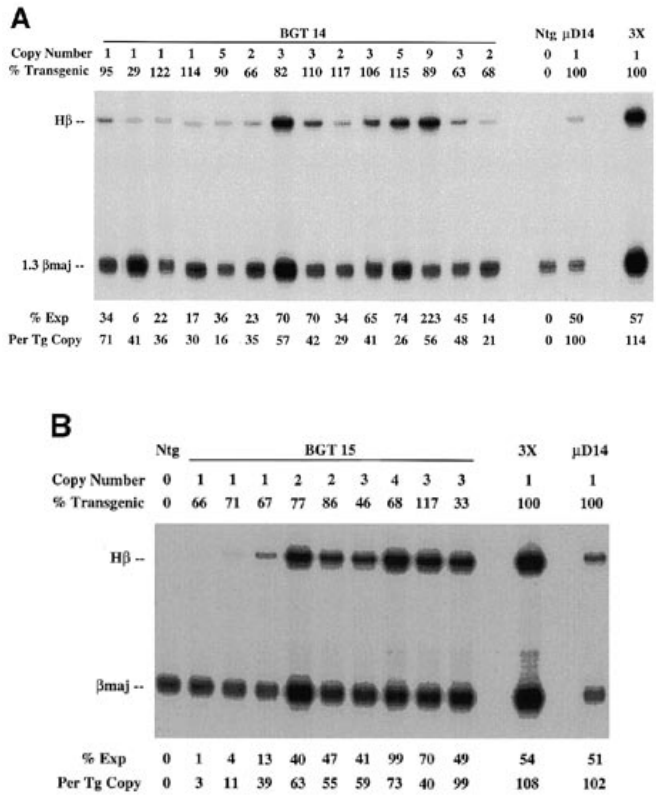

Figure 5. Expression of globin mRNA in transgenic mice containing the 3.0 kb LCR cassette. (**A**) S1 nuclease analysis of fetal liver RNA of 15.5 day F_0 BGT14 transgenic mice showing that all four single copy animals express to high levels. Transgenic samples loaded in the same order as in Figure 4B. (**B**) S1 nuclease analysis of fetal liver RNA of 15.5 day F_0 BGT15 transgenic mice showing that the single copy animals express variable transgene levels. Transgenic samples loaded in the same order as in Figure 4C. Hβ, human β-globin protected probe fragment; βmaj, mouse β major protected probe fragment; Ntg, non-transgenic; μD, one copy μD14 microlocus line; 3X, probe excess control.

junctions, whereas only dimer junctions exist in DM transgene concatamers. The observation that FL but not DM transgene concatamers expressed well at low copy numbers of three to seven suggests that: (i) reproducible expression is only obtained in the presence of monomer junctions; (ii) transgenes adjacent to dimer junctions are not usually expressed; (iii) therefore dimerization of short 5′HS2 cores is detrimental to LCR activity. This finding indicates that 5′HS2 cores must be separated by spacer DNA in order to activate a transgene and implies that simple oligomerization of several hypersensitive site cores will not create an LCR that is fully functional at single copy (16,17,19).

To date the only LCR constructs that are fully functional in single copy transgenic mice contain all four hypersensitive sites with well spaced cores (4,5,27,29,33). Therefore, to obtain full levels of expression from single copy β-globin vectors it will be necessary to employ large LCR cassettes rather than core elements. To this end, we used a 0.85 kb fragment of 5′HS3 as the foundation of a novel 3.0 kb LCR cassette and flanked it by 1.15 kb 5′HS4 and 0.95 kb 5′HS2 fragments. The 815 bp promoter in this BGT14 construct reproducibly expressed at single copy and the average level of expression of 45% per copy indicates that the BGT14 cassette may be suitable for gene therapy constructs. Surprisingly, the same 3.0 kb LCR did not direct reproducible expression at high levels from the 265 bp promoter in single copy BGT15 transgenes. Therefore, the sequences between –815 and –265 bp must either: (i) contain binding sites for *trans*-acting factors that participate in single copy transgene activation; (ii) serve as spacer DNA to facilitate interaction of the LCR with the basal promoter element (34,35).

Full expression is only obtained from single transgene copies of the 10.5 kb microlocus construct, which differs from the BGT14 construct in having 5′HS1, the 1555 bp promoter and 2.5 kb of additional sequences in 5′HS2, 5′HS3 and 5′HS4 (Fig. 4A). The importance of these sequences remains to be assessed in single copy transgenic mice. The 7.2 kb BGT14 construct itself is too large to be incorporated into AAV vectors, but is within the packaging capacity of a retrovirus if the drug resistance gene is omitted from the vector.

In conclusion, single copy transgenic mice provide the most comprehensive assay for LCR activity and should be useful for evaluating *in vivo* expression from potential gene therapy constructs while avoiding the vagaries of producing a stable high titer virus until necessary. We have optimized our microinjection conditions for generating single copy transgenic founder mice by merely adjusting the DNA concentration. Although we excluded mosaic animals from some of our analyses by establishing transgenic lines, founder animals that contain unique transgene end fragments in both the 5′ and 3′ directions and have been carefully screened on Southern blots for mosaicism in the fetal liver by comparison with a known single copy transgenic line are equally useful.

ACKNOWLEDGEMENTS

We thank S.Philipsen for discussions and for providing the 5′HS3 (footprints 1–3) core, J.Strouboulis for line 72 RNA, N.Solomon, C.Wilson and A.Huang for animal care and all those who tailed mice, especially A.Bygrave and A.Elefanty. We benefitted from helpful discussions with N.Dillon and M.Antoniou. This work was supported by grants to J.E. from the Medical Research Council of Canada (MRC) and to F.G. from MRC (UK).

REFERENCES

1 Dillon,N. (1993) *Trends Biotechnol.*, **11**, 167–173.
2 Mulligan,R.C. (1993) *Science*, **260**, 926–932.
3 Grosveld,F., van Assendelft,G.B., Greaves,D.R. and Kollias,G. (1987) *Cell*, **51**, 975–985.
4 Gaensler,K.M., Kitamura,M. and Kan,Y.W. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 11381–11385.
5 Peterson,K.R., Clegg,C.H., Huxley,C., Josephson,B.M., Haugen,H.S., Furukawa,T. and Stamatoyannopoulos,G. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 7593–7597.
6 Ryan,T.M., Behringer,R.R., Martin,N.C., Townes,T.M., Palmiter,R.D. and Brinster,R.L. (1989) *Genes Dev.*, **3**, 314–323.
7 Tuan,D., Solomon,W., Li,Q. and London,I.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 6384–6388.
8 Forrester,W.C., Thompson,C., Elder,J.T. and Groudine,M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1359–1363.
9 Caterina,J.J., Ryan,T.M., Pawlik,K.M., Palmiter,R.D., Brinster,R.L., Behringer,R.R. and Townes,T.M. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1626–1630.
10 Talbot,D., Philipsen,S., Fraser,P. and Grosveld,F. (1990) *EMBO J.*, **9**, 2169–2177.
11 Talbot,D. and Grosveld,F. (1991) *EMBO J.*, **10**, 1391–1398.
12 Chang,J.C., Liu,D. and Kan,Y.W. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 3107–3110.
13 Miller,J.L., Walsh,C.E., Ney,P.A., Samulski,R.J. and Nienhuis,A.W. (1993) *Blood*, **82**, 1900–1906.

14 Novak,U., Harris,E.A., Forrester,W., Groudine,M. and Gelinas,R. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 3386–3390.
15 Takekoshi,K.J., Oh,Y.H., Westerman,K.W., London,I.M. and Leboulch,P. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 3014–3018.
16 Einerhand,M.P., Antoniou,M., Zolotukhin,S., Muzyczka,N., Berns,K.I., Grosveld,F. and Valerio,D. (1995) *Gene Ther.*, **2**, 336–343.
17 Leboulch,P., Huang,G.M., Humphries,R.K., Oh,Y.H., Eaves,C.J., Tuan,D.Y. and London,I.M. (1994) *EMBO J.*, **13**, 3065–3076.
18 Miller,J.L., Donahue,R.E., Sellers,S.E., Samulski,R.J., Young,N.S. and Nienhuis,A.W. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 10183–10187.
19 Sadelain,M., Wang,C.H., Antoniou,M., Grosveld,F. and Mulligan,R.C. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 6728–6732.
20 Plavec,I., Papayannopoulou,T., Maury,C. and Meyer,F. (1993) *Blood*, **81**, 1384–1392.
21 Ren,S., Wong,B.Y., Li,J., Luo,X., Wong,P.M. and Atweh,G.F. (1996) *Blood*, **87**, 2518–2524.
22 Miller,A.D. (1992) *Nature*, **357**, 455–460.
23 Larochelle,A., Vormoor,J., Lapidot,T., Sher,G., Furukawa,T., Li,Q., Shultz,L.D., Olivieri,N.F., Stamatoyannopoulos,G. and Dick,J.E. (1995) *Hum. Mol. Genet.*, **4**, 163–172.
24 Ellis,J., Talbot,D., Dillon,N. and Grosveld,F. (1993) *EMBO J.*, **12**, 127–134.
25 Ellis,J. and Grosveld,F. (1993) *Semin. Dev. Biol.*, **4**, 359–369.
26 Philipsen,S., Pruzina,S. and Grosveld,F. (1993) *EMBO J.*, **12**, 1077–1085.
27 Strouboulis,J., Dillon,N. and Grosveld,F. (1992) *Genes Dev.*, **6**, 1857–1864.
28 Talbot,D., Collis,P., Antoniou,M., Vidal,M., Grosveld,F. and Greaves,D.R. (1989) *Nature*, **338**, 352–355.
29 Ellis,J., Tan-Un,K.C., Harper,A., Michalovich,D., Yannoutsos,N., Philipsen,S. and Grosveld,F. (1996) *EMBO J.*, **15**, 562–568.
30 Fiering,S., Epner,E., Robinson,K., Zhuang,Y., Telling,A., Hu,M., Martin,D.I., Enver,T., Ley,T.J. and Groudine,M. (1995) *Genes Dev.*, **9**, 2203–2213.
31 McCune,S.L. and Townes,T.M. (1994) *Nucleic Acids Res.*, **22**, 4477–4481.
32 Wood,W.G. (1996) *Trends Genet.*, **12**, 204–206.
33 Bungert,J., Dave,U., Lim,K.C., Lieuw,K.H., Shavit,J.A., Liu,Q. and Engel,J.D. (1995) *Genes Dev.*, **9**, 3083–3096.
34 Milot,E., Strouboulis,J., Trimborn,T., Wijgerde,M., de Boer,E., Langeveld,A., Tan-Un,K., Vergeer,W., Yannoutsos,N., Grosveld,F. and Fraser,P. (1996) *Cell*, **87**, 105–114.
35 Wijgerde,M., Grosveld,F. and Fraser,P. (1995) *Nature*, **377**, 209–213.