Variegated transgene expression in mouse mammary gland is determined by the transgene integration locus

(β-lactoglobulin/epigenetic silencing/position-effect variegation)

KENNETH W. DOBIE†, MURIEL LEE‡, JUDITH A. FANTES§, ELIZABETH GRAHAM†, A. JOHN CLARK†, ANTHEA SPRINGBETT†, RICHARD LATHE*, and MARGARET MCCLENAUGHANT‡

*Centre for Genome Research, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JQ, United Kingdom; ‡Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, United Kingdom; and §Medical Research Council Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, United Kingdom

Communicated by Allan C. Spradling, Carnegie Institution of Washington, Baltimore, MD, February 9, 1996 (received for review November 2, 1995)

ABSTRACT Mice carrying an ovine β-lactoglobulin (BLG) transgene secrete BLG protein into their milk. To explore transgene expression stability, we studied expression levels in three BLG transgenic mouse lines. Unexpectedly, two lines exhibited variable levels of transgene expression. Copy number within lines appeared to be stable and there was no evidence of transgene rearrangement. In the most variable line, BLG production levels were stable within individual mice in two successive lactations. Backcrossing demonstrated that genetic background did not contribute significantly to variable expression. Tissue in situ hybridization revealed mosaicism of transgene expression within individual mammary glands from the two variable lines; in low expressors, discrete patches of cells expressing the transgene were observed. Transgene protein concentrations in milk reflected the proportion of epithelial cells expressing BLG mRNA. Furthermore, chromosomal in situ hybridization revealed that transgene arrays in both lines are situated close to the centromere. We propose that mosaicism of transgene expression is a consequence of the chromosomal location and/or the nature of the primary transgene integration event.

β-Lactoglobulin (BLG) is a major ovine milk protein. The function of BLG is unknown, though the crystal structure of bovine BLG is consistent with a role in vitamin A transport (1). We previously reported that mice carrying a sheep BLG transgene secrete BLG into their milk (2); BLG regulatory regions can direct expression of biomedical proteins into the milk of transgenic mice and sheep (3–5). In this context, it is important that transgene expression is stable. Unstable transgene expression has been described previously; Palmiter et al. (6) reported that the level of herpes simplex virus thymidine kinase expression could vary by more than an order of magnitude among progeny of the same founder. Although other transgene insertions express to variable degrees within individual cell lines or transgenic mouse lines (refs. 7–19; M. Mehtali and R.L., unpublished data), there has been no common explanation for the instability of expression. Unstable expression may be due to strong selection against the transgene, for instance by the failure of sperm fertility engendered by testicular thymidine kinase expression (7, 8) or by the toxicity of high-level hepatic expression of plasmaminogen activator (9). A transgene inserted into the X chromosome (10) or an X-autosome translocation (20) generates mosaic expression due to stochastic X chromosome inactivation. Silencing has also been observed when the transgene integrates into repeat sequence or satellite DNA (11, 12), whereas different levels of transgene expression between animals of the same lineage have been attributed to strain-specific modifier genes (13–15).

Mosaic patterns of expression were also observed in transgenic animals bearing intestinal fatty-acid binding protein fusion transgenes (16). Here, mosaicism was attributed to a deficit of cis-acting elements in the transgene. Mosaic expression patterns were also observed in mice carrying different tyrosinase fusion transgenes (17–19); these reports suggested that the striated coat color was an inherent property of the transgene.

To address the stability of transgene expression, we examined three transgenic mouse lines harboring an intact BLG gene. In line 14, BLG expression levels were stable, whereas in lines 7 and 45, BLG levels varied significantly. We examine the nature of this variable expression and show that it reflects variegated patterns of transgene expression within the mammary gland. We demonstrate that variable expression is independent of genetic background and is a property of the transgene locus.

MATERIALS AND METHODS

Mice, Sampling, and Milk Protein Analysis. Transgenic lines 7, 14, and 45 (2) were maintained by systematic crossing to C57BL/6 × CBA F1 hybrid mice. Litters were standardized to five pups per mother at birth. Milk and tissue samples were collected from 6- to 8-week old transgenic mice (hemizygous for the transgene array) at day 11 of lactation. Backcross experiments used CBA and C57BL/6 mice (Harlan Olac, Bicester, U.K.). Transgenic mice were identified by a PCR assay (21). Milk collection and SDS/20% PAGE analysis of milk proteins were performed as described (22).

DNA Analysis. Genomic mammary DNA was obtained from frozen tissue by standard procedures. EcoRI-digested DNA samples (10 μg) were electrophoresed and transferred to ZetaProbe (Bio-Rad) membranes. BLG and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probes were gel purified. A mouse total genomic DNA probe was prepared from nontransgenic mammary tissue. BLG, GAPDH, and mouse genomic DNA probes were labeled with [α-32P]dCTP using a commercial labeling system (Multiprime, Amersham). Following hybridization, results were quantified using a Molecular Dynamics PhosphorImager.

RNA Analysis. Frozen mammary tissue was homogenized in 2 ml Rnazol B (Biogenesis Ltd.) and total mRNA extracted. Samples (10 μg) were electrophoresed on 1.5% denaturing Mops/formaldehyde agarose gels, transferred, and hybridized with BLG and GAPDH probes.

Tissue in Situ Hybridization. PstI fragments from the BLG (424 bp; ref. 23) and β-casein (440 bp; ref. 24) cDNA,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
respectively, were subcloned into pGEM-4Z in both orientations. dUTP\(^{35}\)S-labeled single-stranded antisense probes and control sense probes were synthesized by in vitro transcription using standard procedures (25). Mammary tissue was collected from hemizygous mice at day 11 of lactation and processed for in situ hybridization (26).

**Chromosome Fluorescence in Situ Hybridization (FISH).** Lymphocyte cells from mouse spleens were cultured in RPMI 1640 medium with 10% fetal calf serum and 50 \(\mu\)g/ml lipopolysaccharide, at a density of 1 \(\times\) 10\(^7\) cells/ml for 45 hr at 37°C. These cells were used for Giemsa banding and karyotyping of chromosomes (27). Genomic ovine BLG clone pS1tgXS (2) was nick-translated with biotin-16-dUTP (28) and used to assign transgenes to specific metaphase chromosomes from hemizygous line 7, 14, and 45 mice using avidin-fluorescein isothiocyanate (FITC) (29). Digital images were obtained with a Bio-Rad confocal microscope. Double-color FISH used BLG and mouse major satellite DNA (30) probes (see Fig. 4 legend for details) and results were visualized as described (31).

**RESULTS**

**Lines 7 and 45 Exhibit Variable Expression of BLG.** Lines 7 and 14 harbor an identical 16 kb BLG transgene, whereas line 45 carries an 11 kb 3' truncated transgene; the 3' truncation does not affect the transcription unit nor transgene function (2). Hemizygous transgenic female mice were mated, milk collected at day 11 of lactation, and BLG protein levels determined.

The ranges of BLG expression levels were different among the three lines (Fig. 1). Line 14 expressed BLG in a stable fashion with only 3.5 mg/ml separating the lowest and the highest expressor. However, BLG levels in milk of line 45 animals ranged from 16 to 30 mg/ml, whereas in line 7 there was a 8-fold difference between the lowest and highest expressors (3 to 23.9 mg/ml). Table 1 shows the standard deviations (SDs) and coefficients of variation (c.v.) for BLG protein expression in the three lines. As expected, lines 7 and 45 had larger SD values than line 14. The c.v. for line 45 was similar to that of line 14, but this reflects its higher average expression level and also a possible upper limit to protein production. We conclude that the particular transgene insertion locus in line 7 determines variation in expression levels of BLG protein, for the identical transgene resulted in stable expression in line 14 animals.

To establish whether variation occurred at the steady-state mRNA level, quantitative Northern blot analyses using BLG and control (GAPDH) hybridization probes were performed. BLG mRNA levels were more variable in lactating gland from animals of lines 7 and 45 than in line 14 animals (Table 1). The correlation coefficients for protein and RNA levels were 0.44, 0.54, and 0.54 for lines 14 (\(n = 9\)), 45 (\(n = 9\)) and 7 (\(n = 8\), after the exclusion of one outlier), respectively. The results show variable levels of RNA expression in lines 7 and 45; the relationship between protein and RNA was similar in all three lines.

**Variable Expression Is Not Due to Rearrangement of the Transgene Locus.** We were concerned that variable expression might be due to modifications to the structure of the transgene locus. Therefore, we performed Southern blot analysis on DNA prepared from lactating mammary tissue. Lines 7 and 45 carried 25 and 17 transgene copies, respectively, whereas line 14 harbored 2 copies (32). In all samples, the BLG probe identified a single major 4.3-kb band with no evidence of variations in banding patterns to suggest rearrangement of the transgene arrays. Quantitative hybridization using probes corresponding to ovine BLG and to total genomic DNA revealed that the transgene copy number was indistinguishable within lines, even among mice from line 7 exhibiting the full spread of BLG levels. All lines had similar c.v. values (Table 1) and this contrasts with the protein and RNA data in which the SD and c.v. values differed considerably for the three lines. Furthermore, the c.v. values for protein and RNA in line 7 exceeded that for DNA. While these analyses would not reveal subtle changes in the structure of the transgene array, we surmise that variable expression is not due to stochastic loss or rearrangement of the transgene array in different individual line 7 and 45 mice.

**BLG Expression Does Not Vary Between Lactations.** To determine whether expression levels exhibit temporal variation within individual line 7 animals, milk BLG protein levels were measured in 28 individuals at day 11 of two successive lactations (Fig. 24). The close correlation between BLG levels measured in two lactations argues that individual levels are fixed by day 11 of the first lactation.

**Variable Expression Is Not Due to Genetic Background.** The transgenic mice studied were of a C57BL/6 \(\times\) CBA hybrid background. Variable transgene expression among line 7 mice might be due to segregation of alleles of a modifier gene(s). We therefore backcrossed transgenic animals from line 7 to C57BL/6 and CBA animals for three generations; milk samples were collected at day 11 of lactation and the levels of BLG determined (Fig. 2B).

The variance and mean BLG levels were not significantly different between the third backcross C57 population and the mixed CBA/C57 population (\(P > 0.05\)). Although variance and mean values were significantly different between the third backcross CBA and CBA/C57 populations (\(P < 0.05\)), the difference in the range of expression levels was small. We have since generated data from fourth generation backcross animals (96.9% C57, \(n = 14\); 96.9% CBA, \(n = 13\)) and confirm that variance values for both backgrounds are again similar to the original CBA/C57 population (\(P > 0.05\)). There was also no clear trend in BLG expression levels through sublines within each backcross; for example, a parent expressing 11 mg/ml gave rise to daughters with expression levels of 5 mg/ml and 22 mg/ml.

It is predicted that the presence of single or multiple modifier loci would be revealed by major alterations in both the variance and mean values for expression levels over three generations. Because no such alterations were observed by the fourth backcross, we conclude the variable expression of the transgene within line 7 is not due to heterogeneity of genetic background.

**Mosaic Expression of BLG in the Mammary Gland.** We wished to determine whether levels of BLG expression were reflected in differential expression of the transgene in the mammary epithelium. In situ hybridization studies were performed on abutting tissue sections collected at day 11 of lactation from individuals of lines 7, 14, and 45. Endogenous \(\beta\)-casein mRNA was detected throughout the mammary epithelium in all three transgenic lines, demonstrating that all secretory epithelial cells of the gland have the capacity to express a milk protein gene (Fig. 3 B, D, and F). Sections from

![Fig. 1. BLG expression level profiles in milk collected from line 7, 14, and 45 hemizygous mice.](image-url)
low-expressing line 7 individuals showed small clusters of cells strongly expressing BLG mRNA, surrounded by nonexpressing cells (Fig. 3A). High expressors from line 7 had the opposite pattern of expression with small regions of cells negative for BLG mRNA (Fig. 3C). The proportion of positive cells in any one section correlated with the level of BLG in the milk; higher expressors had a greater proportion of cells positive for transgene mRNA. Line 14 was distinctly different because BLG expression was present in all the epithelial cells examined, matching the pattern of β-casein mRNA expression (Fig. 3 E and F). In lines 7 and 14, the introduced transgene is identical, arguing that mosaic expression is due to the location and/or nature of the transgene array. Line 45 also exhibited mosaic expression of BLG, although less so than in line 7 (Fig. 3G). This may reflect the higher average expression level in line 45; individuals from line 7 and 45 with the same levels of BLG protein exhibited a similar degree of mosaicism. Although this analysis does not permit precise quantitation of per-cell expression levels, the results suggest that variations in the level of BLG protein and mRNA expression in line 7 and 45 individuals appear to be due to variations in the proportion of mammary epithelial cells expressing the transgene.

**The Transgene Locus Lies Near a Centromere in Lines 7 and 45.** Variegated expression of endogenous genes (33) and transgenes (34) has been described in *Drosophila* and linked to proximity to the centromere. To determine the subchromosomal localization of the BLG transgenes, we applied *in situ* hybridization to chromosome spreads prepared from animals of different transgenic lines. G-banding and karyotyping in conjunction with FISH using a BLG probe demonstrated that the transgene arrays had integrated into chromosomes 15 (line 7), 7 (line 14), and 5 (line 45). Further, the BLG locus had integrated midarm in line 14 (Fig. 4a) but was close to the centromere in lines 7 and 45. To refine the localization of the transgene array relative to the centromere, we used a BLG probe in conjunction with a major satellite (196) probe that predominantly decorates centromeric regions. In all line 7 spreads, the BLG locus was close to but out with the centromere (Fig. 4b). A crude estimate from 56 chromatids indicated that the array lies within the 3.5% of the chromosome proximal to centromeric heterochromatin. In line 45 the resolution of the experiment made the array appear to be within major satellite DNA (Fig. 4c, chromosome arrow). However, the transgene signal was observed as clearly separate from major satellite DNA in interphase nuclei where the chromosomes are decondensed (Fig. 4c, nucleus arrow). By quenching BLG signals, it was clear that integrated arrays were not associated with major translocations of this satellite DNA.

**DISCUSSION**

The introduction of transgenes to the mouse germline by pronuclear injection of DNA usually results in tandemly repeated and head-to-tail arrays at a single random site within the genome. Though the transgene insertion site can clearly have an influence upon tissue-specificity and level of expression, it has often been assumed that expression of the transgene, once integrated, is stable within a transgenic line. However, processes have been described, such as the “RIP”
In situ hybridization analysis of BLG and β-casein mRNA expression patterns in mouse mammary tissue. (A and B) Hybridization of BLG and β-casein probes to tissue from a low-expressing line 7 mouse (4 mg/ml). (C and D) BLG, β-casein, high-expressing line 7 mouse (17 mg/ml). (E and F) BLG, β-casein, line 14 mouse (8 mg/ml). (G) Line 45 mouse (25 mg/ml). (H) Control, hybridization with a β-casein sense probe, line 7 mouse (17 mg/ml). (Bar = 0.5 mm.)

Figure 3. In situ hybridization analysis of BLG and β-casein mRNA expression patterns in mouse mammary tissue.

phenomenon in Neurospora (35) and cosuppression in plants (36), that result in transgene silencing and variegated expression. We therefore entered into a study of transgene expression stability using three mouse lines that express the ovine BLG gene selectively within the lactating mammary gland. Only one expressed the transgene in a stable manner, whereas wide variations in the individual level of BLG expression were observed among individuals of the other two lines.

In the most variable line (line 7), expression levels are fixed within the individual. Further, the genetic background of this line does not play a significant role in this variation. In this pedigree, as well as in a second line (line 45), mammary tissue in situ hybridization experiments revealed mosaic expression of the transgene. In contrast, all the secretory cells within transgenic glands expressed the gene for an endogenous milk protein (β-casein), arguing against mosaicism due to cycling of active and inactive domains under hormonal control (37); recently, uniform expression of four endogenous milk proteins was observed throughout the mouse mammary epithelium by day 1 of lactation (38). While our observations do not rigorously rule out threshold effects due to suboptimal levels of lactogenic hormones leading to stochastic onset of milk protein gene expression, we think this unlikely. First, sections were taken from glands at a time when milk production is near-maximal and all secretory cells are producing β-casein. Second, line 14, harboring the identical construct to that present in the highly variable line 7, shows stable expression throughout the gland.

We believe that the patching of cells in which the transgene is active or inactive is most consistent with clonal expansion. Our results suggest epigenetic silencing of transcription occurs stochastically in individual progenitor cells, which is then transmitted through cell division to daughter cells, giving rise to mosaic patterns of expression. The closest precedent for our observations is afforded by the striped coat color patterns recorded by Mintz and colleagues (17, 18) in animals harboring tyrosinase fusion transgenes. All lines appeared to reproduce the striped pattern of expression, and animals either fully-
pigmented or unpigmented were not reported; this contrasts with our observations that a transgene insertion can give rise both to animals in which the transgene is predominantly silent and to animals in which the transgene is expressed in the majority of target cells. Robertson et al. (39) have reported differing levels of expression in red blood cells for a globin/lacZ construct among transgenic mouse lines; underlying mechanisms may be similar, resulting in phenotypic differences that reflect the clonal origin of mosaic patches (phenoclines; ref. 40).

The phenomenon of position-effect variegation in Drosophila affords an informative parallel. Chromosomal rearrangement that relocates an endogenous gene locus close to the centromere can result in silencing of the locus in a variable proportion of cells (33). Silencing is attributed to chromosomal condensation or heterochromatinization brought about by proximity to the heterochromatic centromere. We argue that a similar process may occur in mouse. First, in both mouse lines exhibiting mosaic transgene expression, the array is integrated in the vicinity of the centromere, whereas in the stably-expressing line the transgene is inserted some distance away. Second, gene variegation in Drosophila can extend over 50–60 polytene bands (1–2 megabases) from a rearrangement breakpoint (41). While no precise measure can be provided, our estimates from FISH chromosomal spreads suggest that the transgene insertion of line 7 is some 4–5 megabases from centromeric heterochromatin. This is broadly consistent with the situation found in Drosophila, particularly if relative genomes sizes are taken into account. Third, Dorer and Henikoff (34) recently reported that heterochromatinization in Drosophila is not only dependent upon proximity to the centromere but also on the number of repeats present at the transgene locus. We note that both lines displaying mosaic expression harbor some 20 transgene copies, whereas the stably expressing line 14 contains only two copies. Heterochromatic features of an unusually large (1100 kb) β-globin transgene integrated into a peritelomeric region were reported recently (42). Furthermore, Mehtali et al. (43) described increased extinction of transgene activity (as judged by reporter gene expression and DNA methylation measurements) with increased copy number. It is therefore plausible to suggest that both integration site and number of transgene copies may predispose to the mosaic expression phenotype; indeed, proximity to repetitive DNA per se (either centromeric satellite DNA or repeat transgene copies) could be responsible for variable expression.

Our studies provide the first evidence for variegated expression of a transgene between individuals within a single transgenic line. Such mosaic expression has implications for investigations using gene addition technology. Gene addition is now being tested as a means of causing tissue specific cell ablation in the treatment of brain tumors (44), or in the study of the consequences of the loss of thyroid hormone production without surgery in mice (45). Variegated expression in such experiments would prevent 100% ablation and therefore result in variable numbers of cells being ablated. Similarly, genetic manipulation of animal organs for transplant purposes (46) might be prejudiced by unpredictable transgene expression. In biotechnological applications, such as the production of proteins of biomedical interest in the milk of transgenic livestock (3, 4), mosaic expression will reduce the yield of product. Since this project was completed, other colleagues have reported variable expression within transgenic lines (B. Binas and T. Burdon, personal communication; S. Morley and J. J. Mullins, personal communication) suggesting that the phenomenon may not be uncommon. Particularly intriguing is the possibility that some endogenous genes within the mammalian genome, especially those close to heterochromatic regions [e.g., immunodeficiency, centromeric instability, facial anomalies (ICF) syndrome (47)] or present as tandem duplications at a single locus, may be susceptible to the type of mosaic silencing we describe here.

We are grateful to Andy Dawson, Eric Thompson, Roberta Wallace, and Frances Thompson for assistance with experiments. We thank Arthur Mitchell for the mouse major satellite DNA probe, and Norrie Russell, Roddy Field, and Elliot Armstrong for photography and
graphics. We would also like to thank Bert Binas, John West, Duncan Davidson, and Nick Hastie for helpful discussions and for reading the manuscript.