PLOS GENETICS

Independent domains for recruitment of PRC1 and PRC2 by human XIST

Thomas Dixon-McDougall[ID], **Carolyn J. Brown**[ID]*

Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada

* carolyn.brown@ubc.ca

Abstract

XIST establishes inactivation across its chromosome of origin, even when expressed from autosomal transgenes. To identify the regions of human XIST essential for recruiting heterochromatic marks we generated a series of overlapping deletions in an autosomal inducible XIST transgene present in 8p of the HT1080 male fibrosarcoma cell line. We examined the ability of each construct to enrich its unified XIST territory with the histone marks established by PRC1 and PRC2 as well as the heterochromatin factors MacroH2A and SMCHD1. Chromatin enrichment of ubH2A by PRC1 required four distinct regions of XIST, and these were completely distinct from the two domains crucial for enrichment of H3K27me3 by PRC2. Both the domains required, as well as the impact of PRC1 and PRC2 inhibitors, suggest that PRC1 is required for SMCHD1 while PRC2 function is necessary for MacroH2A recruitment, although incomplete overlap of regions implicates roles for additional factors. This cooperativity between factors contributes to the requirement for multiple separate domains being required for each feature examined. The independence of the PRC1/PRC2 pathways was observed when XIST was expressed both autosomally or from the X chromosome suggesting that these observations are not purely a result of the context in which XIST operates. Although independent domains were required for the PRC1 and PRC2 pathways overall all regions tested were important for some aspect of XIST functionality, demonstrating both modularity and cooperativity across the XIST lncRNA.

Author summary

XIST functions as a long, non-protein coding, RNA to initiate various pathways for the silencing of one of the two X chromosomes in female placental mammals. CRISPR-directed mutations of an inducible human XIST construct in somatic cells allowed us to discover which regions of the RNA are required for chromatin modification and protein recruitment. This was the first large-scale dissection of human XIST domains, and every function assessed was dependent on multiple regions of XIST, suggesting considerable interactions between domains of XIST. We observed similarities, but also differences, with the domains previously identified in mouse Xist and demonstrated the presence of independent pathways for chromosome reorganization in humans as well as ascribing new functionality to regions of XIST. The ability of XIST to inactivate large sections of

chromosomes from which it is expressed makes it both an exciting potential therapeutic for chromosome number abnormalities as well as a paradigm for how non-coding RNA genes are able to regulate cellular biology.

Introduction

XIST was one of the first long noncoding RNAs (lncRNA) to be identified [1], and its role in establishing the complex heterochromatin of the inactive X chromosome (Xi) continues to yield new insights into the process of X-chromosome inactivation (XCI) and the functionality of lncRNAs. Most studies of Xist have been conducted in mice; yet we know there are substantial differences in both the RNA and the XCI process between species [2]. Human and mouse XIST/Xist show a similar exon/intron structure, including two major splicing isoforms, and ~67% sequence conservation across the 15–19 kb lncRNAs [3]. Both contain a series of tandem repeats (labelled A-F), of which only the A repeat is highly conserved in sequence and size across eutheria [4,5]. The F and E repeats, which have limited tandem repeat structure, are also retained in both humans and mice. The most divergent repeat regions are B and C–the regions reported in mouse to be critical for recruitment of the polycomb repressive complexes PRC1 and PRC2 [6]. The human B repeat region is split into Bh and B repeats, and human has only a fraction of one C repeat, while mouse has 14 copies. Additionally, in humans there are more copies of the D repeats than in mouse [3,6,7]. Thus, the study of human XIST adds an important comparator to our growing understanding of mouse Xist functionality. In addition, roles for XIST variation in human disease will require an understanding of the functional domains of XIST, as will reducing the size of XIST to enable its potential as a 'chromosome therapeutic' to silence trisomic chromosomes [8,9].

As inactivation normally occurs early in mammalian development, many alternative model systems for the study of XIST and XCI have been described, most of them utilizing mouse models, reviewed in [10]. Model systems to study human XIST are more limited. The second X chromosome in human female ESCs is generally a mix of inactivated, eroded and active states [11]. Rare *in vivo* studies in humans emphasize differences between human and mouse XCI, with both human X chromosomes expressing XIST prior to random silencing [12]. In contrast, mouse has initial paternal inactivation that is reactivated prior to subsequent random inactivation in embryonic tissues, as well as regulatory elements (including the antisense Tsix) that are not conserved outside of rodents [13–15]. Transgenic human XIST has been used in induced pluripotent stem cells to generate very complete autosomal silencing [8], and also shown to induce many features of XCI in human somatic cells [16,17]. Recently, CRISPR-induced deletions in aneuploid 293 cells have shown the importance of some human regions for maintenance of XCI, and particularly for splicing [18].

Both the human and mouse Xi are seen to acquire repressive epigenetic alterations, with H3K27me3 and H2AK119ub1 (ubH2A) the result of PRC2 and PRC1 activity respectively, reviewed in [19]. SMCHD1 and MacroH2A are acquired later in mouse differentiation, with SMCHD1 recruitment in mouse dependent on PRC1 [20–22]. There has been considerable controversy over the order of PRC1 and PRC2 recruitment and the nature of the elements and intermediary proteins involved, reviewed in [23–25]. A series of proteomic analyses identified numerous Xist-interacting proteins, including HNRNPK, which is now believed to bind the B/C repeats and recruit PRC1, thereby enabling PRC2 recruitment [7,26–28]. Over 200 Xist-binding proteins, including other hnRNPs, splicing factors, RNA modifying factors and

chromatin modifying factors have been identified through various screening approaches [26,29,30].

We previously examined XIST induced from nine different integration sites in the somatic cell line HT1080, and while all sites recruited some UbH2A and suppressed Cot1 RNA expression, the recruitment of H3K27me3, MacroH2A and SMCHD1 was variable between integration sites and did not appear to be dependent on each other [17]. To follow up investigating the ability of XIST to modify its surrounding chromatin we chose to perform further studies using the 8p integration site, as it has shown reliable recruitment of heterochromatic marks, while also being distinct from the unique combinations of DNA elements of the X chromosome. Further, while no lethality was observed when XIST was expressed from the Xq integration site in the HT1080 cells, silencing of the hemizygous X was more likely to be disruptive to the cells and it was feared that this could confound further analysis of XIST activity.

In this study we dissected the modularity of human XIST by comparing the ability of clones with deletions tiled across the XIST RNA in our inducible construct in HT1080 cells to recruit H3K27me3, UbH2A, MacroH2a and SMCHD1. Recruitment of each mark required multiple distinct regions, and the regions involved in recruiting H3K27me3 were distinct from ubH2A. The observed independence of PRC1 and PRC2 recruitment in our inducible system was reinforced by finding that treatment with a PRC1 inhibitor impacted only PRC1 and SMCHD1, while inhibition of PRC2 did not impact ubH2A, only H3K27me3, as well as MacroH2A. While the need for PRC1 for SMCHD1 recruitment is consistent with results from mouse development, the independence of PRC1 and PRC2 recruitment demonstrates recruitment by human XIST in the HT1080 cells is clearly reliant upon different regions of the lncRNA.

Results

Induction of autosomal XIST establishes heterochromatin within 5 days

Induction of the 8p human XIST cDNA transgene in the HT1080 fibrosarcoma cell line has been shown to recruit the PRC1/2 established histone marks, ubH2A and H3K27me3 [17]. The XIST cDNA sequence remained largely transcriptionally inactive in uninduced cells due to binding of TetR to elements within the CMV promoter sequence (Fig 1A). Treatment of the cells with 1μg/ml doxycycline (dox) for five days (5ddox) resulted in the XIST RNA forming a discrete unified domain that was frequently visibly enriched for heterochromatin marks such as H3K27me3 using immunofluorescence and fluorescent *in situ* hybridization (IF-FISH) (Fig 1B). To capture cell heterogeneity and overcome operator bias we developed a method of quantifying the extent of enrichment in individual cells. We calculated the relative fluorescent intensity and variation of a given heterochromatin mark at the site of XIST RNA in the nucleus compared to a cross section of the nucleus. In each cell the relative enrichment of the heterochromatin mark at the XIST RNA locus was expressed as a z-score, calculated relative to the variability of heterochromatin density across the nucleus and XIST RNA signal (Fig 1C). Z-scores were effective means of representing the enrichment of a heterochromatin mark, as it expresses the enrichment of a chromatin mark at the XIST RNA cloud in terms of the standard deviation seen across the nucleus as a whole for that mark, allowing enrichment to be quantified and compared across cells and constructs for a specific chromatin mark. A z-score of 1 indicated that the median fluorescent intensity of a heterochromatin mark at the XIST RNA cloud was a full standard deviation above the average across the nucleus. Substantial heterogeneity was observed between cells and this was also observed for the H3K27me3 enrichment of a female cell line (IMR90), with the standard deviation of 2.4 and a median z-score for the population of 3.8 when measured across a test sample of 52 cells (S5 Table). In mouse embryonic stem cells (ESCs), enrichment of ubH2A and H3K27me3 occurs rapidly upon Xist expression
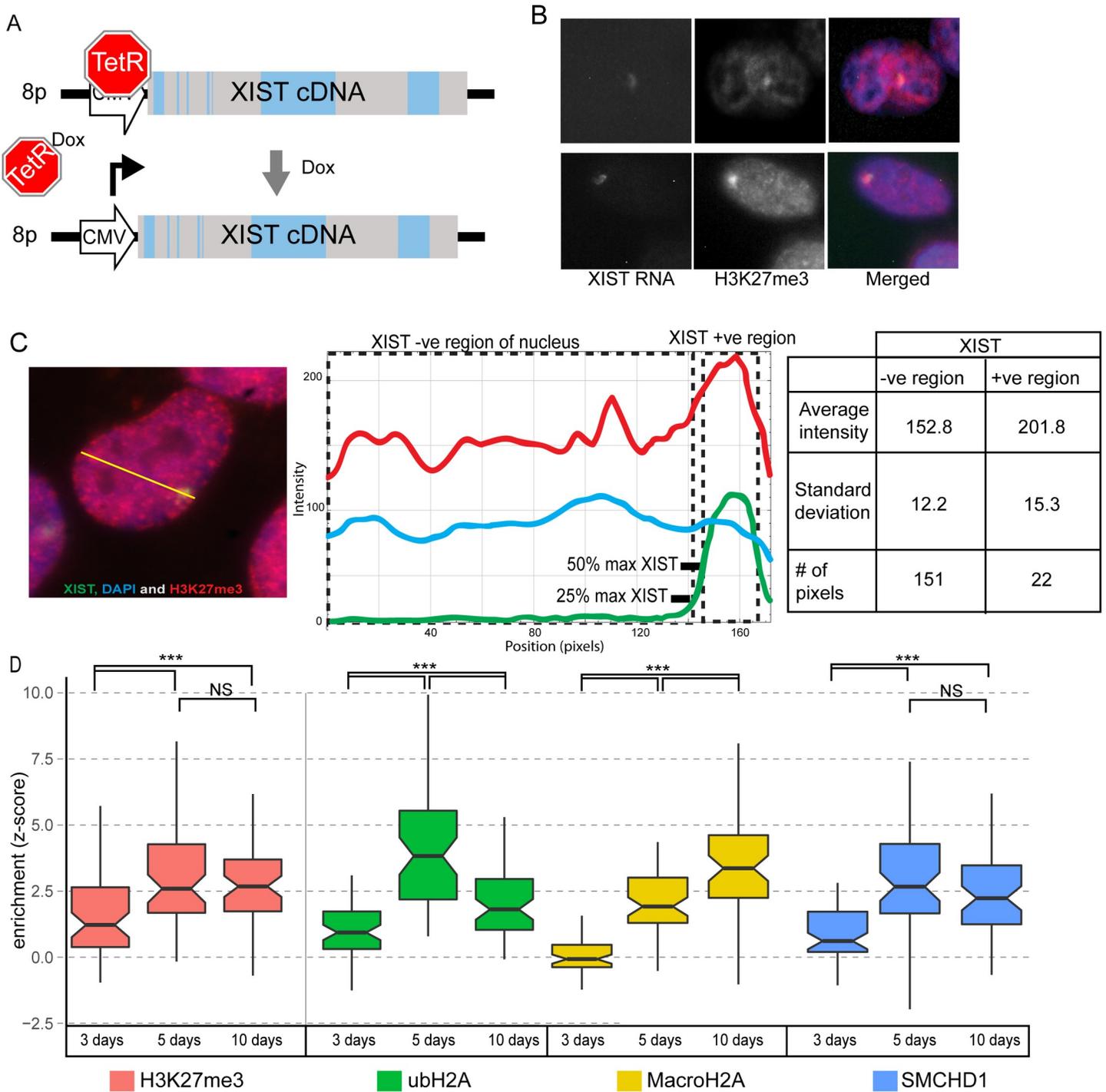
Fig 1. Autosomally expressed XIST RNA maximally enriches its surrounding chromatin with heterochromatin by day 5. A) Diagram of XIST cDNA construct under control of a Doxycycline (dox) inducible CMV promoter integrated into chromosome 8p of HT1080 cells. B) Example images of IF-FISH labelled HT1080 cells expressing XIST RNA with visible H3K27me3 enrichment. C) Schematic of how the levels of a chromatin mark (e.g. H3K27me3) were measured at the XIST RNA cloud relative to the nuclear background. The yellow line is a tangent drawn in imageJ (see Materials and Methods) for capture of pixels as drawn in the center panel with red = heterochromatin (H3K27me3), blue = DAPI and green = XIST. D) Enrichment (z-score) of H3K27me3, ubH2A, MacroH2A and SMCHD1 after different periods of XIST induction from chromosome 8p. 60 cells were measured for each condition, with the centre denoting the median, notch indicating confidence interval and box extending from the 25th to 75th percentile of each population. Significance calculated using the Mann Whitney U test (*** p < 0.001).

[31]; however, in these human differentiated cells we considered that recruitment might be slower given that epigenetic states are less plastic than in embryonic cells. Thus, to determine if marks were fully established by 5 days of induction, we calculated the relative enrichment of the PRC-associated marks H3K27me3 and ubH2A at the XIST RNA cloud following 3,5 or 10 days of dox (Fig 1D). Both marks were still accumulating following 3 days of XIST induction, with the median z-score for ubH2A enrichment 0.96 and H3K27me3 slightly higher at 1.48 across each population of 60 cells. By day 5 of dox induction the median z-score of the enrichment of H3K27me3 increased to 2.6, and remained constant after 10 days dox (z-score = 2.7). The enrichment of ubH2A increased more dramatically by day 5, with a median z-score of 4.1; however, enrichment then declined by day 10 to z-score = 1.8, which was still more than double the level at day 3 (Fig 1D). Changes were not attributable to differing levels of XIST RNA, as qPCR demonstrated that XIST RNA levels were not statistically different between 2 and 5 days of dox induction (Rq 2ddox = 0.68 relative to 5ddox, p = 0.24, S1A Fig).

In addition to examining the PRC1/2 recruited marks we wanted to examine the enrichment of SMCHD1 and MacroH2A by XIST induction. Using the same procedure to quantify the relative distribution of these marks at the XIST RNA cloud we observed a clear enrichment of both marks (Fig 1D). MacroH2A enrichment differed from the other marks in how long it required for enrichment to be observed, as after 3 days of XIST induction the median z-score for the population of cells was -0.02. Unlike the other marks examined MacroH2A had not reached its maximal level by day five (z-score = 1.95) showing a further statistically significant increase by day 10 (z-score 3.40, p = 1.1×10^{-4}). By comparison, the dynamics for SMCHD1 enrichment seemed more rapid compared to MacroH2A, as by day five of XIST induction SMCHD1 reached a peak level (z-score = 2.70, Fig 1D) that was greater than the level at day 3 (z-score 0.66, p = 2.2×10^{-7}) but did not differ significantly from the later examination at day 10 (z-score 2.27, p = 5.3×10^{-2}). These results suggested that in this model most chromatin changes induced by XIST had occurred within 5 days, with only MacroH2A continuing to accumulate beyond this point.

Repeat regions of XIST are not required for expression or localization of the lncRNA

The human and mouse XIST/Xist lncRNAs are large with functionality assigned to multiple of the mouse repeat regions. To assess the presence of functional domains within human XIST we generated a series of deletions, including not only the repeats but also non-repeat regions. Cell lines with partial XIST cDNA constructs were recombined into the 8p integration site [17]: Exon 1 consisted of all but the 3' ~3.6kb of XIST and ΔPflMI had an internal ~3.8kb removed that included Repeat B and C, and also the more diverged repeat D region (Fig 2A, [5]). In addition, an XIST deletion construct, ΔΔ, was created that removed both regions from the ΔPflMI and Exon 1 constructs. While these constructs provided a framework for examining large sections of XIST, a more complete dissection of XIST was desired to identify all potential regions crucial for the recruitment of both PRCs by XIST. Therefore, a series of nine additional deletion constructs were created by excising sections of the full length XIST construct integrated into chromosome 8p of the HT1080 cells using CRISPR (Fig 2B). These were designed with an emphasis placed on specifically excising the tandem repeat sequences separately from one another, although the human B and C repeats were too small and close to each other to be separately removed. The long non-repeat sequences of XIST were also removed individually, such that the entire XIST transgene would be interrogated for potential roles in XIST-mediated heterochromatin recruitment. The gRNAs selected to generate the various deletion constructs were designed to create small partially overlapping deletion sequences
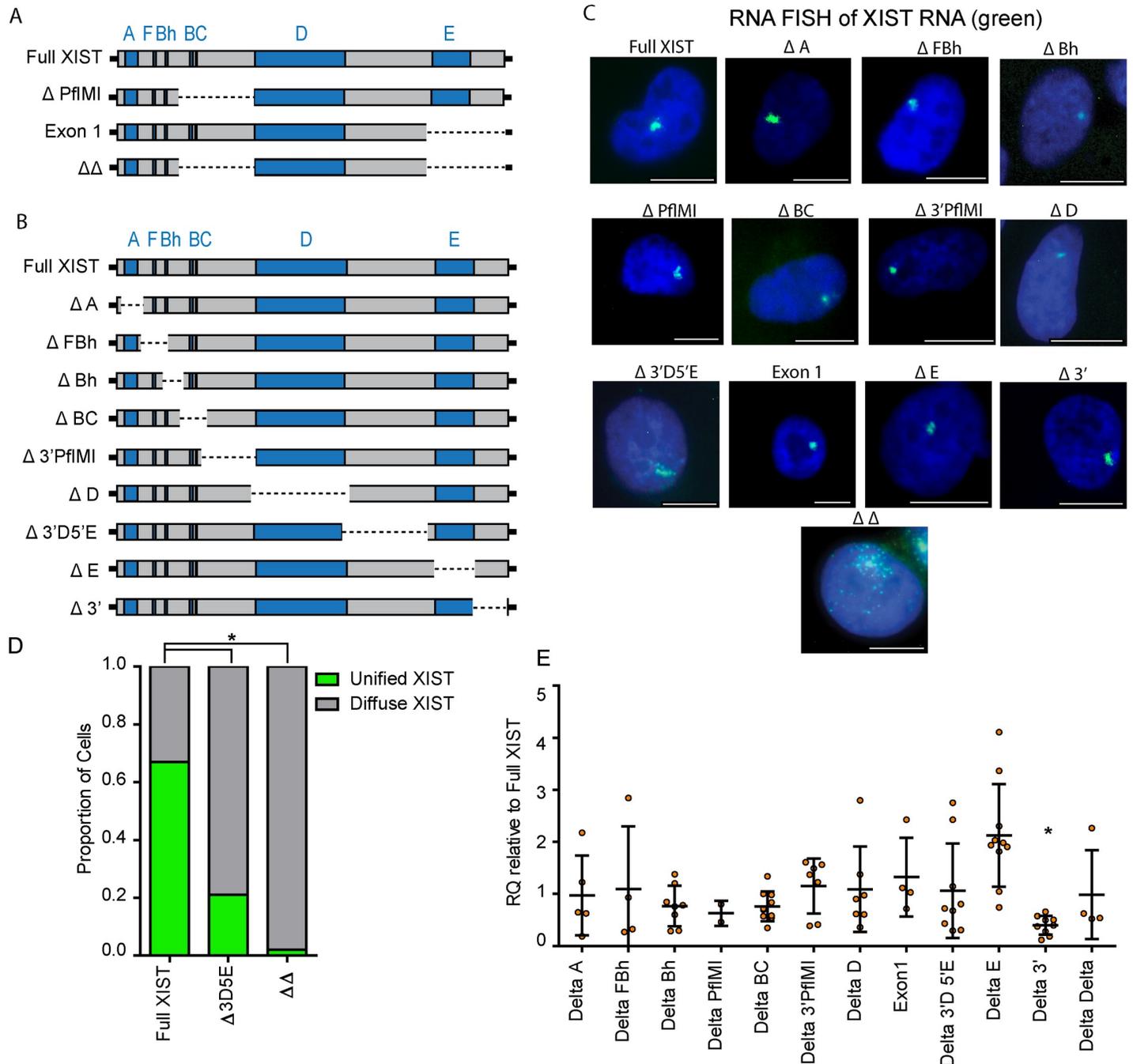
Fig 2. XIST deletion constructs used to investigate the regions of XIST crucial for chromatin remodelling. A) Illustration of the original full length and partial XIST constructs independently recombined into the FRT site on chromosome 8p of the HT1080 cell line. B) The extent of the XIST deletion constructs generated through the use of CRISPR technology from the Full XIST HT1080 cell line. C) FISH images illustrating the typical XIST RNA cloud (green) produced by each of the deletion constructs following 5ddox induction (see S2 Fig for more images). D) Quantification of the proportion of XIST RNA clouds that were either clearly unified or punctate in appearance between the Δ3D5E and ΔΔ constructs in relation to Full XIST. Proportions were compared using Fisher exact test (* p < 0.01). E) Relative XIST RNA levels across numerous biological replicates for each of the deletion constructs. XIST RNA levels were calculated using the endogenous control gene PGK1 and were expressed relative to a Full XIST control and a t-test was used to calculate statistical significance, with the threshold adjusted for multiple testing (* p = 2.05x10^{-5}).

between adjacent deletions to maximize their resolution investigating XIST (S1 Table). Independently generated clonal cell lines were isolated for each deletion construct, and the region surrounding the excision of each confirmed through sequencing (S2 Table and S1 Information; sequence across cut sites of CRISPR deletion constructs).

Inducing XIST expression with doxycycline resulted in a unified XIST RNA signal that could be readily identified when examined by IF-FISH. Most of the XIST RNA signals for each of the deletion constructs as well as Full XIST occupied similar relative space within the nucleus (S1B Fig and S3 Table) and were clearly unified into a single clearly distinct domain within the nuclei (Figs 2C and S2). Only the Delta 3' deletion construct differed from Full XIST in its relative size within the nucleus (S1B Fig, p = 2.06x10^{-18}). The female control cell line, IMR90, was observed to have a similar absolute diameter for XIST RNA clouds to the full-length construct expressed in the HT1080s (~16 pixels for each), however as the IMR90 nuclei had a noticeably greater measurable length (165 vs 102 pixels) the relative area of the IMR90 nuclei occupied by XIST was smaller (S1B Fig and S3 Table, p = 1.38x10^{-5}).

When examining the XIST RNA clouds visually the ΔΔ constructs (lacking both the PflMI region and latter exons of XIST) displayed a clearly diffuse punctate XIST signal that was immediately distinguishable from Full XIST. The only other deletion construct that noticeably affected the distribution of XIST RNA in the nucleus was the Δ 3D5E construct lacking the non-repeat sequence of XIST between Repeats D and E. Quantification demonstrated that 67% of Full XIST RNA signals across 330 cells localized into one clearly unified XIST RNA signal with no observable gaps between regions of peak signal intensity. In contrast only 21% of XIST RNA signals in 380 Δ3D5E cells were clearly unified, and only 2% of the 271 ΔΔ cells examined had a unified XIST RNA signal (Fig 2D). While beyond the scope of this current research these results suggested that redundant and potentially additive elements involved in XIST RNA unification were present across distinct regions of XIST.

The impact of the various deletion constructs on the relative levels of XIST RNA induced by 5ddox was examined (Fig 2E and S4 Table). Only the Δ 3' construct was observed to be statistically different (p = 2.0x10^{-5}), with less expression than Full XIST, suggesting that there may be some elements in that 3' region that contribute to transcript stability. The lower transcript levels of the Δ 3' region were not observed in the encompassing Exon 1 deletion construct (RQ = 1.32, p = 0.46, Fig 2E), suggesting that the effect was based on more complex mechanisms than simply the absence of stabilizing factors.

Distinct regions of XIST crucial for PRC1 and PRC2 recruitment

Our foremost objective was to identify the regions of XIST that were crucial for recruitment of the two types of PRC-established marks, H3K27me3 and ubH2A. We first examined H3K27me3 across the HT1080 Δ constructs to identify which region(s) of *XIST* were essential for recruitment and/or activation of PRC2. H3K27me3 fluorescence at the Full *XIST* RNA cloud was noticeably enriched (median z-score = 2.59, sd = 1.70) which can be conceptualized as the average H3K27me3 intensity at XIST being roughly equal to the 99th percentile of H3K27me3 fluorescent intensity measured in the nucleus (Fig 3A). Most of the deletion constructs showed comparable levels of H3K27me3 enrichment, indicating that over 80% of XIST could be removed without noticeably disrupting this process. Two regions of XIST were observed to be crucial for XIST to enrich the surrounding chromatin with H3K27me3, demonstrating a statistically significant difference from Full XIST levels of enrichment. The ΔFBh construct completely lacked H3K27me3 enrichment across its population of cells, with an average z-score of -0.147 (p = 1.11x10^{-13}, Fig 3A and S5 Table). The overlapping deletion ΔBh showed an attenuated enrichment of H3K27me3 (median z-score = 0.787) that still differed
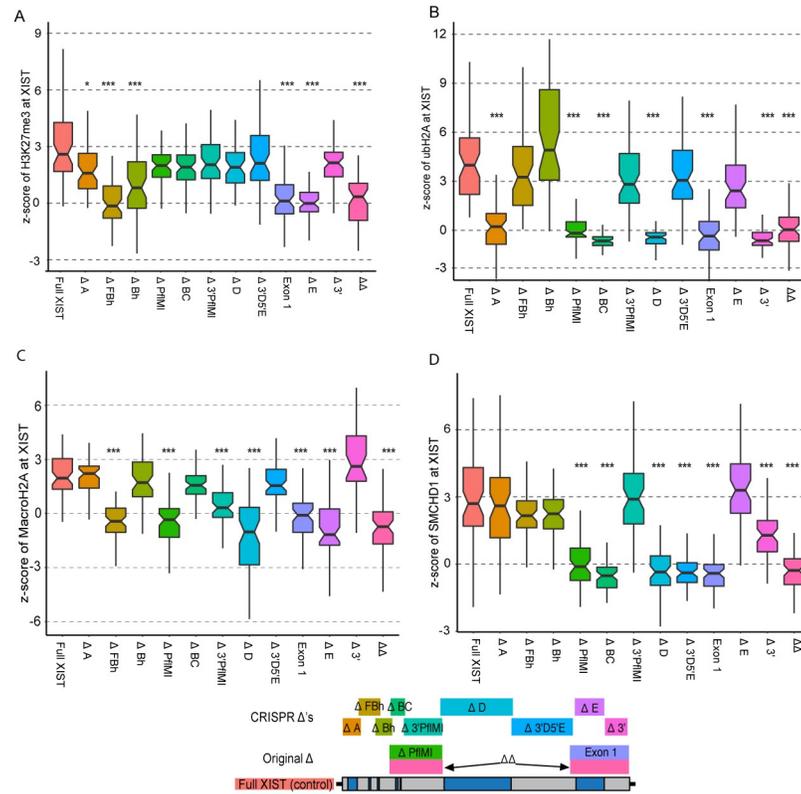
Fig 3. Multiple distinct regions of XIST are necessary for each type of chromatin modification. Each deletion construct induced for 5 days was tested for its relative enrichment (z-score) of one of the Xi associated heterochromatin marks (59–61 cells per construct). The XIST deletion constructs were tested for their ability to enrich their chromatin with A) H3K27me3, B) ubH2A, C) MacroH2A, D) SMCHD1. We were blinded to the identity of every cell analyzed in this study and only unblinded once all data had been collected into a single table. The population distribution of enrichment across the cells of each deletion construct was statistically compared to Full XIST using a Mann Whitney test with multiple testing correction (48 tests in total, * p-value < 1.04×10^{-3}, ** p-value < 2.08×10^{-4}, *** p-value < 2.08×10^{-5}). The control Full XIST cells (pink) shown in this figure are the same 5ddox treated cells shown previously in Fig 1D.

significantly from Full XIST (p = 8.47×10^{-07}, Fig 3A). The ΔA construct also had a statistically attenuated enrichment of H3K27me3 (1.59, p = 1.16×10^{-4}); however, the magnitude of this effect was clearly less extreme than for the adjacent ΔFBh construct, as ΔA and ΔFBh were strongly dissimilar (p = 1.49×10^{-9}). Repeat E of XIST was identified as a second, entirely distinct, region of XIST indispensable for H3K27me3 enrichment, as the absence of this region in the ΔE construct resulted in no observable enrichment and a strong statistical difference from Full XIST (median z-score = -0.010, p = 2.97×10^{-17}, Figs 3A and S3). The role of Repeat E was supported by the similar scores seen for the overlapping deletions of Exon 1 and ΔΔ.

Previous work has indicated a linkage between PRC1 and PRC2 recruitment in mouse models for XCI, so we investigated whether a similar link might exist for human XIST with the same regions involved in recruitment of both complexes. The deletion constructs were tested for their ability to enrich their surrounding nuclear territory with ubH2A to identify the crucial regions of XIST for PRC1 recruitment (Figs 3B and S4 and S6 Table). Full XIST showed greater relative enrichment of ubH2A than any other heterochromatin feature, although with a much greater standard deviation across the 60 cells (median z-score = 4.18 +/- 3.18). The constructs that had disrupted H3K27me3 enrichment still showed strong ubH2A enrichment,

as did constructs lacking the non-repeat regions on either side of Repeat D. Enrichment of ubH2A was found to be highly dependent upon the repeat regions in the first exon of XIST with ΔA, ΔBC and ΔD all showing no evidence of enrichment in the population of cells examined, and all strongly statistically different from Full XIST (median z-scores ≤ 0.223, p $\leq 2.84 \times 10^{-17}$). In addition, the 3' most deletion construct, Δ3', also failed to enrich chromatin with ubH2A (median z-score = -0.635, p = 4.80×10^{-21}) suggesting that this small ~630nt region was also crucial for PRC1 activity and/or recruitment. It was unlikely that this effect was due to the lower transcript levels of the Δ3' construct cell lines as the encompassing deletion of the Exon 1 constructs also failed to enrich for ubH2A (median z-score = -0.352) despite having XIST transcript levels slightly greater than Full XIST (Fig 2E). These results revealed four distinct regions of XIST crucial for the enrichment of ubH2A at the XIST RNA territory, none of which were crucial for PRC2 enrichment and *vice versa*.

Overlap in regions of XIST crucial for PRC1/2 with those for SMCHD1 and MacroH2A

Given our observation that PRC1 and PRC2 were recruited by independent domains of XIST in our human cells, we wished to determine if either of the marks that are reported to be recruited later during mouse XCI would be dependent on similar domains. MacroH2A is a well-established Xi-associated mark, known to be recruited by Xist, but had not previously been associated with other components of the XCI pathway. The formation of a Macrochromatin Body is seen late in mouse XCI, reviewed in [32] and in this model MacroH2A accumulated at a slower rate than either PRC1/2 established marks (Fig 1D). By examining which regions of XIST are crucial for MacroH2A enrichment we sought to reveal additional hierarchies for chromatin remodelling during XCI. MacroH2A had a somewhat weaker relative enrichment compared to the other Xi-associated factors described here, with the Full XIST population having a median z-score of 1.96 (Figs 3C and S5 and S7 Table). As with the other XCI factors examined, numerous distinct regions of XIST were identified as crucial for MacroH2A enrichment. The deleted regions in the ΔFBh and ΔE constructs identified as crucial for H3K27me3 enrichment were also found to be crucial for MacroH2A enrichment (median z-score ≤ -0.410, p $\leq 6.08 \times 10^{-17}$, Fig 3C). A third broad region encompassing Repeat D and the region upstream of it (Δ3'PflMI) was found to be essential for MacroH2A enrichment (median z-score ≤ 0.313, p $\leq 3.06 \times 10^{-09}$, Fig 3C). This overlap between the domains suggested potential communication between the factors, with later enrichment of MacroH2A than H3K27me3 suggesting that MacroH2A may be enabled by PRC2 recruitment or activation by XIST, with additional factors located within Repeat D and the upstream region of degenerate D repeats.

SMCHD1 recruitment by Xist has been reported to require mouse repeats B and C to recruit PRC1 [20]. Since we found ubH2A enrichment in humans to also associate with Repeats B and C, despite the divergence in size and position of these repeats between species, we questioned whether the association between PRC1 and SMCHD1 might also exist in humans. We compared the enrichment of SMCHD1 at the XIST RNA cloud in the deletion constructs relative to the Full XIST control as described in the previous sections (Figs 3D and S6 and S8 Table). The enrichment of SMCHD1 at Full XIST following 5ddox induction was found to be 2.704, and the three 5' most Δ constructs had similar levels of enrichment (median z-score ≥ 2.16), indicating that those regions that had been associated with H3K27me3 enrichment were dispensable for SMCHD1 enrichment. The ΔBC and ΔD constructs that had been unable to enrich ubH2A were also found to be incapable of enriching SMCHD1 (median z-score ≤ -0.350). The Delta 3' terminal region that had been incapable of recruiting ubH2A,

however, showed a weak but still observable enrichment of SMCHD1 (median z-score = 1.291), suggesting that the region contributed to SMCHD1 enrichment without being essential. Finally, the non-repeat region of XIST between Repeat D and E was also found to be crucial for SMCHD1 enrichment as the Δ3D5E deletion failed to enrich its nuclear territory with the factor and differed significantly compared to Full XIST (median z-score = -0.386, p = 6.35×10^{-17}). Thus, while the B-C-D and 3' regions were required for both ubH2A and SMCHD1, additional elements were involved in the recruitment of each by XIST.

Inhibitors confirm independence of PRC1/SMCHD1 from PRC2/MacroH2A

Perhaps the most surprising insight from our analysis of the XIST Δ constructs was the apparent independence of recruitment of PRC1 and PRC2 by XIST in human somatic cells. To test this observation we treated the Full XIST 8p cells with either the PRC2 small molecule inhibitor GSK343 or the PRC1 small molecule inhibitor PRT4165. Both of these inhibitors only affect the enzymatic activity of PRC2 or PRC1 rather than damaging the complexes and both inhibitors have been demonstrated to be highly specific for their target enzyme [33, 34]. Using these inhibitors would also test our hypotheses that SMCHD1 and MacroH2A were reliant on PRC1 and PRC2, respectively. For the 5 days that the cells were undergoing dox treatment to induce XIST expression we concurrently treated the cells with the inhibitors at levels reported to be effective at inhibiting either PRC2 or PRC1, but that did not cause visible signs of stress to the cells or disrupt the XIST RNA cloud.

Chemical inhibition of PRC2 with 5μM GSK343 resulted in a clear disruption of H3K27me3 enrichment compared to the uninhibited control (median z-score = 0.162, p = 7.24E-16; Figs 4A and S7 and S9 Table). These results suggested that the treatment was effective at disrupting the enrichment of H3K27me3 by XIST. We then examined whether inhibition of PRC2 affected PRC1 mediated ubH2A enrichment by XIST. In keeping with the expectation of independence based on the examination of the Δ constructs, we observed that ubH2A was still strongly enriched at the XIST RNA cloud even when PRC2 was inhibited (median z-score = 3.053, p = 0.22, Fig 4A). We also examined whether disruption of PRC2 affected MacroH2A, as the overlapping regions of XIST crucial for each suggested a potential functional link. We observed that inhibition of H3K27me3 completely disrupted MacroH2A enrichment (median z-score = 0.224) and this was strongly statistically different from the uninhibited induced Full XIST control (p = 1.30×10^{-10}, Fig 4A). Inhibition of PRC2, however, had no effect on the enrichment of SMCHD1 (z-score = 2.64, p = 0.64, Fig 4A). These observations led to the conclusion that the activity of PRC2 at the XIST RNA cloud was an essential step for the enrichment of MacroH2A, but not PRC1 or SMCHD1.

Inhibition of PRC1 with 50μM of PRT4165 had the anticipated effect of preventing the XIST RNA cloud from becoming enriched with ubH2A relative to the nuclear background (median z-score = 0.469, significance relative to Full XIST p = 5.28×10^{-18}, Figs 4A and S7 and S9 Table). Inhibition of PRC1 and the lack of ubH2A at the XIST RNA cloud did not affect the ability of XIST to enrich H3K27me3 (median z-score = 2.483), or MacroH2A (median z-score = 2.56), further supporting the independence of PRC1 and PRC2 recruitment by XIST (Fig 4A). PRC1 activity, however, was found to be crucial for SMCHD1 enrichment, as PRC1 inhibition resulted in a significant loss of SMCHD1 enrichment at the XIST RNA cloud (median z-score = -0.196, p = 5.28×10^{-18}, Fig 4A). Therefore, a role for ubH2A enrichment for the recruitment of SMCHD1 to the XIST RNA territory may have been conserved between mice and humans.
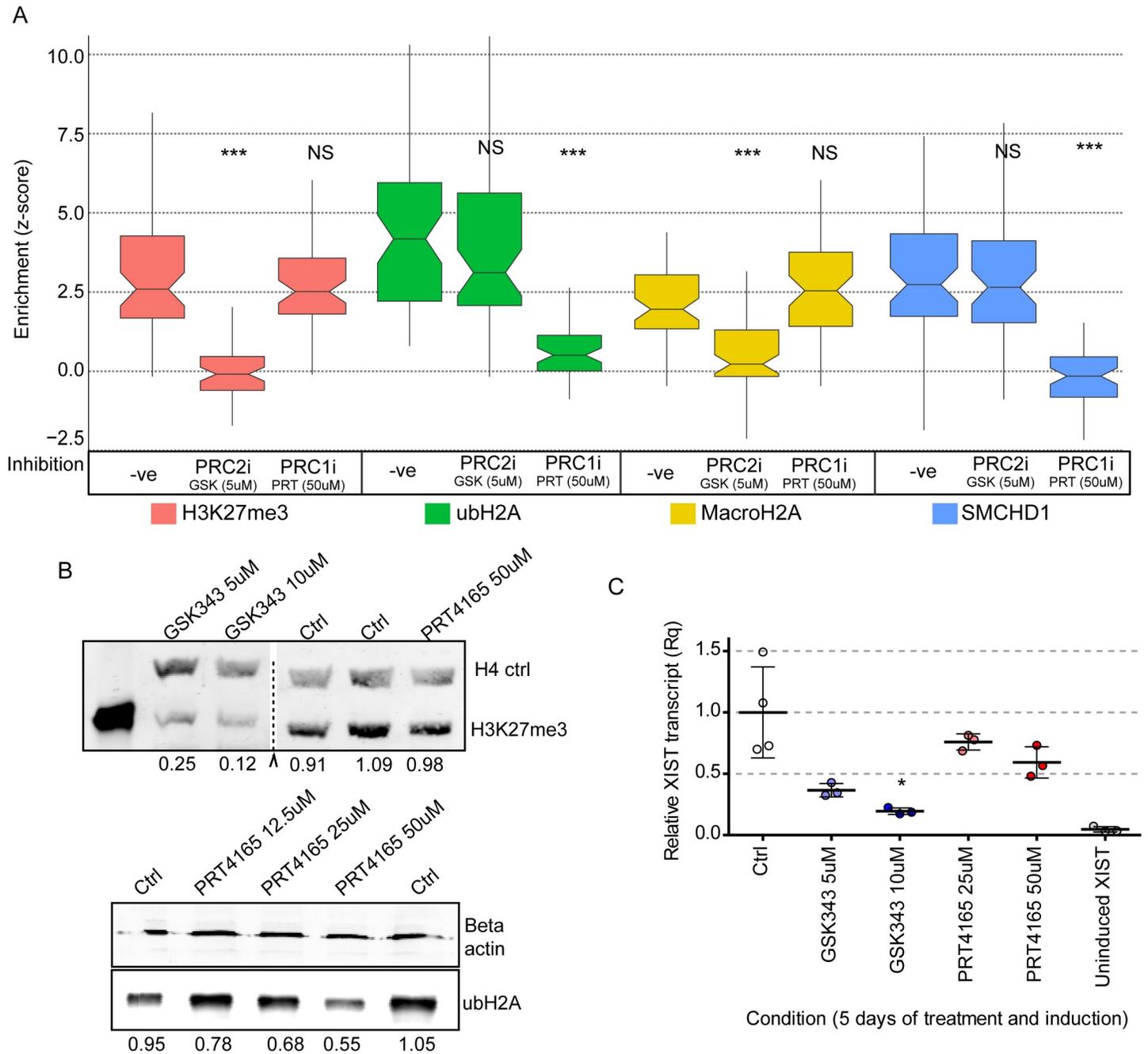
Fig 4. The two polycomb complexes operate independently to promote further heterochromatin formation. A) The effect of PRC2 inhibition with GSK343 or PRC1 inhibition with PRT4165 on the relative enrichment of H3K27me3 (pink), MacroH2A (yellow), SMCHD1 (blue) and ubH2A (green) at the XIST RNA cloud relative to the average level in the nucleus across a population of 60–61 cells. We were blinded to all cells and chromatin marks identity until after all data and calculations were completed. Statistical significance of the effect of each inhibitor on a chromatin mark was calculated using the Mann Whitney test with adjusted p value (*** $p < 1\times10\text{-}6$). The uninhibited Full XIST cells used as controls are the same 5ddox treated cells shown previously in Fig 1D. B) Western blotting images demonstrating the levels of ubH2A and H3K27me3 after cells had undergone chemical inhibition with either PRT4165 or GSK343. The dotted line denotes where several bands from an unrelated set of tests were spliced out digitally. The numbers underneath each row of bands denotes the relative concentration of the chromatin mark being tested compared to the control populations. C) Relative XIST RNA levels for each inhibitor concentration relative to control 5ddox. The dots denote independent biological replicates and statistical significance was calculated using a t-test (* $p < 0.05$).

https://doi.org/10.1371/journal.pgen.1009123.g004

Neither inhibitor completely removed its affected mark throughout the nucleus, allowing our continued use of the quantitative approach outlined in Fig 1C. We quantified the impact of the inhibitors by performing western blotting for H3K27me3 and ubH2A in the 8p Full XIST cells after 5 days of inhibition treatment and dox were compared to cells after 5 days of dox induction without inhibitors. Following 5 days of inhibition with GSK343 there was 25% the relative amount of H3K27me3 remaining in the cells compared to the control populations (Fig 4B). Following 5 days of PRC1 inhibition with PRT4165 roughly only 55% of ubH2A remained in the HT1080 cells relative to cells without the inhibitor. The PRC1 inhibition did not produce a statistically significant effect on XIST RNA levels compared to the uninhibited XIST transcript levels (Fig 4C). The PRC2 inhibitor by contrast produced a clear dose-dependant decrease in XIST transcript levels (Fig 4C). Additional concentrations of GSK343 were tested to better understand the effect of this inhibitor on XIST activity. Treatments with lower (1.0 and 2.5μM) amounts of GSK343 resulted in a weak yet technically significant change in H3K27me3 enrichment (z-score = 1.84). MacroH2A enrichment, roughly mirrored these changes but was not significantly different, suggesting that a moderate decrease in H3K27me3 levels was not strong enough to affect the recruitment of MacroH2A (z-score = 1.62, S8 Fig and S10 Table). The relative levels of XIST following PRC2 inhibition with the 5μM concentrations used in these experiments were comparable to those seen in the Δ3' cell lines, and coupled with the clear enrichment of ubH2A in both cases it seemed that this lower expression level did not cause an obvious disruption to XIST function.

XIST-mediated interactions between factors similarly observed for X chromosome

We previously demonstrated that the location of the XIST construct impacted the ability to recruit heterochromatin marks [17], and thus we analyzed the effect of PRC2 and PRC1 inhibition using an HT1080 cell line with the inducible XIST construct integrated into Xq. The X chromosome offered a more favorable context for XIST-mediated chromatin remodelling as the enrichment of all heterochromatin marks examined was greater than what had been seen in chromosome 8p (Fig 5A and S11 Table). Despite the greater magnitude of enrichment in the X chromosome, the results were similar to those with the 8p integration. PRC2 inhibition with 5μM of GSK343 resulted in a loss of H3K27me3 enrichment and MacroH2A enrichment that was strongly statistically significant but did not significantly impact ubH2A enrichment (Fig 5A and S11 Table). Conversely, inhibition of PRC1 with 50μM of PRT4165 resulted in a loss of ubH2A and as a result SMCHD1 enrichment, with no negative impact on the enrichment of H3K27me3 by XIST (Fig 5A and S11 Table). Treatment of the Xq XIST cells with GSK343 resulted in a decreased level of XIST transcripts as had been observed in the 8p cell, though the effect did not reach the threshold of significance (Fig 5B). PRT4165 treatment had no effect on XIST transcript levels (Fig 5B). Overall, our results suggest that the PRC2/MacroH2A pathway and PRC1/SMCHD1 pathway are both initiated independently of each other by XIST, and that these pathways are not a direct result of the location of the XIST gene.

Discussion

LncRNAs have been suggested to function as modular scaffolds for protein binding and the internal repeats of XIST/Xist have exemplified the concept [35,36]. In this study we sought to characterize domains of the human XIST, focussing not only on the repeat regions, but also interrogating the non-repetitive regions of the lncRNA. We generated 13 deletions that removed from 800 bp to over 3 kb of the cDNA. For each deletion we characterized the expression and localization of XIST as well as the ability to recruit H3K27me3, H2Aub, SMCHD1
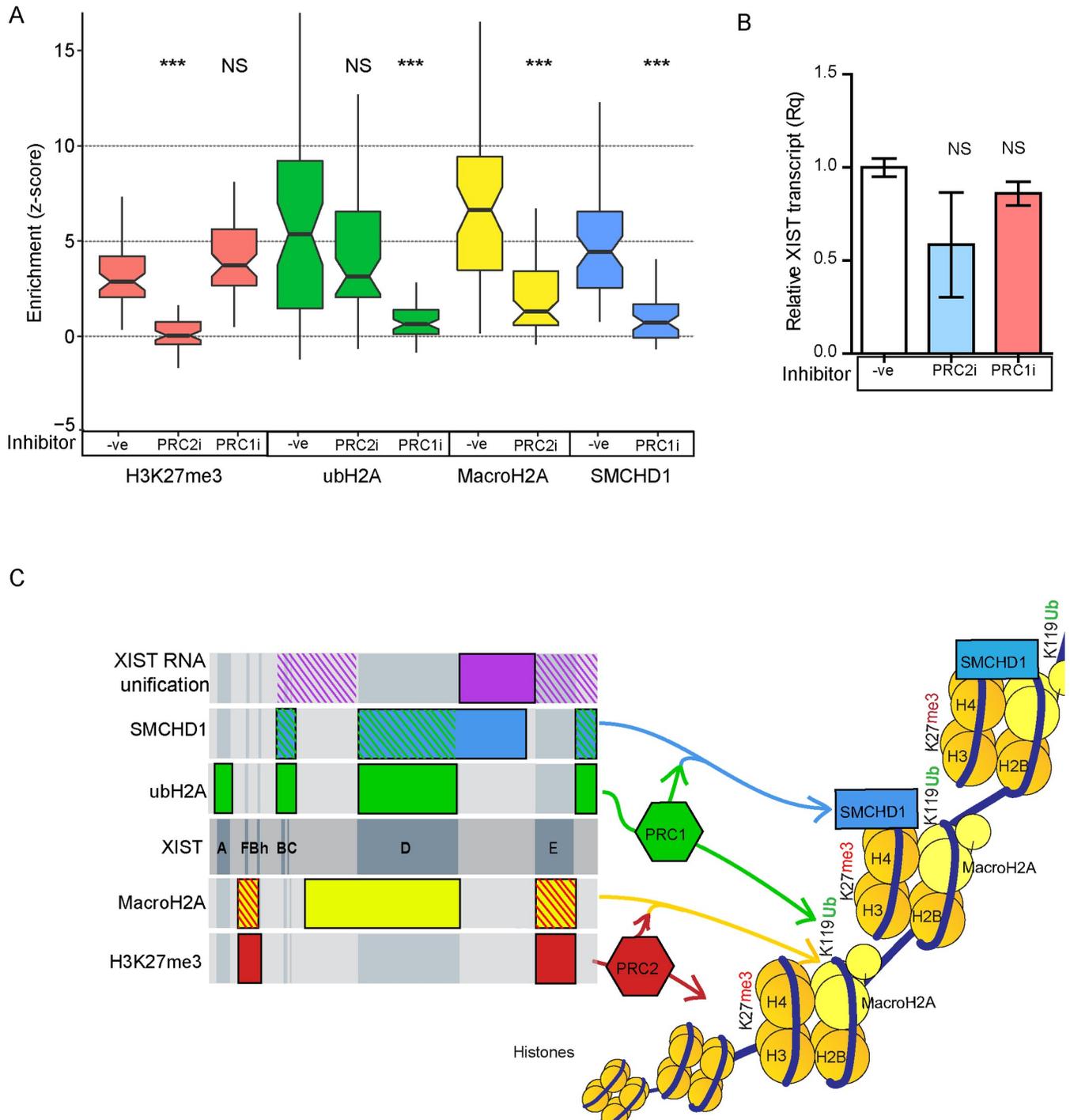
Fig 5. Testing the effect of polycomb inhibition on XIST expressed from Xq and summary of results. A) The effect of PRC2 or PRC1 inhibition on the activity of Full XIST induced from chromosome Xq in HT1080 cells. PRC2 activity was inhibited with 5uM of GSK343 and PRC1 activity was inhibited with 50uM of PRT4165. The relative enrichment of H3K27me3 (pink), MacroH2A (yellow), SMCHD1 (blue) and ubH2A (green) at the XIST RNA cloud relative to the average level in the nucleus is shown across a population of 60–61 cells. Statistical significance of the effect of each inhibitor on a chromatin mark was calculated using the Mann Whitney test with adjusted p value (*** $p < 1\text{x}10\text{-}6$). B) Relative XIST RNA levels for each inhibitor concentration relative to control 5ddox measured using qPCR and compared using a t-test relative to the uninhibited control (p = 0.05). C) Summary of the regions and pathways proposed to be crucial for XIST and PRC mediated chromatin remodelling. PRC2 (red) and PRC1 (green) were identified to operate through entirely independent regions of XIST and to not affect each other in this context. PRC2 catalytic activity was crucial for MacroH2A (yellow) enrichment while PRC1 catalytic activity was critical for SMCHD1 (blue) enrichment. Regions of XIST identified as important for the recruitment of these chromatin features were marked in the relevant solid colour, while dashed lines denoted the PRC dependent regions that overlapped with the late chromatin marks. XIST RNA unification (purple) was found to be affected by the loss of a large internal non-repeat sequence (solid colour) as well as two redundant regions of XIST (dashed lines).

and MacroH2A. As summarized in Fig 5, every feature required multiple regions of XIST to become enriched upon the XIST-coated chromatin, and there was no single critical region that was necessary for all features. While repeat-containing regions were crucial for the features analyzed it is worth noting that the traditionally less studied non-repeat regions were also found to be critical in our system.

Each feature required two or three regions of XIST separated by kilobases of RNA, that when independently deleted had no impact upon recruitment of the feature. Previous work had proposed that long-range interactions were endemic across XIST through the use of RNA duplex mapping approaches [35,37]. One could envision specific RNA secondary structures formed through long range interactions that may allow for complexes to bind to XIST, or distinct protein interactions with different RNA regions that are required to assemble essential protein complexes. While not exclusive of each other, the latter explanation may fit with evidence of five hubs of prolific protein binding, around the A repeats, F repeats, B-C-D repeats, E repeats and 3' end of XIST [35]. Studies in mouse models have argued that the multivalency of the E repeat region allows homo- and hetero-typic interactions with partial redundancy that function in promoting phase separation and formation of the inactive X compartment [38,39]. Such aggregation might be impacted by overall density of protein binding along the RNA.

Strikingly, our results showed that recruitment of the marks established by PRC1 and 2 relied upon entirely distinct domains in the human somatic cells with induced XIST expression. Similar to mouse, PRC1 recruitment required the BC region, which is considerably smaller in humans with only part of a single C repeat and a separation of the B into two clusters of repeats [3]. The D repeat region was also required for human PRC1 recruitment, consistent with structural analysis suggesting a large protein-interaction domain containing from B-C to beyond the D repeats in exon 1, as well as HNRNPK binding to the human D repeats [35]. Thus, the B-C-D region may recruit PRC1, likely through HNRNPK, with weighting in humans towards the requirement including the D repeat region as B-C is smaller than in mouse. We identified two additional regions that were also essential for PRC1 recruitment— the A repeats, and the 3' end of XIST. Again, in mouse models, it has been shown that silencing is necessary for the spread of UbH2A into genes [31], and it is possible that we only detect UbH2A when it has spread beyond initial deposition. The A repeats are also critical for binding of RBM15 for m6A deposition at the 5' and extreme 3' ends of XIST [35], so it is possible that modification of the RNA affects the ability of XIST to bind PRC1 [26,40,41].

The recruitment of H3K27me3 was impacted by only two regions—the F repeat region and the E repeat region. These regions do not overlap the B-C region seen to be necessary for murine PRC1 recruitment, nor the regions discussed above as being necessary for human PRC1 recruitment. Many regions and pathways have been reported to be responsible for recruitment of PRC2 in mouse [2], including repeats F and B for recruitment via JARID2 [42]. RNA immunoprecipitation of human XIST found EZH2 and SUZ12 bound near repeat E, and

long-range arcs to the repeat F region [43]. Another consideration is whether the different regions of XIST modulate unique aspects of PRC2 recruitment versus activity, potentially as a result of the differing effects of the various core factors (SUZ12) and cofactors (e.g. JARID2) involved. Differences in abundance of factors, such as JARID2 which is relatively low in the HT1080 cells, might shift the balance in recruitment pathways; however, the dramatic reduction of H3K27me3 upon loss of PRC1 in mouse differentiating ESCs [6] is very different from the independent recruitment of PRC1 and PRC2 observed in these human somatic cells.

SMCHD1 recruitment overlapped regions required for PRC1, and was also inhibited by PRT4165, a specific PRC1 inhibitor, which is consistent with mouse studies showing SMCHD1 recruitment requires UbH2A [20]. In our system, SMCHD1 also required the non-repetitive region at the end of exon 1 that has been seen to form numerous duplexes with the BC region of XIST. SMCHD1 did not require the A repeat region, perhaps reflecting that the A-repeat induced silencing or RNA modification that is required to allow detectable spread of UbH2A is not required for the recruitment of SMCHD1 to the XIST-coated chromosome.

There was also overlap between the regions of XIST required for H3K27me3 and MacroH2A recruitment, and inhibition with GSK343 substantially reduced the recruitment of both to the XIST-coated domain, indicating that MacroH2A recruitment was dependent upon the catalytic activity of PRC2. We observed that the complete loss of H3K27me3 either in the deletion constructs of 8p or through chemical inhibition of PRC2 resulted in XIST not being able to recruit MacroH2A. However, slight attenuations in H3K27me3 levels either through lower doses of PRC2 inhibitor or as seen in the ΔBh construct did not noticeably impact MacroH2A levels, suggesting that even moderate enrichment of H3K27me3 may be sufficient for MacroH2A enrichment. The discovery of this novel association has yet to be investigated in mouse models, though both have been observed to be lost when Xist is deleted [32,44]. At time of writing the mechanisms underlying the connection between PRC2 and MacroH2A during XCI are unknown. MacroH2A is enriched at H3K27me3 enriched chromatin throughout the genome, reviewed in [45], but appears to be established independently [46]. It remains to be determined whether MacroH2A directly interacts with XIST; however, the lack of its presence in the interactome screens may implicate intermediate factors, which perhaps bind upstream of the D repeat [26,29,39,47].

Overall, we observe similarities and differences from previous studies of the functionality of Xist/XIST. Many of these have utilized mouse ES cells, and thus our study differs in both developmental timing as well as the species being investigated. The difference between mouse and human sequences are discussed above; however, another important difference between humans and mice is the limited developmental window for Xist function, while we have seen that XIST can induce many features of XCI in human somatic cells, although the cells studied here are cancer-derived and thus may have a less restrictive chromatin state. The differences between the broad contexts of human and mouse XCI are extensive and have recently been well reviewed [48]. A few notable differences include the relatively unique occurrence of imprinted Xist expression, differentiation-dependent protein binding and the more extensive inactivation of X-linked genes in mice relative to humans [15,49].

In our system XIST is induced from a viral-derived inducible promoter, with the promoter inducing expression to approximately the same level as seen for endogenous XIST, an important consideration given the myriad and cooperative binding of proteins to XIST. Previous reports suggested that repeats D, F and A are involved in expression of XIST [50–54], which would not impact our promoter, although a recent report implicates feedback between transcription of mouse Xist and turnover of the RNA [55]. The effects of repeat A deletion seem to be both size and model system dependent in mouse, with variable effects on expression in particular [56–58]. We previously reported a larger deletion of the A region and surrounding

region, Delta XB, in the HT1080 model system. Induction of the Delta XB construct no longer localized into discernable unified RNA clouds and were therefore not usable for IF-FISH analysis so our analysis was limited to the smaller Delta A deletion (sequences present in these constructs are compared in S2 Information). Only the specific removal of the 3' non-repeat region of XIST was found to decrease transcript levels in our study. Work in mouse models identified that truncation of the 3' terminal of mouse Xist led to a reduction in the half-life of transcripts so it may be that some stabilizing elements have been conserved between XIST/Xist transcripts [59]. The unification of the XIST RNA transcripts themselves into a single domain was found to be generally very stable, with only the synergistic deletion of both >3kb regions of the ΔΔ construct or the deletion of the large non-repeat region spanning between repeat D and E noticeably affecting RNA unification. The non-repeat region spanning D to E was also found to be crucial for SMCHD1 enrichment through ubH2A independent mechanisms. SMCHD1 has been associated with the compartmentalization of the Xi, thus it is possible that the two processes of unification and SMCHD1 enrichment may be critically facilitated by this region of XIST as shown in summary Fig 5 [60,61].

We only examined 4 chromatin marks in a somatic cell model, yet we found that every region of XIST we examined was critical for some aspect of XCI. Thus, despite the size of XIST, and the only limited conservation with mouse Xist, including deviation in extent of tandem repeats, the lncRNA seems highly adapted to function in multiple independent pathways. While our results contribute to the concept of modularity of lncRNAs, they also emphasize the need to shift the paradigm of the functional domain of a lncRNAs away from being single linear sequences towards being based on secondary and tertiary arrangements. Future studies and research examining how to regulate these complex and long-range interactions between regions of XIST and other non-coding RNAs seem likely to yield new breakthroughs in the field of RNA biology and potential insights into the utility of XIST as a therapeutic for chromosome abnormality disorders such as trisomy 21.

Materials and methods

Cell culturing protocols for the HT1080 cell lines

Dox-inducible XIST cDNA constructs in autosomes of the male fibrosarcoma cell line, HT1080, were generated as described by Kelsey et al and Chow et al [16,17]. In brief, inducible XIST cDNA constructs (Delta PflMI, Exon 1 or Delta Delta; Fig 1) were integrated into an 8p FRT site in HT1080 cells previously transfected with pcDNA6/TR for Tet-Repressor protein expression. The full XIST cDNA (Full XIST) was previously described, and corresponded to the short-isoform of XIST, the mouse homolog of which has also been reported to be fully functional [62]. The integrated XIST cDNA constructs were controlled by a CMV promoter, that was blocked by TetR and induced in the presence of 1ug/ml doxycycline. The HT1080 cells were grown at 37˚C with 5% CO_2 in DMEM supplemented with 10% Fetal Calf Serum (Sigma-Aldrich) by volume, 100 U/ml Penicillin-Streptomycin, non-essential amino acids and 2mM L-Glutamine. The chemical inhibitors GSK343 (Sigma-Aldrich) and PRT4165 (Sigma-Aldrich) were dissolved in DMSO and added to the DMEM media at the concentrations listed during the chemical inhibition assays. An equal volume of DMSO (without inhibitors) as was used in the inhibition treatments was added to the media of the HT1080 cells used as controls during the inhibition experiments to ensure that any differences between the populations of cells were a result of the inhibitors. Media with chemical inhibitors or doxycycline was replaced daily.

CRISPR modifications of XIST constructs

Excising specific regions of the XIST construct involved transiently introducing two guide RNAs along with a Cas9 gene into HT1080 2-3-0.5a Full XIST 1c.1 cell lines. The gRNA vector

pSPgRNA plasmid (Addgene #47108) gifted by Charles Gersbach contained BbsI digestible sites, that when cleaved (NEB #R0539) allowed a desired target sequence to be inserted as part of the gRNA gene. The gRNA target sequences were chosen with the E-CRISP online tool (http://www.e-crisp.org/E-CRISP/) provided by Deutsches Krebsforschungszentrum. The tools default settings were set to 'strict', and FASTA sequences for each region to be targeted (1kb) by a gRNA were pasted into the program. gRNAs between 22 and 19bp were included and off-target analysis was carried out using Bowtie2 against the Homo sapiens GRCh38 genome, along with an additional test for potential interference with the puromycin resistance gene. Sense and antisense oligos of the target sequence were ordered, annealed and phosphory-lated using T4 PNK (NEB #M0201). The phosphorylated double-stranded target sequences were ligated using T7 ligase (NEB #M0318) into the BbsI digested pSPgRNA plasmid and were transformed into DH5a competent cells (ThermoFisher #18265017). The pSPgRNA and the Cas9 producing plasmid pSpCas9(BB)-2A-Puro (PX459) gifted by the Zhang lab [63] (Addgene #62988) were purified using Qiaquick miniprep kits (Catalog # 27115). The concentration and purity of the purified plasmids was determined using a spectrophotometer, and in cases where the plasmids were too dilute (less than 0.5 ug/ml) they were concentrated using a speed-vac and remeasured.

The inducible XIST constructs within the HT1080 2-3-0.5a cells were modified by transiently transfecting two unique gRNA plasmids (pSPgRNA) and a Cas9 plasmid providing transient puromycin resistance (pSpCas9(BB)-2A-Puro) using Lipofectamine 3000 (ThermoFisher #L3000008). The day before the transfection the HT1080 cells were split into 24 well plates, at around ~30–40% confluency. On the day of transfection 0.5ug of combined plasmid DNA was mixed with 1.5ul of the Lipofectamine 3000 reagent as directed by ThermoFisher's protocol and the whole mixture was pipetted into a well of a 24 well plate. Extensive optimization determined that a molar ratio of three of each gRNA plasmid per Cas9 plasmid in the transfection mixture resulted in the greatest overall efficiency. The cells were left overnight to absorb the plasmids and 24 hours after transfection they were treated with 1ug/ml puromycin in the media. Treatment in puromycin continued for three days at which point only cells that had undergone transfection with the pSpCas9(BB)-2A-Puro plasmid were alive. The remaining cells were then transferred to 100mm plates (roughly 20–30 per plate) and the individual cells were allowed to grow into colonies over the next two weeks. Single cell colonies were picked and transferred to 24 well plates, where they were allowed to grow until nearly confluent. When nearly confluent, the cells were split, with roughly nine tenths of the cells transferred into 1.5ml eppendorf tubes and incubated at 55°C overnight in 100ul of Mouse Homogenization Buffer (50 mM KCl, 10mM Tris-HCl, 2mM MgCl2, 0.1mg/ml gelatin, 0.45% IGEPAL CA-630, 0.45% Tween 20) with 1.2ul of 10mg/ml Proteinase K (Protocol provided by Andrea Korecki of the Simpson Lab, Centre for Molecular Medicine and Therapeutics, UBC). The Proteinase K was inactivated by incubating the samples at 95°C for ten minutes. DNA from the colonies was tested by PCR using primers spanning the regions to be deleted (S2 Table), and running the products on a gel. The colonies that produced bands of the correct size were sent to UBC's Sequencing + Bioinformatics Consortium for Sanger sequencing. To avoid the possibility of clones arising from a single progenitor, clones that were not from different transfections needed to have at least one base pair variable at their deletion sites to be deemed independent.

RNA isolation, reverse transcription and quantitative real-time PCR (RT-qPCR) of cDNA

RNA was obtained from cells either by directly treating cells with TRIZOL (Invitrogen) in t25 flasks or by trypsinizing and collecting adherent cells, pelleting them at 5000rcf for one minute

and removing the supernatant before treating with TRIZOL. Both techniques were effective at obtaining high quality RNA and no observable difference between the results using either technique was observed. Purification of the cellular RNA was carried out according to the Invitrogen protocol in the TRIZOL user guide. The RNA obtained in this manner was measured using spectrometry to determine the concentration and purity of RNA. 5ug of RNA was transferred into 50ul DNAse1 (Roche) reactions consisting of 5ul 10x DNAse1 buffer, 1 U/ul RNAseI, 10 units of DNAse1 and the remaining volume of DEPC ddH$_2$O. The DNAse1 reactions were incubated at 35˚C for 20 minutes and then heat inactivated at 75˚C for 10 minutes. The DNAse-treated RNA was then reverse transcribed using the M-MLV Reverse Transcriptase (Invitrogen #28025013) in a reaction volume consisting of 1ug of RNA, 4ul of 5x first strand buffer, 0.25mM dNTP, 0.01mM DTT, 1ul random hexamers, 1U/ul RNAse Inhibitor and 1ul of M-MLV. The final volume of 20ul was obtained with DEPC treated ddH$_2$O and the reaction volume was gently mixed by inversion before letting sit at room temperature for five minutes. The reaction mixture was then incubated at 42˚C for 2 hours, then inactivated at 95˚C for five minutes. The newly produced cDNA was then immediately stored at -20˚C until it was ready to be used in subsequent tests.

To test the relative expression of XIST across cell lines quantitative real-time PCR (qPCR) was carried out using EVAgreen (Biotium) and Maxima Hot Start Taq (Thermo Scientific) on a StepOnePlus Real-Time PCR System (Applied Biosystems, Darmstadt, Germany). At least three technical replicates were run for each combination of cDNA and primers tested. The cDNA itself for each sample was diluted two-fold with ddH$_2$O, and the cDNA for each qPCR reaction consisted of 1.5ul of that communal dilution. Each qPCR reaction was performed in 20ul reaction volume consisting of 0.2mM dNTP, 5mM MgCl$_2$, 0.25 nmoles of sense and antisense primers, 2ul of 10x Maxima Hot Start Taq Buffer, 0.16ul of Maxima Hot Start Taq and 1ul of EVA green. The remaining 18.5ul of the reaction mixture (not including the remaining 1.5ul of cDNA) consisted of ddH$_2$O. For the sake of improved consistency, master mixes for each combination of primers were made and 18.5ul aliquots of that mastermix were placed in each reaction well prior to the addition of the cDNA. The PCR was performed in MicroAmp Fast Optical 96-Well Reaction Plates or Optical 8-Well Strips (Applied Biosystems) by first heating to 95˚C for 5 minutes before running through 40 cycles of 95˚C for 15 seconds, 60˚C for 30 seconds and finally 72˚C for one minute. The levels of XIST expressed in each cell line were compared to an endogenous control, PGK1, and the same primers for both genes were used across all tests and the XIST primers bound the transcript 5' of the sites deleted. Expression of XIST in each construct was compared to a 5ddox'd Full length XIST construct whose RNA was purified and reverse transcribed simultaneously with the construct to be tested. The relative quantification of XIST in each sample was calculated using the ΔΔCT method where $\Delta\Delta CT = (CT_{XIST\_test} - CT_{PGK1\_test})/(CT_{XIST\_control} - CT_{PGK1\_control})$ and expressed as rq = $2^{-\Delta\Delta Ct}$.

Immunofluorescence and RNA FISH

A detailed protocol listing all the steps necessary to replicate the IF-FISH work done in this study can be found online at protocols.io: dx.doi.org/10.17504/protocols.io.bqtdmwi6.

Adherent cells were transferred onto glass coverslips two days before being permeabilized and fixed, at which point growth media was removed, then the cells were washed in PBS and then in CSK buffer for 3–5 minutes. The cells were then permeabilized in chilled CSK media supplemented with Triton-X (5%v/v) for 8 minutes. The newly permeable cells were then treated in 4% paraformaldehyde in PBS for an additional 8 minutes before being stored in 70% ethanol at 4˚C.

The fluorescent RNA probes for FISH were created from template DNA complementary to the XIST RNA. The probes either targeted the 5' region of XIST including the A and F repeats, or the 3' most 3.6kb region of the short XIST isoform's cDNA. These probes were nick-translated (Abbott Molecular) with either Green 496 dUTP or Red 598 dUTP (Enzo) then precipitated and resuspended in 50ul of DEPC ddH$_2$O and stored at -20°C in the dark.

IF and RNA FISH were performed jointly as most investigations hinged on the relative positions of various factors to the XIST RNA cloud. Permeable and fixed cells on coverslips were incubated face down on 100ul droplets of PBT (1% v/v Bovine Serum Albumin and 0.1% v/v Tween-20 in PBS) supplemented with 0.4U/ul RNAse Inhibitor (Ribolock) for twenty minutes. These cells were then incubated at room temperature for 4–6 hours in an RNAse inhibited 100ul PBT droplet, containing 1ul of the primary antibody of interest (S11 Table). To prevent the coverslip and antibody solution from drying they were sealed between two layers of Parafilm, creating a semi-air tight Parafilm pocket. The coverslips could also be left overnight in this Parafilm pocket without any significant repercussions, so long as they were kept at 4°C rather than room temperature.

Just prior to retrieving the coverslips from their Parafilm pockets, hybridization mixtures were created from 5ul of the nick translated fluorescent RNA probe mixed with 10ul of human Cot-1 (ThermoFisher #15279011) and 2ul of Salmon Sperm DNA (ThermoFisher #15632011), the latter two ingredients being included to prevent the probe from binding non-specifically. The hybridization mixtures were then completely dried in a speedvac while the coverslips were retrieved from their Parafilm pockets with the primary antibody solution and washed three times in PBST (PBS and 0.1% Tween-20) before being incubated for 1 hour at room temperature with the fluorescently-labelled secondary antibody (1:100 PBT supplemented with RNase Inhibitor, 1ul secondary antibody, S12 Table). From this point onwards the coverslips were kept in dark containers to avoid photobleaching. Next, the coverslips were washed for five minutes in PBST three times, to remove the unbound secondary antibodies. They were then fixed in 4% paraformaldehyde for 10 minutes. The coverslips were washed once in PBS to remove most of the paraformaldehyde and then were submerged for two minutes each in 70%, 80% and finally 100% ethanol to dehydrate them, before air drying the coverslips at room temperature for ~10 minutes. The desiccated probes were resuspended in 10ul of deionized formamide (Sigma) and heated to 80°C for 10 minutes. An additional 10ul of hybridization buffer (20% BSA and 20% Dextran Sulfate in 4x SSC) was added to the hybridization mixture, which was gently mixed and pipetted as a droplet on a piece of parafilm onto which a coverslip was placed, cell side down. A second piece of parafilm was placed on top and the edges were sealed, creating a parafilm pocket where the probes could hybridize overnight at 37°C.

The next day the coverslips were retrieved and incubated in an equal measure of deionized formamide (Sigma) and 4x SSC (Invitrogen) at 37°C for 20 minutes. The coverslips were then incubated in 2x SSC at 37°C and 1x SSC at room temperature, for 15 minutes each. DAPI staining was performed by placing the coverslips in a 0.1ug/ml solution of DAPI in pure methanol at 37°C for 15 minutes. The excess DAPI was rinsed off in methanol and the coverslips were mounted on glass slides using a hardset antifade mounting media (Vectashield). After letting the media harden, cells were photographed using a confocal fluorescence microscope (DMI 6000B) and camera (MicroPublisher 5.0 RTV, Qimaging) to capture and compile the red (secondary antibody), green (XIST RNA probe) and blue (DAPI) fluorescent channels using the Openlab program (Perkin Elmer).

The colour channel images obtained for a given cell were converted into composite RGB images using ImageJ (Fiji). The fluorescent intensity of the DAPI (blue), immunofluorescence channel (red), and XIST fluorescence (green), were measured using the BAR plugin and drawing a straight line bisecting the point of greatest XIST fluorescence and the maximal width of

the nucleus possible without intersecting a nucleolar territory. The fluorescent intensity at each position across the line was recorded and the positions were bifurcated into either the XIST +ve category or XIST -ve category. The pixels that had a level of green intensity greater than 50% of the maximum along the range of XIST fluorescence were defined as XIST +ve while those less than 25% were defined as XIST negative. Those very few positions with XIST signal intensity between 25%-50% were not included in the analysis to better delineate the two groups. The mean and standard deviation of Immunofluorescence intensity in the XIST +ve and XIST -ve groups were determined and used to calculate the z-score for each cell. Either 59–61 cells were analyzed in every condition described throughout this work and the population distribution of z-scores were compared using a Mann-Whitney (M.W.) U test. The M.W. test compares the distribution of those enrichment values collected across a population and directly examines whether the distribution of values between populations are similar or different from each other. The threshold for statistical significance was adjusted based on the number of tests being performed (e.g. p-value = 0.05/# of tests). The identities of all the cells tested and the heterochromatin marks being analyzed in each case were blinded throughout this process and the identities only revealed after all testing and analysis had been completed.

Estimations of the relative size of the XIST RNA clouds within the nucleus were performed by dividing the number of XIST +ve pixels by the absolute number of pixels that could be measured within the nucleus of a cell. As all the deletion constructs were integrated into HT1080 cell lines it was anticipated that cell to cell variability in the maximal length of the line bisecting a cell would average out when the analysis was performed using all the cells analyzed in this paper. The relative distribution of XIST RNA in the HT1080 cells was also compared to a female control cell line, IMR90.

Western blotting

A detailed protocol listing all the steps necessary to replicate the western blotting done in this study can be found online at protocols.io: dx.doi.org/10.17504/protocols.io.bqubmwsn.

Protein was extracted from the HT1080 cell lines using RIPA buffer, prepared according to the Cold Spring Harbor protocol, supplemented with 1μl of Protease Inhibitor Cocktail from Roche per 100μl of RIPA buffer. A volume of 500μl of RIPA buffer was used per 2 million cells used. Samples were gently agitated by shaking for 30 minutes at 4°C to allow for the breakdown of cells, then were centrifuged at max speed (>14,000 rcf) at 4°C for 20 minutes.

The acrylamide gel was prepared in Bio-Rad 1.5mm casting plates and gel casting stand. The lower gel consisted of 10ml of 12% acrylamide lower running gel [4.2ml of 29:1 acrylamide/bis-acylamide (Bio Rad), 2.5ml of 4x lower gel buffer (1.5M Tris and 0.4%SDS in distilled water, pH of 8), 3.3ml of distilled water, 10μl of TEMED from Fisher Scientific and 40μl of 10% ammonium persulfate (which was prepared fresh each time) and topped with 100% isopropyl alcohol ~2cm from the top of the plates. Once the lower gel solidified in the plates, the upper gel mixture replaced the isopropyl alcohol [0.75ml 29:1 acrylamide/bis-acylamide, 0.45ml of 4x upper gel buffer (0.5M Tris and 0.4%SDS in distilled water, pH of 6.8), 1.8ml of distilled water, 10μl of of 10% ammonium persulfate and 5μl of TEMED (Fisher Scientific)] and a gel comb was inserted.

The solidified gel was loaded into a Bio-Rad electrode assembly and buffer tank. The 1x running buffer was diluted from a 10x stock (0.25M Tris, 1.92M Glycine and 1% w/v SDS in distilled water) using distilled water. 15μl of each protein extract was mixed with 2X SDS gel-loading buffer (4% SDS, 0.2% bromophenol blue, 20% glycerol and 200mM dithiothreitol) according to the Cold Spring Harbor protocol and the resulting mixture was heated to 95°C for 2–3 minutes. The samples were run with 180 volts until the leading band reached the

bottom of the gel. BenchMark Pre-stained Protein Ladder from Thermo Fisher was loaded into one of the wells abutting the samples to provide a reference for the size of the bands.

To transfer the protein samples to a piece of 0.2μm nitrocellulose paper (ThermoFisher) the gel and nitrocellulose paper were placed together inside a Bio Rad Core Assembly Module according to the manufacturer's instructions. The Core Assembly Module was then inserted as directed into the Bio Rad protein transfer tank and a cold ice pack was placed in the Transfer Tank as well. The tank was then filled with Transfer Buffer diluted from a cold 10x stock [0.25M Tris and 1.9M glycine] with 20% methanol and 70% distilled water per final volume. and was hooked up to a Bio Rad PowerPac HC power supply. Protein transfer was carried out at 90 volts for one hour at 4˚C with constant gentle agitation of the Transfer buffer using a magnetic stir bar to ensure dispersion of heat.

After the proteins were transferred to the nitrocellulose membrane the membrane was incubated with a 40ml blocking buffer (0.1% v/v Tween-20 and 3% Bovine Serum Albumin) for 1 hour while gently being agitated to ensure complete coverage of the buffer. After blocking, the nitrocellulose was placed inside a 50ml falcon tube containing 20ml of primary antibody solution (3% w/v BSA, 0.1% v/v Tween-20 in TBS plus the relevant antibodies listed in S12 Table). The membrane was incubated in this solution overnight at 4˚C on a tube rotator to allow even coating of the entire surface of the membrane with antibodies. The next day the nitrocellulose membrane was washed four times for 5 minutes each with TBST (0.1% v/v Tween-20 in TBS). The nitrocellulose membrane was then in 20ml of secondary antibody solution (3% w/v BSA, 0.1% v/v Tween-20 in TBS plus either goat anti-mouse or goat anti-rabbit fluorescently labelled secondary antibodies, S12 Table) for 1 hour while being gently agitated. The nitrocellulose was washed twice with TBST then once with TBS for five minutes each. The protein and antibody bound membrane was then imaged at the relevant wavelengths using an LI-COR Odyssey machine from BioAgilytix and the software package, Image Studio.

Supporting information

S1 Fig. Relative XIST transcript levels expressed from chromosome 8p in HT1080 cell line as well as relative size of the XIST RNA cloud in the deletion constructs. A) XIST transcript levels were measured using RT-qPCR and determining the relative transcript levels compared to the endogenous control gene PGK1. Four biological replicates for each time point of XIST induction in the 8p HT1080 cell line were tested and the relative levels of XIST were normalized to the average level of the 5 day time point, as it was the time point which had become standard in previous examinations of the model system. The mean and standard deviation for each condition are indicated by lines, and individual dots representing the relative expression of each replicate. Increasing darkness of shading was used to indicate increasing length of *XIST* induction. The statistical significance of a difference between the 5 day treatment and other time points was calculated by two-tailed unpaired t-test (* $p < 0.05$, ** $p < 0.01$). B) Approximation of the relative size of XIST RNA clouds across all of the IF-FISH labelled cells analyzed in this paper. For each cell the number of XIST +ve pixels were divided by the absolute number of pixels that could be measured across the length of the cell without intersecting the nucleolar compartments. This information was used to check for any signs of variability in the proportion of the nucleus taken up by XIST RNA. A female control cell line (IMR90) was included in this analysis as a reference however its nuclei was typically much larger than that of the HT1080s, resulting in it occupying a proportionally smaller region of the nuclei. Significance was calculated using M.W. test. (*** $p < 7.14 \times 10^{-5}$).
(TIF)

S2 Fig. XIST RNA clouds of the different HT1080 deletion cell lines following five days of induction. XIST RNA was labelled by FISH in green and the DNA of the nuclei was labelled with DAPI. The white box denotes the area shown in the grayscale image of XIST RNA accompanying each image of a cell nuclei. The greyscale images for each cell show only the fluorescent channel of XIST RNA.
(TIF)

S3 Fig. Enrichment of H3K27me3 by the Full XIST and Delta XIST constructs. Four example IF-FISH images for each type of XIST inducible construct are in rows to the left of the appropriate label. XIST RNA is labelled green, H3K27me3 labelled red and the DNA labelled blue with DAPI. The z scores calculated for all the analyzed cells for each type of construct are shown in the dot plot, with each cell represented by a single dot. The red dashed line indicates the median z-score (2.59) of the Full XIST construct to provide a point of comparison.
(TIF)

S4 Fig. Enrichment of ubH2A by the Full XIST and Delta XIST constructs. Four example IF-FISH images for each type of XIST inducible construct are in rows to the left of the appropriate label. XIST RNA is labelled green, ubH2A labelled red and the DNA labelled blue with DAPI. The z scores calculated for all the analyzed cells for each type of construct are shown in the dot plot, with each cell represented by a single dot. The red dashed line indicates the median z-score (4.2) of the Full XIST construct to provide a point of comparison.
(TIF)

S5 Fig. Enrichment of MacroH2A by the Full XIST and Delta XIST constructs. Four example IF-FISH images for each type of XIST inducible construct are in rows to the left of the appropriate label. XIST RNA is labelled green, MacroH2A labelled red and the DNA labelled blue with DAPI. The z scores calculated for all the analyzed cells for each type of construct are shown in the dot plot, with each cell represented by a single dot. The red dashed line indicates the median z-score (2.0) of the Full XIST construct to provide a point of comparison.
(TIF)

S6 Fig. Enrichment of SMCHD1 by the Full XIST and Delta XIST constructs. Four example IF-FISH images for each type of XIST inducible construct are in rows to the left of the appropriate label. XIST RNA is labelled green, SMCHD1 labelled red and the DNA labelled blue with DAPI. The z scores calculated for all the analyzed cells for each type of construct are shown in the dot plot, with each cell represented by a single dot. The red dashed line indicates the median z-score (2.7) of the Full XIST construct to provide a point of comparison.
(TIF)

S7 Fig. Enrichment of heterochromatin marks following chemical inhibition. Four example IF-FISH images for each type of XIST inducible construct are in rows to the left of the appropriate label. XIST RNA is labelled green, the heterochromatin features being examined are labelled red and the DNA labelled blue with DAPI. The identity of the heterochromatin marks is adjacent to the photos. The z scores calculated for all the analyzed cells for each type of construct are shown in the dot plot, with each cell represented by a single dot.
(TIF)

S8 Fig. Effect of additional GSK343 doses on XIST activity in 8p. A) The effect of additional PRC2 inhibition with GSK343 oH3K27me3 (pink) and MacroH2A (yellow) at the XIST RNA cloud relative to the average level in the nucleus across a population of 59–61 cells. We were blinded to all cells and chromatin marks identity until after all data and calculations were completed. Statistical significance of the effect of each inhibitor on a chromatin mark was

calculated using the Mann Whitney test with adjusted p value (* $p < 8.3 \times 10^{-3}$, ** $p < 1.6 \times 10^{-3}$, *** $p < 1.6 \times 10^{-4}$). B) Relative XIST RNA levels for each inhibitor concentration relative to control 5ddox. The dots denote independent biological replicates and statistical significance was calculated using a t-test (* $p < 0.05$). C) Western blotting images demonstrating the levels of H3K27me3 after cells had undergone chemical inhibition with GSK343, relative levels of H3K27me3 were determined through comparison with H4 and the normalized values are shown beneath each lane.
(TIF)

S1 Table. gRNAs used to generate XIST deletions.
(DOCX)

S2 Table. XIST deletion sizes confirmed by sequencing across deletion. Each of the cell lines successfully generated for each type of deletion construct are listed along with the gRNAs used to create the deletion and the total number of nucleotides lost from the XIST cDNA sequence.
(DOCX)

S3 Table. Relative size of XIST RNA cloud in deletion constructs. The relative size of the XIST +ve region of the nucleus of each cell was measured relative to the maximal length of measurement taken within each cell. The number of cells for each construct along with the mean size of the XIST region and total length measured. The mean relative size of each type of XIST construct within the nucleus and standard deviation (SD) were calculated. Cells from an IMR90 female control cell line were also included.
(DOCX)

S4 Table. XIST expression levels in deletion clones relative to Full XIST. The relative expression (RQ) of each deletion cell line as well as construct is listed along with the standard deviation (SD) across biological replicates (≥ 3 per cell line) as well as the resulting p value of any construct or cell lines difference from Full XIST.
(DOCX)

S5 Table. Summary of H3K27me3 enrichment in deletion constructs. List of the number of cells analyzed, the median z-score calculated as well as the standard deviation (SD) for each construct. The median z-score of a female control cell line (IMR90) was included to exemplify established XCI. The statistical significance of each population of deletion constructs' difference from Full XIST in their enrichment was calculated using the Mann-Whitney U test and the p values are listed.
(DOCX)

S6 Table. Summary of ubH2A enrichment in deletion constructs. List of the number of cells analyzed, the median z-score calculated as well as the standard deviation (SD) for each construct. The statistical significance of each population of deletion constructs' difference from Full XIST in their enrichment was calculated using the Mann-Whitney U test and the p values are listed.
(DOCX)

S7 Table. Summary of MacroH2A enrichment in deletion constructs. List of the number of cells analyzed, the median z-score calculated as well as the standard deviation (SD) for each construct. The statistical significance of each population of deletion constructs' difference from Full XIST in their enrichment was calculated using the Mann-Whitney U test and the p values are listed.
(DOCX)

S8 Table. Summary of SMCHD1 enrichment in deletion constructs. List of the number of cells analyzed, the median z-score calculated as well as the standard deviation (SD) for each construct. The statistical significance of each population of deletion constructs' difference from Full XIST in their enrichment was calculated using the Mann-Whitney U test and the p values are listed.
(DOCX)

S9 Table. Summary of effect of chemical inhibitors on XIST mediated chromatin remodelling when expressed from 8p. List of the number of cells analyzed, the median z-score calculated as well as the standard deviation (SD) for each treatment condition and heterochromatin feature. The statistical significance of each population of inhibitor treated cells' difference from the uninhibited control population was calculated using the Mann-Whitney U test and the p values are listed.
(DOCX)

S10 Table. Summary of effect of additional doses of GSK343 treatment on XIST mediated chromatin remodelling. List of the number of cells analyzed, the median z-score calculated as well as the standard deviation (SD) for each treatment condition and heterochromatin feature. The statistical significance of each population of inhibitor treated cells' difference from the uninhibited control population was calculated using the Mann-Whitney U test and the p values are listed.
(DOCX)

S11 Table. Summary of effect of chemical inhibitors on XIST mediated chromatin remodelling when expressed from the X chromosome of the HT1080 cell line. List of the number of cells analyzed, the median z-score calculated as well as the standard deviation (SD) for each treatment condition and heterochromatin feature. The statistical significance of each population of inhibitor treated cells' difference from the uninhibited control population was calculated using the Mann-Whitney U test and the p values are listed.
(DOCX)

S12 Table. List of Antibodies used for Immunofluorescence and western blotting.
(DOCX)

S1 Information. List of sequences across the cut sites of CRISPR deletion constructs. The sequencing results across the cut size of each of the deletion constructs created through CRISPR were listed below arranged from 5' to 3' along the sense sequence of the *XIST* construct. Bolded letters indicated the gRNA target sequences remaining within the final transcript. The cut site was denoted by vertical bars (|) which encapsulated the summary of the deletion size and any other characteristics of note. The sequences of the repeats were underlined to provide landmarks when referring to the location of the deletions in each instance.
(DOCX)

S2 Information. Sequence removed from CRISPR generated Delta A#12 construct and previously generated Delta XB. The sequences of XIST absent from both constructs aiming to examine the effect of loss of repeat A on the activity of XIST are listed. The names of the construct are shown in bold along with the nucleotide positions within the XIST transcript that have been removed. The sequence corresponding to Repeat A has been underlined.
(DOCX)

Acknowledgments

Author Contributions

References

1. Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, et al. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature. 1991 Jan 3; 349(6304):38–44. https://doi.org/10.1038/349038a0 PMID: 1985261

2. Loda A, Heard E. Xist RNA in action: Past, present, and future. PLoS Genet. 2019 Sep; 15(9): e1008333. https://doi.org/10.1371/journal.pgen.1008333 PMID: 31537017

3. Yen ZC, Meyer IM, Karalic S, Brown CJ. A cross-species comparison of X-chromosome inactivation in Eutheria. Genomics. 2007 Oct; 90(4):453–63. https://doi.org/10.1016/j.ygeno.2007.07.002 PMID: 17728098

4. Minks J, Baldry SE, Yang C, Cotton AM, Brown CJ. XIST-induced silencing of flanking genes is achieved by additive action of repeat a monomers in human somatic cells. Epigenetics Chromatin. 2013 Aug 1; 6(1):23. https://doi.org/10.1186/1756-8935-6-23 PMID: 23915978

5. Nesterova TB, Slobodyanyuk SY, Elisaphenko EA, Shevchenko AI, Johnston C, Pavlova ME, et al. Characterization of the genomic Xist locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. Genome Res. 2001 May; 11(5):833–49. https://doi.org/10.1101/gr.174901 PMID: 11337478

6. Almeida M, Pintacuda G, Masui O, Koseki Y, Gdula M, Cerase A, et al. PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. Science. 2017 Jun 9; 356(6342):1081–4. https://doi.org/10.1126/science.aal2512 PMID: 28596365

7. Brown CJ, Hendrich BD, Rupert JL, Lafrenière RG, Xing Y, Lawrence J, et al. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell. 1992 Oct 30; 71(3):527–42. https://doi.org/10.1016/0092-8674(92)90520-m PMID: 1423611

8. Chiang J-C, Jiang J, Newburger PE, Lawrence JB. Trisomy silencing by XIST normalizes Down syndrome cell pathogenesis demonstrated for hematopoietic defects in vitro. Nat Commun. 2018 Dec 5; 9 (1):5180. https://doi.org/10.1038/s41467-018-07630-y PMID: 30518921

9. Jiang J, Jing Y, Cost GJ, Chiang J-C, Kolpa HJ, Cotton AM, et al. Translating dosage compensation to trisomy 21. Nature. 2013 Aug 15; 500(7462):296–300. https://doi.org/10.1038/nature12394 PMID: 23863942

10. Brockdorff N, Bowness JS, Wei G. Progress toward understanding chromosome silencing by Xist RNA. Genes Dev. 2020 Jun 1; 34(11–12):733–44. https://doi.org/10.1101/gad.337196.120 PMID: 32482714

11. Patel S, Bonora G, Sahakyan A, Kim R, Chronis C, Langerman J, et al. Human Embryonic Stem Cells Do Not Change Their X Inactivation Status during Differentiation. Cell Rep. 2017 Jan 3; 18(1):54–67. https://doi.org/10.1016/j.celrep.2016.11.054 PMID: 27989715

12. Petropoulos S, Edsgärd D, Reinius B, Deng Q, Panula SP, Codeluppi S, et al. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. Vol. 165, Cell. 2016. p. 1012–26. https://doi.org/10.1016/j.cell.2016.03.023 PMID: 27062923

13. Migeon BR, Chowdhury AK, Dunston JA, McIntosh I. Identification of TSIX, encoding an RNA antisense to human XIST, reveals differences from its murine counterpart: implications for X inactivation. Am J Hum Genet. 2001 Nov; 69(5):951–60. https://doi.org/10.1086/324022 PMID: 11555794

14. Okamoto I, Otte AP, Allis CD, Reinberg D, Heard E. Epigenetic dynamics of imprinted X inactivation during early mouse development. Science. 2004 Jan 30; 303(5658):644–9. https://doi.org/10.1126/science.1092727 PMID: 14671313

15. Mak W, Nesterova TB, de Napoles M, Appanah R, Yamanaka S, Otte AP, et al. Reactivation of the paternal X chromosome in early mouse embryos. Science. 2004 Jan 30; 303(5658):666–9. https://doi.org/10.1126/science.1092674 PMID: 14752160

16. Chow JC, Hall LL, Baldry SEL, Thorogood NP, Lawrence JB, Brown CJ. Inducible XIST-dependent X-chromosome inactivation in human somatic cells is reversible. Proc Natl Acad Sci U S A. 2007 Jun 12; 104(24):10104–9. https://doi.org/10.1073/pnas.0610946104 PMID: 17537922

17. Kelsey AD, Yang C, Leung D, Minks J, Dixon-McDougall T, Baldry SEL, et al. Impact of flanking chromosomal sequences on localization and silencing by the human non-coding RNA XIST. Genome Biol. 2015 Oct 2; 16:208. https://doi.org/10.1186/s13059-015-0774-2 PMID: 26429547

18. Lee HJ, Gopalappa R, Sunwoo H, Choi S-W, Ramakrishna S, Lee JT, et al. En bloc and segmental deletions of humanXISTreveal X chromosome inactivation-involving RNA elements. Nucl Acids Research. 2019. 47(8) 3875–3887. https://doi.org/10.1093/nar/gkz109 PMID: 30783652

19. Monfort A, Wutz A. The B-side of Xist [Internet]. Vol. 9, F1000Research. 2020. p. 55. https://doi.org/10.12688/f1000research.21362.1 PMID: 32047616

20. Jansz N, Nesterova T, Keniry A, Iminitoff M, Hickey PF, Pintacuda G, et al. Smchd1 Targeting to the Inactive X Is Dependent on the Xist-HnrnpK-PRC1 Pathway. Cell Rep. 2018 Nov 13; 25(7):1912–23. e9. https://doi.org/10.1016/j.celrep.2018.10.044 PMID: 30428357

21. Wang C-Y, Colognori D, Sunwoo H, Wang D, Lee JT. PRC1 collaborates with SMCHD1 to fold the X-chromosome and spread Xist RNA between chromosome compartments. Nat Commun. 2019 Jul 3; 10 (1):2950. https://doi.org/10.1038/s41467-019-10755-3 PMID: 31270318

22. Costanzi C, Pehrson JR. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature. 1998 Jun 11; 393(6685):599–601. https://doi.org/10.1038/31275 PMID: 9634239

23. Brockdorff N. Polycomb complexes in X chromosome inactivation. Philos Trans R Soc Lond B Biol Sci. 2017 Nov 5; 372(1733).

24. Colognori D, Sunwoo H, Wang D, Wang CY, Lee JT. Xist Repeats A and B Account for Two Distinct Phases of X Inactivation Establishment. Dev Cell 2020 Jul 6; 54(1) 21–32.e5. https://doi.org/10.1016/j.devcel.2020.05.021 PMID: 32531209

25. Almeida M, Bowness JS, Brockdorff N. The many faces of Polycomb regulation by RNA. Curr Opin Genet Dev. 2020 May 10; 61:53–61. https://doi.org/10.1016/j.gde.2020.02.023 PMID: 32403014

26. Chu C, Zhang QC, da Rocha ST, Flynn RA, Bharadwaj M, Calabrese JM, et al. Systematic discovery of Xist RNA binding proteins. Cell. 2015 Apr 9; 161(2):404–16. https://doi.org/10.1016/j.cell.2015.03.025 PMID: 25843628

27. Pintacuda G, Wei G, Roustan C, Kirmizitas BA, Solcan N, Cerase A, et al. hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing. Mol Cell. 2017 Dec 7; 68(5):955–69.e10. https://doi.org/10.1016/j.molcel.2017.11.013 PMID: 29220657

28. Cooper S, Grijzenhout A, Underwood E, Ancelin K, Zhang T, Nesterova TB, et al. Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. Nat Commun. 2016 Nov 28; 7:13661. https://doi.org/10.1038/ncomms13661 PMID: 27892467

29. McHugh CA, Chen C-K, Chow A, Surka CF, Tran C, McDonel P, et al. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature. 2015 May 14; 521(7551):232–6. https://doi.org/10.1038/nature14443 PMID: 25915022

30. Monfort A, Di Minin G, Postlmayr A, Freimann R, Arieti F, Thore S, et al. Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells. Cell Rep. 2015 Jul 28; 12(4):554–61. https://doi.org/10.1016/j.celrep.2015.06.067 PMID: 26190100

31. Żylicz JJ, Bousard A, Žumer K, Dossin F, Mohammad E, da Rocha ST, et al. The Implication of Early Chromatin Changes in X Chromosome Inactivation. Cell. 2019 Jan 10; 176(1–2):182–97.e23. https://doi.org/10.1016/j.cell.2018.11.041 PMID: 30595450

32. Pinter SF. A Tale of Two Cities: How Xist and its partners localize to and silence the bicompartmental X. Seminars in Cell & Developmental Biology. 2016 Aug; 56:19–34.

33. Alchanati I, Teicher C, Cohen G, Shemesh V, Barr HM, Nakache P, et al. The E3 Ubiquitin-Ligase Bmi1/Ring1A Controls the Proteasomal Degradation of Top2α Cleavage Complex–A Potentially New Drug Target. Vol. 4, PLoS ONE. 2009. p. e8104. https://doi.org/10.1371/journal.pone.0008104 PMID: 19956605

34. Verma SK, Tian X, LaFrance LV, Duquenne C, Suarez DP, Newlander KA, et al. Identification of Potent, Selective, Cell-Active Inhibitors of the Histone Lysine Methyltransferase EZH2. ACS Med Chem Lett. 2012 Dec 13; 3(12):1091–6. https://doi.org/10.1021/ml3003346 PMID: 24900432

35. Lu Z, Guo JK, Wei Y, Dou DR, Zarnegar B, Ma Q, et al. Structural modularity of the XIST ribonucleoprotein complex. Nature Commun. 2020. 11(1), 1–14 https://doi.org/10.1038/s41467-020-20040-3 PMID: 33268787

36. Tattermusch A, Brockdorff N. A scaffold for X chromosome inactivation. Hum Genet. 2011 Aug; 130(2). https://doi.org/10.1007/s00439-011-1027-4 PMID: 21660507

37. Weidmann CA, Mustoe AM, Jariwala PB, Calabrese JM, Weeks KM. Analysis of RNA-protein networks with RNP-MaP defines functional hubs on RNA. Nat Biotechnol. 2020. In Press. https://doi.org/10.1038/s41587-020-0709-7 PMID: 33077962

38. Pandya-Jones A, Markaki Y, Serizay J, Chitiashvili T, Mancia W, Damianov A, et al. A protein assembly mediates Xist localization and gene silencing. Nature, 132, jcs235093. http://doi.org/10.1038/s41586-020-2703-0

39. Strehle M, Guttman M. Xist drives spatial compartmentalization of DNA and protein to orchestrate initiation and maintenance of X inactivation. Curr Opin Cell Biol. 2020 Jun; 64:139–47. https://doi.org/10.1016/j.ceb.2020.04.009 PMID: 32535328

40. Minajigi A, Froberg JE, Wei C, Sunwoo H, Kesner B, Colognori D, et al. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation [Internet]. Vol. 349, Science. 2015. p. aab2276–aab2276. https://doi.org/10.1126/science.aab2276 PMID: 26089354

41. Lu Z, Carter AC, Chang HY. Mechanistic insights in X-chromosome inactivation. Philos Trans R Soc Lond B Biol Sci. 2017 Nov 5; 372(1733). https://doi.org/10.1098/rstb.2016.0356 PMID: 28947655

42. da Rocha ST, Boeva V, Escamilla-Del-Arenal M, Ancelin K, Granier C, Matias NR, et al. Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. Mol Cell. 2014 Jan 23; 53(2):301–16. https://doi.org/10.1016/j.molcel.2014.01.002 PMID: 24462204

43. G Hendrickson D, Kelley DR, Tenen D, Bernstein B, Rinn JL. Widespread RNA binding by chromatin-associated proteins. Genome Biol. 2016 Feb 16; 17:28. https://doi.org/10.1186/s13059-016-0878-3 PMID: 26883116

44. Csankovszki G, Panning B, Bates B, Pehrson JR, Jaenisch R. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. Nat Genet. 1999 Aug; 22(4):323–4. https://doi.org/10.1038/11887 PMID: 10431231

45. Sun Z, Bernstein E. Histone variant macroH2A: from chromatin deposition to molecular function. Essays Biochem. 2019 Apr 23; 63(1):59–74. https://doi.org/10.1042/EBC20180062 PMID: 31015383

46. Sun Z, Filipescu D, Andrade J, Gaspar-Maia A, Ueberheide B, Bernstein E. Transcription-associated histone pruning demarcates macroH2A chromatin domains. Nat Struct Mol Biol. 2018 Oct; 25(10):958–70. https://doi.org/10.1038/s41594-018-0134-5 PMID: 30291361

47. Gilbert SL, Pehrson JR, Sharp PA. XIST RNA associates with specific regions of the inactive X chromatin. J Biol Chem. 2000 Nov 24; 275(47):36491–4. https://doi.org/10.1074/jbc.C000409200 PMID: 11006266

48. Patrat C, Ouimette J-F, Rougeulle C. X chromosome inactivation in human development. Development. 2020 Jan 3; 147(1). https://doi.org/10.1242/dev.183095 PMID: 31900287

49. Agrelo R, Souabni A, Novatchkova M, Haslinger C, Leeb M, Komnenovic V, et al. SATB1 defines the developmental context for gene silencing by Xist in lymphoma and embryonic cells. Dev Cell. 2009 Apr; 16(4):507–16. https://doi.org/10.1016/j.devcel.2009.03.006 PMID: 19386260

50. Lv Q, Yuan L, Song Y, Sui T, Li Z, Lai L. D-repeat in the XIST gene is required for X chromosome inactivation. RNA Biol. 2016; 13(2):172–6. https://doi.org/10.1080/15476286.2015.1137420 PMID: 26786668

51. Chapman AG, Cotton AM, Kelsey AD, Brown CJ. Differentially methylated CpG island within human XIST mediates alternative P2 transcription and YY1 binding. BMC Genet. 2014 Sep 9; 15:89. https://doi.org/10.1186/s12863-014-0089-4 PMID: 25200388

52. Royce-Tolland ME, Andersen AA, Koyfman HR, Talbot DJ, Wutz A, Tonks ID, et al. The A-repeat links ASF/SF2-dependent Xist RNA processing with random choice during X inactivation. Nat Struct Mol Biol. 2010 Aug; 17(8):948–54. https://doi.org/10.1038/nsmb.1877 PMID: 20657585

53. Hoki Y, Kimura N, Kanbayashi M, Amakawa Y, Ohhata T, Sasaki H, et al. A proximal conserved repeat in the Xist gene is essential as a genomic element for X-inactivation in mouse. Development. 2009 Jan 1; 136(1):139–46. https://doi.org/10.1242/dev.026427 PMID: 19036803

54. Wang Y, Zhong Y, Zhou Y, Tanaseichuk O, Li Z, Zhao JC. Identification of a Xist silencing domain by Tiling CRISPR. Sci Rep. 2019; 20; 9(1):article number 2408. https://doi.org/10.1038/s41598-018-36750-0 PMID: 30787302

55. Rodermund L, Coker H, Oldenkamp R, Wei G, Bowness J, Rajkumar B, et al. Time-resolved structured illumination microscopy reveals key principles of Xist RNA spreading. bioRxiv. 2020;

56. Nesterova TB, Wei G, Coker H, Pintacuda G, Bowness JS, Zhang T, et al. Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. Nat Commun [Internet]. 2019 Dec 1 [cited 2019 Sep 7]; 10(2019):1–15. https://doi.org/10.1038/s41467-019-11171-3 PMID: 31311937

57. Colognori D, Sunwoo H, Kriz AJ, Wang CY, Lee JT. Xist Deletional Analysis Reveals an Interdependency between Xist RNA and Polycomb Complexes for Spreading along the Inactive X. Mol Cell. 2019 Apr 4; 74(1):101–117.e10 https://doi.org/10.1016/j.molcel.2019.01.015 PMID: 30827740

58. Trotman JB, Lee DM, Cherney RE, Kim SO, Inoue K, Schertzer MD, et al. Elements at the 5′ end of Xist harbor SPEN-independent transcriptional antiterminator activity. Nucleic Acids Res [Internet]. 2020 Oct 9 [cited 2021 Feb 20]; 48(18):10500–17. https://doi.org/10.1093/nar/gkaa789 PMID: 32986830

59. Smola MJ, Christy TW, Inoue K, Nicholson CO, Friedersdorf M, Keene JD, et al. SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. Proc Natl Acad Sci U S A. 2016 Sep 13; 113(37):10322–7. https://doi.org/10.1073/pnas.1600008113 PMID: 27578869

60. Gdula MR, Nesterova TB, Pintacuda G, Godwin J, Zhan Y, Ozadam H, et al. The non-canonical SMC protein SmcHD1 antagonises TAD formation and compartmentalisation on the inactive X chromosome. Nat Commun. 2019 Jan 3; 10(1):30. https://doi.org/10.1038/s41467-018-07907-2 PMID: 30604745

61. Wang C-Y, Jégu T, Chu H-P, Oh HJ, Lee JT. SMCHD1 Merges Chromosome Compartments and Assists Formation of Super-Structures on the Inactive X. Cell. 2018 Jul 12; 174(2):406–21.e25. https://doi.org/10.1016/j.cell.2018.05.007 PMID: 29887375

62. Yue M, Ogawa Y. CRISPR/Cas9-mediated modulation of splicing efficiency reveals short splicing isoform of Xist RNA is sufficient to induce X-chromosome inactivation. Nucleic Acids Res. 2018 Mar 16; 46 (5):e26. https://doi.org/10.1093/nar/gkx1227 PMID: 29237010

63. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013 Nov; 8(11):2281–308. https://doi.org/10.1038/nprot.2013.143 PMID: 24157548