Additional Material and Methods

pLuc vector (5.63 kb) derives from pTLuc upon deletion of promoter sequences (116 bp). pL1Luc (6.47 kb) and pL1TLuc (6.59 kb) were cloned by adding the promoter region of the LINE1 element, carrying the CpG island, identified as L1 DNA fragment (*Eco*R I-*Xba*I segment, 836 bp long; AF364302, nucleotides 4422-5258) to pLuc and pTLuc vectors, respectively.

The pHSΔATluc, pHSΔBTluc and pHSΔABTLuc (deletion constructs were derived from pHSTLuc with the following deletions: (nucleotide positions are given with respect to X76647): ΔA, nucleotides 2622 to 2654; ΔB, nucleotides 2657 to 2682, and; ΔAB, nucleotides 2622 to 2682. An scheme of these deletions is given in Figure 8.

Additional Results & Discussion

Statistical evaluation of all transgenic flies analysed, using the hierarchical log-linear model, supported our interpretation (Table II). Indeed, the large number of transgenic flies analysed, as compared to transgenic mice, enabled us to evaluate the results obtained with appropriated statistical models (i.e. hierarchical log-linear model, Table II) in order to derive more sustainable conclusions.

For statistical analysis we first verified the goodness of fit of the model by the maximum likelihood ratio and chi-square tests, both statistically significant ($p<0.01$), and only original observed variables were considered: each fly was treated as a single individual carrying a construct and displaying a given eye colour, without taking in consideration the average percentage of pigment previously used (Figure 4).

Two categorical variables were considered: “transgenic construct” (with eight possible values: “w”, BRwBR, 5’HS4w5’HS4, TE3wTE3, XwX, HSwHS, LCRwLCR and LCRmutwLCRmut) and “eye colour” (with five possible values: yellow, pale orange, orange, brown and red), resulting in a 8 x 5 table of associations. Variable “transgenic line” is hierarchically related with variable “transgenic construct”. Analysis of frequencies (crosstabs) and log-linear model between variables “eye colour” and “transgenic line” showed a similar behaviour for all lines belonging to each construct (not shown). 3059 heterozygous (Table IIA) and 3060 homozygous individuals (Table IIB) were used.

Numbers shown in Table II correspond to Z-values, a descriptive statistical parameter normally distributed N(0,1) and generated by the saturated log-linear model. Positive values of Z indicate proportionally higher probability of association. Conversely, negative values of Z indicate proportionally lower probability of association. These values show, for instance, that transgenic flies carrying the reference construct “w” are most likely to display dark eye colours (Table IIA, red; Table IIB, brown), and less likely to display light eye colours (Table IIA, yellow; Table IIB, yellow and pale orange). On the other hand, individuals carrying the positive control BRwBR are most likely to display light eye colours (Table IIA, orange; Table IIB, pale orange, orange or brown), and less likely to display dark eye colours (Table IIA, red; Table IIB, red), in good agreement with its known boundary features. Experimental HSwHS and LCRwLCR constructs show high probabilities to display light eye colours (Table IIA, yellow; Table IIB, yellow and pale orange) and low probabilities to display dark eye colours (Table IIA, orange and red; Table IIB, red), opposite to the effects observed for both the reference “w” construct and the negative

control TE3wTE3, that also depicts higher probabilities for flies with dark eye colours (Table IIA and Table IIB, brown and red).

The vertebrate positive control 5'HS4w5'HS4 behaves more similar to "w" and TE3wTE3 than BRwBR constructs, and thus, with this test, does not behave properly as a true insulator sequence. LCRmutwLCRmut flies display an increasing tendency towards darker eye colours in homozygous individuals (Table IIB, orange and red), opposite to what is observed with LCRwLCR flies, suggesting a decrease or lost of insulating effects from what is observed with the same LCRmutwLCRmut construct in heterozygous individuals. XwX flies appear highly dispersed. However, the pattern observed from statistical analysis with XwX flies does not match that of HSwHS or LCRwLCR flies, further suggesting absence or decrease of insulator function.

We evaluated the LINE1 promoter activity *in vitro*, in transient transfection assays using new luciferase reporter constructs (Figure 7). We cloned the LINE1 promoter (L1 fragment, Figure 5A) upstream of a promoter-less construct (pLuc), in the 5'-3' orientation (reverse from the native genomic orientation), thus maintaining the same transcriptional orientation for both the LINE1 promoter and the luciferase reporter gene. We could not detect promoter activity in mouse fibroblasts L929 or B16 melanoma cells (Figure 7). However, the L1 fragment, cloned upstream of a minimal promoter (pTLuc), could weakly alter the basal transcriptional activity in B16 cells ($p=0.001$), less apparently seen in L929 cells ($p=0.079$), suggesting a poorly or non-functional promoter function of the CpG island associated with the LINE1 promoter *in vitro* (Figure 7).

Table II. Statistical analysis of transgenic flies data

A: Heterozygous individuals.

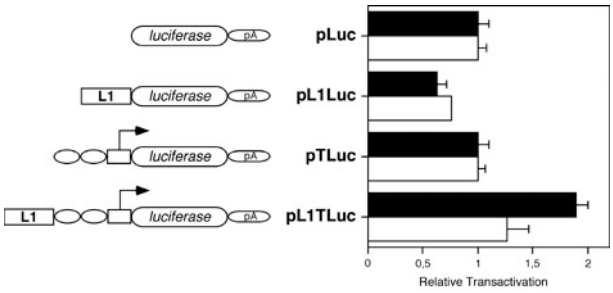
(*)	w	BRwBR	5'HS4w5'HS4	TE3wTE3	XwX	HSwHS	LCRwLCR	LCRmutwLCRmut
yellow	-6,35917	1,00228	-0,90500	-0,77991	-1,02038	4,60447	3,49764	-0,03993
pale orange	-0,46298	1,32284	-4,29008	-3,15679	-0,82629	0,60519	1,13768	5,67043
orange	2,62783	2,52551	-6,47824	-4,56445	1,40692	-2,16975	-0,01697	6,66915
brown	-1,15136	-0,49470	5,95274	2,35165	-0,61775	-0,08563	-0,44650	-5,50845
red	5,34568	-4,35593	5,72058	6,14950	1,05750	-2,95428	-4,17185	-6,79120

B: Homozygous individuals.

(*)	w	BRwBR	5'HS4w5'HS4	TE3wTE3	XwX	HSwHS	LCRwLCR	LCRmutwLCRmut
yellow	-1,97761	-1,62995	-3,17198	-1,60734	6,55275	4,95086	4,20467	-7,32140
pale orange	-3,89189	2,87197	0,91457	-5,38690	-0,99283	4,96280	4,11567	-2,59339
orange	1,90552	4,13552	-5,44550	-3,74484	-3,12783	0,30561	-0,50943	6,48095
brown	4,66643	3,35354	3,89972	6,09988	-2,13004	-1,24353	-1,95442	-12,69158
red	-0,70245	-8,73108	3,80319	4,63920	-0,30205	-8,97574	-5,85649	16,12542

(*) Z-values: Results of hierarchical (saturated) log-linear model using variables "transgenic construct" x "eye colour".

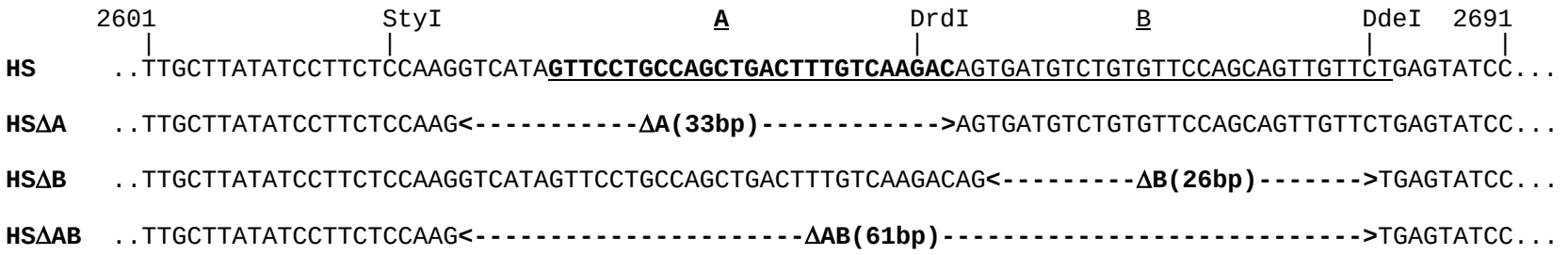
FIGURE 7



Analysis of promoter and transactivation activities of the CpG island located at the 5' end of the LINE1 element by transient transfections in mammalian cells.

Transient transfection of pLuc, pL1Luc, pTLuc and pL1TLuc plasmids in B16 (black bars) and L929 (white bars) cells. L1 DNA fragment corresponds to the CpG island and promoter fragment present within the 5' end of the LINE1 located upstream of the mouse tyrosinase LCR (indicated in Figure 5A). Schemes of these plasmids (left) and the result of transfection in B16 and L929 cells (right) are shown. Results are expressed as relative transactivation, arbitrarily assigning to backbone plasmids (pLuc and pTLuc respectively) the value of x1 and thereafter referring the activity of each construct to its corresponding backbone plasmid. Normalisation of the luciferase reporter values between different transfected constructs is achieved taking into account the activity of a co-transfected *lacZ* reporter plasmid and the number of pmols of experimental plasmid DNA used in each transfection. In pLuc and pL1Luc luciferase constructs, reporter expression values were close to background. See methods for plasmid sizes. Relative transactivations are mean values from triplicate experiments (+/- SD).

Figure 8



DNA sequence alignment of deletion HSΔA, HSΔB and HSΔAB DNA constructs. The original mouse tyrosinase LCR sequence is depicted above displaying the position of box A (**underlined and bold**) and box B (underlined), as described before (Ganss et al., 1994). HSΔA, HSΔB and HSΔAB mutants are shown below with the number of deleted nucleotides indicated. Restriction enzyme sites used for the generation of these mutants are shown, namely: *Sty* I, *Drd* I and *Dde* I. ΔA, nucleotides 2622 to 2654; ΔB, nucleotides 2657 to 2682, and; ΔAB, nucleotides 2622 to 2682. Nucleotide coordinates refer to the entire 3.7 kb Eco RI fragment (EMBL Database, accession number X76647).