Escape from X chromosome inactivation is an intrinsic property of the *Jarid1c* locus

Nan Li[a,b] and Laura Carrel[a,1]

[a]Department of Biochemistry and Molecular Biology and [b]Intercollege Graduate Program in Genetics, Pennsylvania State College of Medicine, Hershey, PA 17033

Edited by Stanley M. Gartler, University of Washington, Seattle, WA, and approved September 23, 2008 (received for review August 8, 2008)

Although most genes on one X chromosome in mammalian females are silenced by X inactivation, some ''escape'' X inactivation and are expressed from both active and inactive Xs. How these escape genes are transcribed from a largely inactivated chromosome is not fully understood, but underlying genomic sequences are likely involved. We developed a transgene approach to ask whether an escape locus is autonomous or is instead influenced by X chromosome location. Two BACs carrying the mouse *Jarid1c* gene and adjacent X-inactivated transcripts were randomly integrated into mouse XX embryonic stem cells. Four lines with single-copy, X-linked transgenes were identified, and each was inserted into regions that are normally X-inactivated. As expected for genes that are normally subject to X inactivation, transgene transcripts *Tspyl2* and *Iqsec2* were X-inactivated. However, allelic expression and RNA/DNA FISH indicate that transgenic *Jarid1c* escapes X inactivation. Therefore, transgenes at 4 different X locations recapitulate endogenous inactive X expression patterns. We conclude that escape from X inactivation is an intrinsic feature of the *Jarid1c* locus and functionally delimit this escape domain to the 112-kb maximum overlap of the BACs tested. Additionally, although extensive chromatin differences normally distinguish active and inactive loci, unmodified BACs direct proper inactive X expression patterns, establishing that primary DNA sequence alone, in a chromosome position-independent manner, is sufficient to determine X chromosome inactivation status. This transgene approach will enable further dissection of key elements of escape domains and allow rigorous testing of specific genomic sequences on inactive X expression.

epigenetics | dosage compensation | transgene

In female mammals, one X chromosome is inactivated in early embryogenesis to equalize X dosage between the sexes (1). Initiation of X chromosome inactivation (XCI) requires a locus that includes the *Xist* gene. At the onset of XCI, *Xist* RNA coats the inactive X, the X is epigenetically modified, and gene silencing is established and then maintained for all subsequent cell divisions (1). Despite the chromosomal nature of X inactivation, some genes ''escape'' XCI and are expressed from both active and inactive Xs (2, 3). How these escape genes remain expressed on the largely inactivated X is an important question that is not yet completely understood.

Many human escape genes cluster (3, 4), suggesting that they are organized in coordinately controlled domains. Intriguingly, mouse differs; escape genes are interspersed among inactivated genes (1, 5). Nevertheless, higher-order chromatin may be similar as several boundaries between escape and inactivated genes in both mice and humans have insulators bound by the CCCTC binding factor protein (CTCF) (6). CTCF is proposed to isolate escape genes from the surrounding inactive heterochromatin (6). However, CTCF binding alone is not sufficient for escape gene expression because a reporter gene flanked by CTCF binding sites was silenced by XCI (7).

What other factors may regulate escape gene expression? X-inactivated and escape domains differ in sequence composition, particularly repetitive element distribution (3, 5, 8–11). Sequences on the X are hypothesized to propagate XCI (12) and to be depleted at escape genes (8, 9). LINE-1 repeats fit such predictions, particularly on the human X (8–10). Distinct distributions of other repeats classify some mouse X genes (5).

X-linked transgenes also test the role of genomic sequences in escape gene expression. Most transgenes are X-inactivated, although a number escape XCI (e.g., refs. 13 and 14). Such transgene studies indicate that, in addition to CTCF (6), locus control regions and matrix attachment sites are also not sufficient to escape XCI (15, 16). It is unclear why some transgenes do escape XCI, because each experiment assesses different transgenes integrated at different X locations. Furthermore, these transgenes do not originate from the X, and it is not clear how nonmammalian sequences or mammalian autosomal genes should respond to XCI. These data notwithstanding, at least 8 transgenes integrated into the X-inactivated *Hprt* locus are subject to XCI (e.g., refs. 7 and 15), suggesting that chromosome location may profoundly influence inactive X expression regardless of transgene composition.

To better understand escape gene regulation, we established a transgene system in mouse XX embryonic stem cells and asked whether a domain that escapes XCI is autonomous or instead takes on properties of the region on the X in which it is located. We specifically asked whether the escape gene *Jarid1c*, in the context of adjacent genomic sequences, would retain its expression pattern upon relocation on the X.

Results

Isolation of Single-Copy X-Linked BAC Transgenes in Mouse Female ES Cells. The locus selected for these studies is well characterized with respect to XCI (5, 6, 17, 18). This region contains the escape gene *Jarid1c* (formerly *Smcx*) flanked by X-inactivated transcripts (5). Large BACs were selected to include potential distant regulatory sequences. BACs RP23-330G24 and RP23-391D18 incorporate the entire *Jarid1c* gene and adjacent transcripts (Fig. 1*A*).

Transgenes were assessed in female mouse ES cells, a well established *ex vivo* model for XCI studies (1). Using this system, sequences can be introduced onto active X chromosomes in undifferentiated ES cells and will insert into sites unrelated to XCI response. Subsequently, transgenes are monitored after ES cell differentiation and XCI. To investigate transgenes at multiple chromosome locations, BAC DNA was transfected into undifferentiated ES cells and lines were screened by FISH to identify random integrants on the X (Fig. 1*B*). Ten of 185 cell

Author contributions: N.L. and L.C. designed research; N.L. and L.C. performed research; N.L. and L.C. analyzed data; and N.L. and L.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed at: Department of Biochemistry and Molecular Biology, Pennsylvania State College of Medicine, Hershey Medical Center, C5757, 500 University Drive, Hershey, PA 17033. E-mail: lcarrel@psu.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0807765105/DCSupplemental.

GENETICS

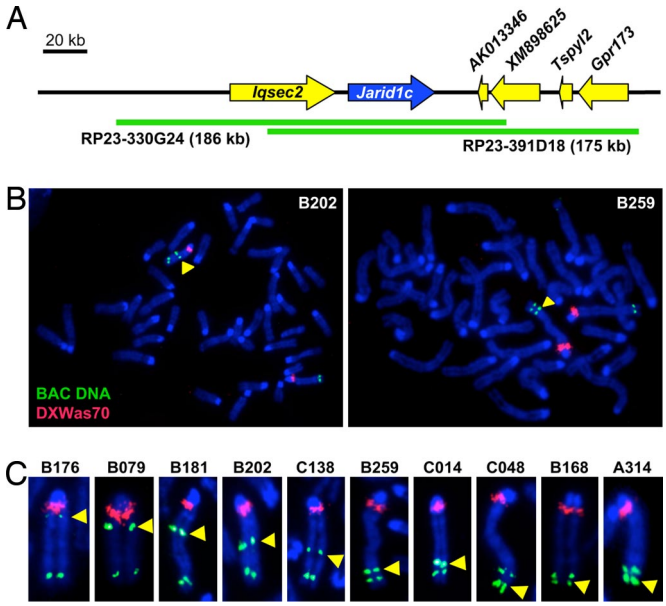


Fig. 1. Isolation of X-linked BAC transgenes. (*A*) A 300-kb region on mouse X chromosome at 148.52–148.82 Mb (UCSC Genome Browser, July 2007 assembly). *Jarid1c* escapes XCI (blue) and is surrounded by X-inactivated genes (yellow) (5). (*B*) FISH to identify X-linked integrants. BAC DNA probes hybridize to transgenic (arrowhead) and endogenous loci. (*C*) Enlarged Xs from 10 independent ES lines with transgenes (arrowheads). All transgenes were derived from BAC RP23-391D18 except A314 (RP23-330G24).

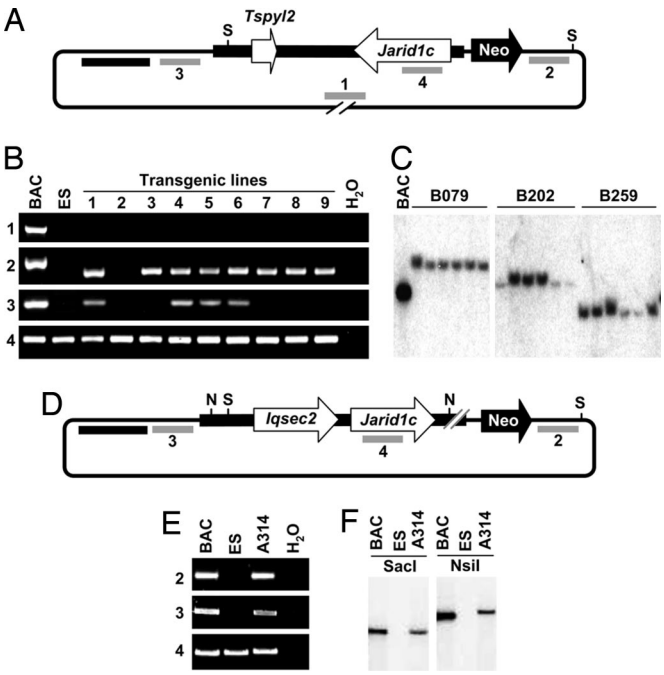


Fig. 2. Characterization of X-linked BAC transgenes. (*A*) Cartoon of BAC RP23-391D18 (not to scale) indicates relevant genes in BAC vector (thin line) and genomic insert (thick line). Location of PCR products (gray bars), Southern probe (black bar), linearization site (double hash mark), and relevant SacI restriction sites (S) are also indicated. (*B*) PCR analysis of transgenes. Transgenic lines B079, B168, B176, B181, B202, B259, C014, C048, and C138 are compared to BAC DNA (BAC) and the parental ES line. (*C*) Southern analysis to evaluate transgene copy number in 6 subclones from each line. Intensity differences between subclones reflect DNA loading differences. Similar results were seen for 2 additional restriction enzymes (data not shown). (*D*) Cartoon of BAC RP23-330G24 (not to scale) labeled as in *A* with relevant SacI (S) and NsiI (N) restriction sites indicated. (*E*) PCR analysis of clone A314. (*F*) Southern analysis indicates that the A314 transgene is single-copy. Integration site sequencing after inverse PCR confirmed that the transgene linearized 4.0 kb from the end of the genomic insert (hash marks). Consequently, by Southern blot, the NsiI but not SacI band for A314 is slightly shifted compared with purified BAC DNA.

lines carried a transgene on the X (Fig. 1*C*). Line A314 was generated from BAC RP23-330G24, and the remaining 9 lines were generated from BAC RP23-391D18.

ES lines were characterized to assess transgene integrity and copy number. PCR identified properly linearized BACs with intact vector sequences on both sides of the genomic insert (Fig. 2 *B* and *E*). By Southern, 4 of the 5 apparently intact lines had a single BAC insert (Fig. 2 *C* and *F*). Altogether, lines B079, B202, B259, and A314 carry intact or largely intact single-copy transgenes and were pursued for subsequent analysis. Although many transgenes are silenced by position effects that are unrelated to XCI response, in all 4 lines RNA FISH confirmed expression of transgenic *Jarid1c* from active Xs in undifferentiated ES cells [supporting information (SI) Fig. S1].

BAC Transgenes Inserted into Regions That Are Normally X-Inactivated. Transgene insertion sites were identified by inverse PCR (19), and locations are indicated in Table 1. To monitor XCI landscape we determined the normal XCI status of adjacent genes on a nontransgenic X chromosome. XCI status was assessed by measuring relative active and inactive X expression of a transcribed polymorphism in the nonrandomly inactivated primary fibroblast cell line B119 (17). The closest annotated genes that were expressed in fibroblasts and ES cells were assayed by using an allele-specific primer extension assay, SNaPshot (3). All genes showed monoallelic expression indicating that they are X-inactivated (Table 1 and Fig. S2). These data suggest that all 4 *Jarid1c* transgenes integrated into regions that are normally X-inactivated, with the caveat that 2 insertion sites are relatively gene-poor. Notably, genome landscape at all transgene integration sites, particularly repetitive element composition, differs from the endogenous *Jarid1c* locus (Fig. S3*A*) (5).

Transgenic *Iqsec2* and *Tspyl2* Are Properly X-Inactivated. Because either X in the ES cells can undergo XCI, before analyzing transgene expression, sequential RNA and DNA FISH was performed to determine how frequently the transgene was on the inactive X (Fig. S4). The ES line contains Xs from 2 different mouse strains, 129 and *Mus castaneus* (CAST), and upon ES cell differentiation the 129 X is inactivated in ≈75% of cells (20). Therefore, the frequency that the transgene is on the inactive X also infers strain origin. The transgenes in lines B079 and B259 were on the inactive X in >70% of cells, indicating integration onto the 129 X. The A314 and B202 transgenes were on the inactive X in <30% of cells, suggesting insertion into the CAST X (Fig. S4). For line A314, sequence obtained at the integration site contained SNPs between the 2 mouse strains and confirmed integration into the CAST X.

Will transcripts that are normally X-inactivated still be subject to XCI at an ectopic transgene location? *Iqsec2* was first tested in line A314 by sequential RNA and DNA FISH (Fig. 3 *A* and *B*). Nuclei were hybridized with an *Iqsec2* probe to detect nascent transcripts and an *Xist* probe that marks the inactive X. Subsequently, after signal fixation and denaturation, a BAC probe was used to detect both endogenous and transgene DNA loci. *Iqsec2* transgene expression was identical to the endogenous gene; in nearly all cells scored (97%), *Iqsec2* RNA signals were detected from the active X but not from the inactive X (Fig. 3*B*). These results indicate that the *Iqsec2* transgene, similar to the endogenous allele, is X-inactivated.

Table 1. Chromosomal location and XCI status of genes near transgene integration sites

Transgene line	Transgenic X strain	BAC genes assayed	Adjacent genes assayed*	Location on X,† Mb	XCI status
B079	129			36.67	
			AK139935	35.83	Inactivated
			Cul4b	35.93	Inactivated
		Jarid1c			Escape
		Tspyl2			Inactivated
			Gria3	38.75	Inactivated
			Birc4	39.45	Inactivated
B202	CAST			88.19	
			EG547215	87.31	Inactivated
		Jarid1c			Escape
B259	129			131.95	
			Zmat1	131.54	Inactivated
		Jarid1c			Escape
		Tspyl2			Inactivated
			Gprasp1	132.28	Inactivated
A314	CAST			166.36‡	
			Mid1	166.12	Inactivated
		Jarid1c			Escape
		Iqsec2			Inactivated

*The closest single-copy annotated genes with expressed SNPs that were expressed in fibroblasts were tested. Using these criteria, very few transcripts were excluded except in the B259 line. At least six transcripts near this gene-rich integration site showed tissue-restricted expression or were multicopy.
†Map locations are from the UCSC Genome Browser (http://genome.ucsc.edu/), build 37, July 2007 assembly.
‡A314 integrated into exon 2 of the *Mid1* gene.

By RNA FISH, *Tspyl2* transcripts were detected from active Xs in only a small proportion of cells (<5%). Inactive X *Tspyl2* transcripts for both the endogenous and transgene loci were absent in these cells, suggesting that both are X-inactivated (an example is shown in Fig. 3*A*). However, with few scorable cells, results did not approach statistical significance (data not shown).

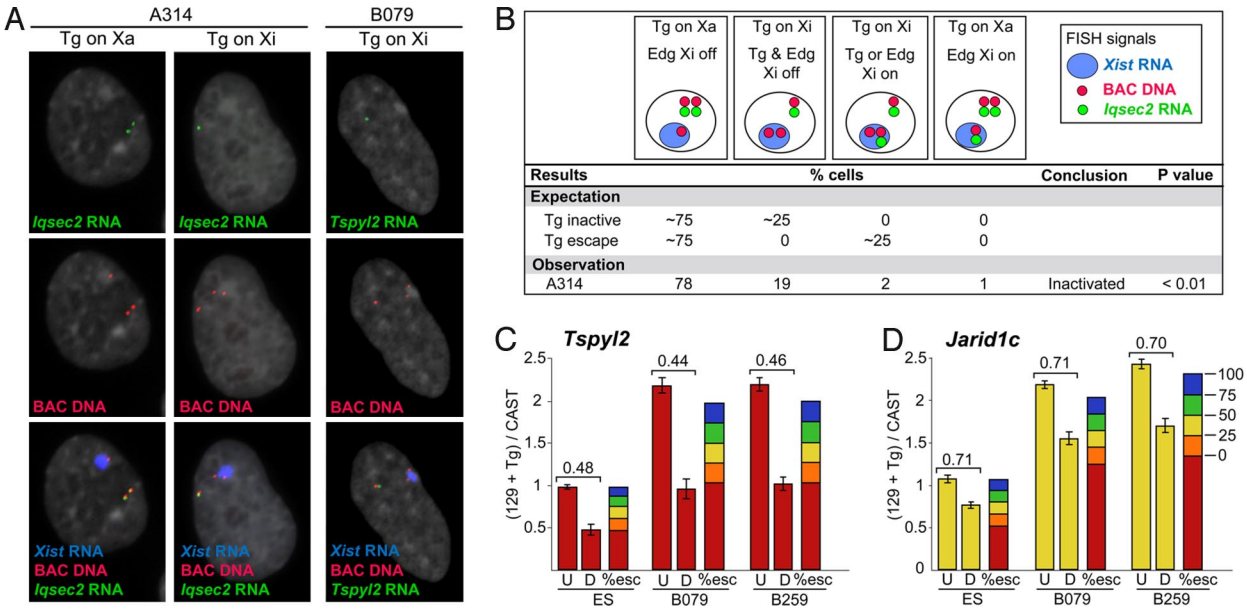
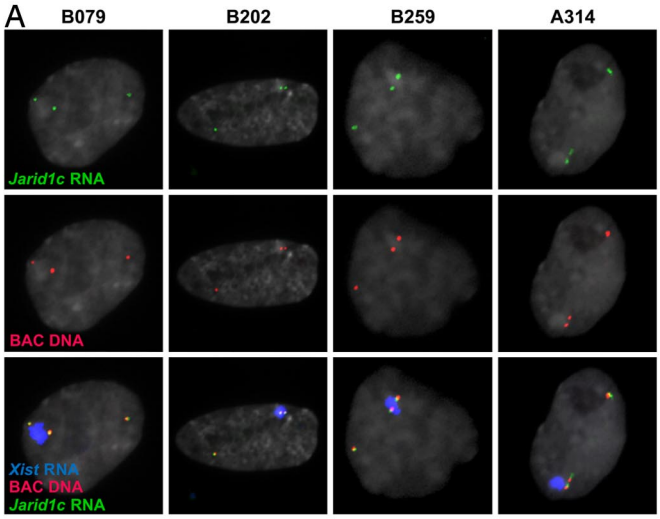


Fig. 3. X inactivation status of transgene transcripts. (*A*) Representative nuclei from sequential RNA and DNA FISH. *Iqsec2* transgene expression was tested in line A314, and examples are shown with transgene on the active X (Xa) and inactive X (Xi). *Tspyl2* expression is also shown in line B079 with the transgene on the Xi. (*B*) Summary of *Iqsec2* FISH results. Only *Xist*-positive cells with 3 clear BAC DNA foci and *Iqsec2* RNA transcripts from all active X loci were scored. The hybridization patterns scored are denoted with transgene (Tg) location and XCI status of transgene and/or endogenous (Edg) locus. The percentage of cells showing each pattern was compared with expectations for a transgene that is inactivated or escapes XCI (>100 nuclei scored). (*C* and *D*) X inactivation status of transgenic *Tspyl2* (*C*) and *Jarid1c* (*D*) in lines B079 and B259 was evaluated by allelic expression using a SNP that differentiates the 129 and transgene alleles from the CAST allele. Undifferentiated (U) and enriched differentiated (D) cells were tested, and relative allelic expression levels are shown. The ratio between U and D for each line is indicated. The expected allelic expression ratios corresponding to different levels of transgene inactive X expression (%esc indicated for 0%, 25%, 50%, 75%, and 100% of active X levels) are shown for comparison with observed results. Experiments were performed in triplicate, and standard deviations are indicated.

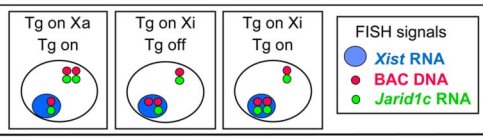
Therefore, XCI status of *Tspyl2* was assessed by comparing allelic expression in transgenic and nontransgenic cell lines. Relative allelic expression was first quantitated in undifferentiated cells (Fig. 3*C*). As expected for lines with 2 active X chromosomes, equivalent expression was seen for the 129 and CAST alleles in nontransgenic ES lines. The polymorphism assayed does not distinguish the endogenous 129 allele from the C57BL/6-derived transgene allele, but both differ from the CAST allele. The 2 lines tested have transgenes on the inactive 129 X in a large percentage of cells (≈75%), allowing inactive X transgene expression to be effectively detected. We compared transgene and 129 expression on one X to CAST expression on the other. In transgenic lines, the combined expression level of the transgene and 129 allele was approximately twice that of the CAST allele (Fig. 3*C*), suggesting that transgenic *Tspyl2* is similarly expressed before XCI. Expression was then tested in enriched populations of differentiated ES cells in which most cells (≈83%) retain two X chromosomes and have undergone XCI. Relative expression levels also reflect that the 129 allele is on the inactive X in only ≈75% of cells (Fig. S4). Using these values in nontransgenic enriched differentiated cells, the escape level, or the amount of 129 allelic expression from the inactive X, could be determined and was compared with expected values. This approach confirmed that *Tspyl2* is X-inactivated (5). Importantly, in the 2 transgenic lines tested, the ratio of enriched differentiated cells to undifferentiated cells was highly similar to the ratio seen for the nontransgenic line, implying that the transgene had the same XCI status as the endogenous allele. Comparison of observed with expected ratios calculated for different levels of XCI escape also suggests that the transgene was X-inactivated (Fig. 3*C*). Therefore, we conclude that transgenic *Tspyl2* in lines B079 and B259 is subject to XCI.

***Jarid1c* Transgenes Escape X Inactivation.** We next asked whether transgenic *Jarid1c* was subject to XCI or, like the endogenous locus, would escape XCI. *Jarid1c* transgenes in lines B079 and B259 were analyzed in the same manner as *Tspyl2*, by comparing allelic expression in undifferentiated and differentiated cells using a polymorphism that distinguishes transgene and 129 alleles from the CAST allele. As seen for *Tspyl2*, comparison of undifferentiated transgenic and nontransgenic cells suggests that before XCI, transgenic *Jarid1c* was expressed at levels that approximate either endogenous allele (Fig. 3*D*). Endogenous *Jarid1c* partially escapes XCI, with inactive X expression levels at ≈25–50% of active X levels (17, 21). Allelic expression was tested upon differentiation and XCI. As expected, in nontransgenic ES cells, the relative ratio of 129 to CAST *Jarid1c* expression was higher than for *Tspyl2* but not identical to that in undifferentiated cells (Fig. 3 *C* and *D*). These data are consistent with *Jarid1c* partially escaping XCI, at levels calculated to be 40% of active X levels (Fig. 3*D*), similar to previous reports (17, 21). Importantly, for the transgenic lines, the allelic expression ratio between differentiated cells and undifferentiated cells was highly similar to the nontransgenic line, supporting the conclusion that transgenic *Jarid1c* escapes XCI with inactive X expression levels predicted to be ≈36% of active X levels (Fig. 3*D*).

To confirm and extend these results, we directly examined transgenic *Jarid1c* expression in differentiated cells by sequential RNA and DNA FISH. At least 86% of nontransgenic cells showed *Jarid1c* nascent transcripts from the inactive X (Fig. S5). Transgenic lines were tested and scored using criteria that ensured that only diploid nuclei with optimal hybridization were analyzed. For all 4 transgenic cell lines, nascent *Jarid1c* RNA transcripts were identified at each DNA locus (2 endogenous and 1 transgenic loci) in >93% of cells (Fig. 4). These results clearly demonstrate that transgenic *Jarid1c* at 4 ectopic locations escapes XCI in all, or at least the vast majority of, cells.



B



Results Expected & Observed		% cells		Conclusion	P value
Tg on Cast X					
Expectation					
Inactivated	~75	~25	0		
Escape	~75	0	~25		
Observation					
A314	71	6	23	Escape	< 0.01
B202	78	7	15	Escape	< 0.05
Tg on 129 X					
Expectation					
Inactivated	~25	~75	0		
Escape	~25	0	~75		
Observation					
B079	23	4	73	Escape	< 0.01
B259	27	6	68	Escape	< 0.01

Fig. 4. Sequential RNA and DNA FISH analysis of *Jarid1c* transgene expression. (*A*) Representative nuclei. (*B*) Summary of FISH results. To ensure that only diploid cells with optimal hybridization signals were considered, scored cells showed 1 *Xist* signal, 3 *Jarid1c* DNA foci (1 transgenic and 2 endogenous), and *Jarid1c* transcripts from all active X loci and at least 1 inactive X locus. The percentage of cells showing each hybridization pattern was compared with expectations for a transgene that is subject to or escapes XCI.

Discussion

BACs and other large-insert constructs have been instrumental for transgenic animal models because they include long-range elements that orchestrate proper spatial and temporal expression (22). BAC transgenes have modeled epigenetic events such as genomic imprinting (23). Nevertheless, even large inserts can be susceptible to integration site effects (24). In this study we extend the use of BAC transgenes to evaluate escape genes, because inactive X expression patterns are likely influenced by genomic sequence environment (3, 9–11). We established that transgenic *Jarid1c* escapes XCI at all 4 sites tested, whereas other transgene transcripts were appropriately X-inactivated. We conclude that escape from XCI is an intrinsic property of the *Jarid1c* locus, requiring sequences in the transgenes. To our knowledge, these studies are the first to functionally define an escape domain derived from an endogenous X-linked locus and give important insight into mechanisms of XCI.

Inactive and escape genes are distinguished by chromatin differences that include DNA methylation and histone modification (1). Reports differ as to whether *Jarid1c* escapes from the onset of XCI or is initially inactivated (25, 26), although, notably,

the *Jarid1c* promoter remains unmethylated throughout the onset and establishment of XCI (6). It is unclear what initially differentiates this promoter from others that become heavily methylated and transcriptionally inactive. Our transgenes were generated from BAC DNA lacking epigenetic modification, and appropriate XCI expression patterns were established. These data indicate that DNA sequence alone, in a chromosome position-independent manner, sufficiently directs Xi expression.

What transgene sequences drive proper inactive X regulation? Escape gene expression likely involves both gene-specific and long-range regulatory elements (1, 5, 10, 18). The BACs overlap to share 112 kb (Fig. 1*A*), and it is likely that this 112 kb defines the maximum *Jarid1c* escape domain. Because repetitive element distribution correlates with XCI response (3, 9–11), it is notable that repeat composition at *Jarid1c* is unique compared with the transgene integration sites, the rest of the X, and the mouse genome (Fig. S3*A*) (5). This suggests that the sequence composition of each BAC is adequate to direct inactive X expression, although this may not be surprising because the transgenes used in these studies are quite large. Nonetheless, these studies now delimit the amount of genomic sequence that may influence X expression.

Transcripts immediately adjacent to *Jarid1c* are X-inactivated, and, therefore, repetitive sequence environment alone cannot explain why genes escape XCI. It is hypothesized that escape genes reside in domains flanked by CTCF (6); CTCF sites are found in the promoter and 5′ UTR of *Jarid1c* (6) and are present in the BAC transgenes. CTCF binding sites are also predicted downstream of *Jarid1c* (http://insulatordb.utmem.edu/) (27), within the transgenes (Fig. S3*B*). Additional experimentation, aided by the transgene system reported here, will be useful to determine whether these putative binding sites and/or other promoter or long-range sites are important for escape gene regulation.

Jarid1c also assumes a more peripheral inactive X location than inactivated genes (26), suggesting that 3-dimensional architecture influences XCI status. If positioning is critical for escape gene expression, transgenic *Jarid1c* should localize similarly. Both endogenous and transgenic *Jarid1c* FISH signals were frequently seen at the periphery of the *Xist* RNA-marked inactive X domain (e.g., Fig. 4*A*), yet more critical analysis will be necessary to confirm this observation. CTCF may aid this function by tethering chromatin loops (26), but other sequences must facilitate exterior localization. Perhaps repetitive sequences are involved, because transcriptionally silenced interspersed repeats are the most internalized sequences within the inactive X domain (26, 28).

From these and other studies, it is clear that regulation on the inactive X chromosome is complex, and the transgene approach described here establishes a tractable system to directly address long-range factors influencing XCI expression. By narrowing sequences necessary for *Jarid1c* expression and functionally defining an inactive X escape domain that is autonomous and chromosome-position-independent, future work can now address the role of specific sequences in escape gene regulation.

Materials and Methods

Cell Lines and Culture Conditions. The female mouse ES cell line EL16 was kindly provided by En Li (Novartis Institute), and subclones have been used for XCI studies (e.g., ref. 20). Subclone SA13 was isolated and stably retains two Xs for >40 doublings (>94% cells are 40,XX). Cell lines were grown essentially as described (20).

To isolate enriched populations of differentiated ES cells for RT-PCR experiments, 10-day differentiated cells were trypsinized, resuspended, and plated for 10 min. Subsequently, media were changed to remove nonadherent cells, and cells were grown for an additional 5–10 days before harvesting. FISH to examine *Xist* RNA and a control X chromosome probe confirmed that ≈83% cells obtained by this method were differentiated and retained two Xs (data not shown).

BAC Modification and Transfection. A neomycin selectable marker from plasmid PL451 (provided by N. Copeland, Institute of Molecular and Cell Biology, Singapore) was inserted into each BAC vector backbone by recombineering as described (29). BAC integrity was confirmed by PCR and pulsed-field gel electrophoresis.

BAC DNA was linearized with SgrAI within the BAC backbone and was transfected into undifferentiated ES cells by electroporation (30). Drug selection (250 μg/ml G418) was initiated at 24 h, and resistant colonies were picked after 8 days. After 1 additional passage, drug selection was removed for all subsequent experiments. Substantial cell death upon reintroduction of G418, despite retention of two Xs in a high percentage of cells, argues that neomycin is not constitutively expressed in all cells and is unlikely to influence transgene expression.

FISH. Probes for DNA FISH included the X-specific repeat DXWas70 (31) and BAC DNA. Probes for RNA FISH were an *Xist* subclone that includes a 7.6-kb genomic XbaI fragment covering most of exon 1, *Jarid1c* genomic sequence from a 19-kb EcoRI subclone that encompasses exons 5–12, an *Iqsec2* genomic probe from 3 overlapping SpeI fragments that include exons 3–8, and an amplified *Tspyl2* probe that encompasses all genomic sequence for the gene. For both RNA and DNA FISH, double-stranded DNA probes were directly labeled with Alexa Fluor 488, 546, 594, or 647 by nick translation using ARES DNA labeling kits (Invitrogen).

Metaphase spreads and DNA FISH analysis were performed as described (4). For each slide, >20 spreads were examined. RNA FISH was performed essentially as described (32). Undifferentiated ES cells were harvested and cytospun onto L-polylysine (Sigma)-treated glass slides. For differentiated cells, embryoid bodies were plated and grown directly on glass slides.

Sequential RNA and DNA FISH experiments were performed by slightly modifying established methods (32). Briefly, after RNA FISH hybridization and washes, signals were fixed in 4% paraformaldehyde in PBS (15 min, 25 °C). Subsequently, slides were denatured in 70% formamide, 2× SSC for 5 min at 75 °C, and DNA FISH was then performed as described (4). RNaseA pretreatment abolished RNA signals, confirming that hybridization conditions specifically detected RNA, not DNA (Fig. S6).

Slides were analyzed on a Nikon ECLIPSE E1000 epifluorescence microscope equipped with a Hamamatsu CCD camera and ImagePro3 software (Media-Cybernetics) or a Nikon TE2000-U microscope outfitted with a Roper Scientific CCD camera and NIS elements software. Each fluorophore was captured individually, pseudocolored, and merged in Photoshop (Adobe Systems). FISH experiments were performed in duplicate, and >100 nuclei were analyzed. Results were evaluated by using the χ^2 statistic.

Inactive X Expression Analysis. RNA was isolated and cDNA was synthesized as described (3). Two strategies were used to evaluate inactive X gene expression. To assess the normal inactive X expression pattern of transcripts in nontransgenic lines, expression was tested in the B119 cell line, an early-passage primary fibroblast cell line derived from a (T16H × CAST) F_1 female newborn mouse in which the CAST X chromosome is inactive in all cells (17). Allele-specific expression was assessed by Q-SNaPshot (3). Primer sequences for all allele-specific assays are indicated in Table S1.

To determine inactive X expression for the *Jarid1c* and *Tspyl2* transgene transcripts, enriched populations of differentiated cells were isolated. Allelic expression was examined by Q-SNaPshot (3). Allelic expression ratios could be used to determine inactive X expression by normalizing for the percentage of cells that were clearly differentiated and the frequency that either X is inactivated. The ''enriched differentiated'' cells tested are mixed cell populations composed of ≈83% of cells that are differentiated and retain two X chromosomes [as determined by FISH (data not shown)]. Furthermore, for the lines tested, B259 and B079, only 75% inactivate the 129 X chromosome (Fig. S4). Assays were repeated 3 times and evaluated by 2-tailed *t* tests. Expected escape values (%esc) were calculated that correspond to varying levels (from 0% to 100%) of inactive X expression relative to levels on the active X chromosome. Because results indicated some variation between the 2 transgene lines, expected values were computed separately for each line.

ACKNOWLEDGMENTS. We thank Hunt Willard and members of the L.C. laboratory for helpful discussions and comments on the manuscript. We gratefully acknowledge Neal Copeland for plasmids and bacterial strains and En Li for the ES cell line. This project was supported in part by start-up funds from Pennsylvania State University, a seed grant from the Pennsylvania Department of Health using Tobacco Settlement Funds, and National Institutes of Health Grant R01-HD056452.

1. Heard E, Disteche CM (2006) Dosage compensation in mammals: Fine-tuning the expression of the X chromosome. *Genes Dev* 20:1848–1867.
2. Disteche CM, Filippova GN, Tsuchiya KD (2002) Escape from X inactivation. *Cytogenet Genome Res* 99:36–43.
3. Carrel L, Willard HF (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* 434:400–404.
4. Miller AP, *et al.* (1995) Three genes that escape X chromosome inactivation are clustered within a 6 Mb YAC contig and STS map in Xp11.21-p11.22. *Hum Mol Genet* 4:731–739.
5. Tsuchiya KD, *et al.* (2004) Comparative sequence and x-inactivation analyses of a domain of escape in human xp11.2 and the conserved segment in mouse. *Genome Res* 14:1275–1284.
6. Filippova GN, *et al.* (2005) Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Dev Cell* 8:31–42.
7. Ciavatta D, Kalantry S, Magnuson T, Smithies O (2006) A DNA insulator prevents repression of a targeted X-linked transgene but not its random or imprinted X inactivation. *Proc Natl Acad Sci USA* 103:9958–9963.
8. Lyon MF (1998) X-chromosome inactivation: A repeat hypothesis. *Cytogenet Cell Genet* 80:133–137.
9. Bailey JA, Carrel L, Chakravarti A, Eichler EE (2000) Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: The Lyon repeat hypothesis. *Proc Natl Acad Sci USA* 97:6634–6639.
10. Carrel L, *et al.* (2006) Genomic environment predicts expression patterns on the human inactive X chromosome. *PLoS Genet* 2:e151.
11. Wang Z, Willard HF, Mukherjee S, Furey TS (2006) Evidence of influence of genomic DNA sequence on human X chromosome inactivation. *PLoS Comput Biol* 2:e113.
12. Gartler SM, Riggs AD (1983) Mammalian X-chromosome inactivation. *Annu Rev Genet* 17:155–190.
13. Goldman MA, *et al.* (1987) A chicken transferrin gene in transgenic mice escapes X-chromosome inactivation. *Science* 236:593–595.
14. Wu H, *et al.* (1992) An X-linked human collagen transgene escapes X inactivation in a subset of cells. *Development* 116:687–695.
15. Chong S, Kontaraki J, Bonifer C, Riggs AD (2002) A functional chromatin domain does not resist X chromosome inactivation: Silencing of cLys correlates with methylation of a dual promoter-replication origin. *Mol Cell Biol* 22:4667–4676.
16. Whyatt D, *et al.* (2000) An intrinsic but cell-nonautonomous defect in GATA-1-overexpressing mouse erythroid cells. *Nature* 406:519–524.
17. Carrel L, Hunt PA, Willard HF (1996) Tissue and lineage-specific variation in inactive X chromosome expression of the murine Smcx gene. *Hum Mol Genet* 5:1361–1366.
18. Tsuchiya KD, Willard HF (2000) Chromosomal domains and escape from X inactivation: Comparative X inactivation analysis in mouse and human. *Mamm Genome* 11:849–854.
19. Li ZH, Liu DP, Liang CC (1999) Modified inverse PCR method for cloning the flanking sequences from human cell pools. *Biotechniques* 27:660–662.
20. Lee JT, Lu N (1999) Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* 99:47–57.
21. Sheardown S, Norris D, Fisher A, Brockdorff N (1996) The mouse Smcx gene exhibits developmental and tissue specific variation in degree of escape from X inactivation. *Hum Mol Genet* 5:1355–1360.
22. Heaney JD, Bronson SK (2006) Artificial chromosome-based transgenes in the study of genome function. *Mamm Genome* 17:791–807.
23. Yevtodiyenko A, Steshina EY, Farner SC, Levorse JM, Schmidt JV (2004) A 178-kb BAC transgene imprints the mouse Gtl2 gene and localizes tissue-specific regulatory elements. *Genomics* 84:277–287.
24. Alami R, *et al.* (2000) Beta-globin YAC transgenes exhibit uniform expression levels but position effect variegation in mice. *Hum Mol Genet* 9:631–636.
25. Lingenfelter PA, *et al.* (1998) Escape from X inactivation of Smcx is preceded by silencing during mouse development. *Nat Genet* 18:212–213.
26. Chaumeil J, Le Baccon P, Wutz A, Heard E (2006) A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev* 20:2223–2237.
27. Bao L, Zhou M, Cui Y (2008) CTCFBSDB: A CTCF-binding site database for characterization of vertebrate genomic insulators. *Nucleic Acids Res* 36:D83–D87.
28. Clemson CM, Hall LL, Byron M, McNeil J, Lawrence JB (2006) The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences. *Proc Natl Acad Sci USA* 103:7688–7693.
29. Lee EC, *et al.* (2001) A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* 73:56–65.
30. LePage DF, Conlon RA (2006) Animal models for disease: Knockout, knock-in, and conditional mutant mice. *Methods Mol Med* 129:41–67.
31. Disteche CM, Gandy SL, Adler DA (1987) Translocation and amplification of an X-chromosome DNA repeat in inbred strains of mice. *Nucleic Acids Res* 15:4393–4401.
32. Clemson CM, McNeil JA, Willard HF, Lawrence JB (1996) XIST RNA paints the inactive X chromosome at interphase: Evidence for a novel RNA involved in nuclear/chromosome structure. *J Cell Biol* 132:259–275.