

Supplemental Figure Legends

Figure S1: Additional information on an alternative allele-specific 4C strategy and 4C analysis by next generation sequencing. **A:** To distinguish between the conformation of the X_a and X_i in female cells we designed an allele-specific 4C approach as outlined in Figure 1A, B. An alternative strategy, which allows the usage of a ‘RFLP recognizing’ restriction enzyme that is sensitive to CpG methylation is depicted here. In this strategy 4C product is first amplified using 20 cycles of PCR, after which it is treated with the RFLP enzyme to create the allele specificity. An example of the allele-specific formation of 4C PCR product is shown in **B**, where *Pcdh11x* primers only amplify in an allele-specific (here: CAST) manner upon *SnaBI* digestion. As a control the *Jarid1C* primers, which were not designed around a *SnaBI* restriction site, were able to amplify similar amounts of PCR product from both digested and undigested template. **C:** 4C sample preparation is similar between array and NGS analysis. The difference between these two options of analyzing 4C data mainly concerns primer design in two ways. One: using the Illumina platform the sample requires adaptor sequences that allow the DNA to bind to the flowcell. By using primers containing these adaptor sequences we limit sequencing to products formed in the PCR reaction, thus improving the signal to noise ratio. The second restriction in primer design concerns the limited read length of NGS on the Illumina platform. Although improving over time, experimental design is such that 36bp read length results in the generation of informative data. This is done by designing the *HindIII* –or read primer as short as possible (typically 18-20bp) and placing it on top of the *HindIII* restriction site. The first 18bp analyzed in the sequencing reaction are from the primer and are used as a barcode to identify captures from each experiment pooled in one sequencing lane. Typically we mix 10–15 experiments per lane, mostly yielding more than 1 million mappable interactions per viewpoint. Because the *HindIII* sequence in the primer is redundant for the capture it adds 6bp to last 18bp of the sequencing read resulting in a total of 24bp of capture sequence that is used for mapping purposes. A

comparison between two experiments, PCR amplified from the same template but analyzed differently, resulted in the identification of highly similar regions, but with higher resolution (**D** and **E**) due to an increase in signal to noise ratio.

Figure S2: The X chromosome inactivation process is well represented in NPCs. Dependent on which X chromosome carries the RFLP, the 4C analysis is directed to either the 129SvJ or the CAST chromosome. To study both the X_a and X_i two cell lines were generated, one where 129SvJ was active (X_a^{129}) and CAST is silenced (X_i^{CAST}) and one with $X_a^{CAST} X_i^{129}$. **A:** NPC identity was confirmed by IF using antibodies against various proteins representing different stages of differentiation. Oct4, loss of embryonic stem cell (ESC) features; Nestin, neural lineage marker (typically >98% positive n=100); Gfap, astrocyte and Tubulin, neuronal marker. Scale bar represents 10 μ m. **B-E:** Verification of different marks of XCI in NPC. The formation of a Xist RNA cloud, an early mark of XCI, was detected by RNA FISH (**B**). PRC2 recruitment to the X_i was confirmed by an enrichment of Ezh2 (data not shown and Fig. 6D) and its mark, H3K27me3 by IF (**C**). Bisulphate sequencing confirmed the presence of DNA methylation as roughly 50% methylated alleles in NPC while in ES cells almost no DNA methylation could be detected (**D**). Rhox6 is XCI independently methylated and served as a control. Gene silencing was confirmed by allele-specific cDNA analysis (details in Methods) (**E**). Interrogated genes and cell types are indicated. 129 and CAST templates (first two lanes of each panel) and ESC served as references to determine the origin of the transcripts found in NPC X_i^{129} and NPC X_i^{CAST} .

Figure S3: Characterization of X_a interacting regions. Active genes located in gene-dense regions spatially separate from inactive genes located in gene-poor areas on the active X chromosome. Preferred interacting regions of the different viewpoints on X_a are interrogated for gene activity (**A**).

Violin plots depict the distribution of normalized absolute expression of all genes on the X chromosome (grey) and regions that were not contacted (-) and contacted (+) respectively by the different viewpoints indicated. P-values were calculated using a Wilkinson test. Average gene density present in the *cis* contacted regions is plotted in **B**. *Cis* interacting regions of the different viewpoints were also queried for LINE and SINE content (**C**). The dashed line represents the average content of the interrogated repeat elements on the X chromosome. On the X_i little interacting regions were identified due to the lack of clustered 4C-captured-fragment-ends.

Figure S4: Distribution of captured sequences. **A:** Distribution of the relative read count of captured sequences in the different 4C experiments. **B:** Distribution of unique coverage in the different 4C experiments. Note the change in distribution of reads and unique coverage comparing the X_a and X_i shifting from local to more distal (>1MB) contacts.

Figure S5: 4C interactions are confirmed by 3D-DNA FISH and comparing 3D-RNA FISH with 3D-DNA FISH suggest co-localization of genes is not mediated by active transcription. Co-localizing frequencies were determined by performing 3D distance measurements on 100 nuclei. Loci were scored co-localizing when the center of both signals was detected within 1 μ m apart. Only cells containing four signals with clear separation between spots belonging to the X_a or X_i (indicated by Xist RNA staining; BAC probe: CT7-399K20) were analyzed. Co-localizing frequencies were calculated using the total number of measurements taken within each experiment. The BACs, representing the different loci used in the FISH experiments including their position on the X chromosome, are shown below panel E. **A:** Representative examples of DNA and RNA FISH experiments showing co-localization of loci on both X chromosomes, on the X_a , on the X_i or non co-localizing loci respectively. For discrimination between X_a and X_i a BAC probe

spanning *Xist* was added to the experiment. Although in the DNA FISH the detection of the X_i resulted in a dotted appearance of the *Xist* RNA cloud, this did not affect its detection. **B:** DNA FISH co-localizing percentages are plotted for the different combinations of loci indicated. Below the graph the green and red bars indicate if corresponding loci were scored interacting or non-interacting by 4C. Co-localizing percentages for non interacting 4C sequences were found to be lower than 4C interacting sequences, although separated further apart on the linear chromosome (compare *Pcdh11x-Jarid1C* and *Taf1/Ogt-Jarid1C*; 4Cneg-*MeCP2* and *Zfx-MeCP2*). **C:** Depicting the cumulative FISH signal distribution of the experiments shown in (B) further supports this observation. **D:** If co-localization would only occur when two genes are actively transcribed, co-localizing percentages are expected to increase focusing only on cells actively transcribing those two genes. This is tested by comparing co-localizing frequencies found in DNA FISH (interrogation of loci, irrespective of gene activity) to frequencies found in RNA FISH (only actively transcribed loci will show a primary transcript signal). Importantly, interrogated genes were not active in all cells (**E**). Comparing RNA and DNA FISH co-localizing frequencies no difference could be observed, even not for genes expressed in as few as 60% of the cells.

Figure S6: Spearman rank correlation analysis, comparing the X_i and X_a for the same locus (as indicated) revealed that the *cis*-interaction profiles of escapees were as dissimilar as those of the non-escaping genes.

Figure S7: Spearman's correlation analysis applied on 4C data shows a partial refolding of X_i upon *Xist* RNA depletion while gene expression is unaffected. A: *Xist* RNA FISH applied on *Xist*^{KO} and control NPC. Both panels depict examples of the RNA FISH; *Xist* RNA signal is shown in red and Dapi signal in blue. **B:**

Graphical illustration of the calculation of Spearman's ρ , using increasing window sizes, to compare interaction profiles identified in the 4C experiments. **C:** *Sox3* is located in a region contacted by the active viewpoints located in gene-dense areas (*Pctk1*, *MeCP2* and *Jarid1C*) on the active X chromosome. These contacts are lost on the X_i , in agreement with the silencing of *Sox3* during XCI (Fig. 6G). After depletion of *Xist* the 4C profile of *Sox3^{Xi}* more resembles the *MeCP2^{Xa}* profile. On the active X chromosome *Slitrk4* captures mainly inactive, gene-poor regions, opposite to sequences captured by *MeCP2*. Comparing the *Slitrk4* 4C profiles with and without *Xist* to the *MeCP2^{Xa}* profile, a stronger anti-correlation is observed in absence of *Xist*. Both these observations are indicative of a partial refolding of the X_i to X_a conformation. **D:** In total 11 genes were subjected to allele-specific expression analysis, 9 silenced (S) and 2 escaping XCI (E). The position of X_i originating transcript signals, indicative of escape, is indicated by the red asterisk. Of all genes tested only *Atp7a* was found to show de-repression, indicating most genes on X_i remain stably repressed upon *Xist* deletion.

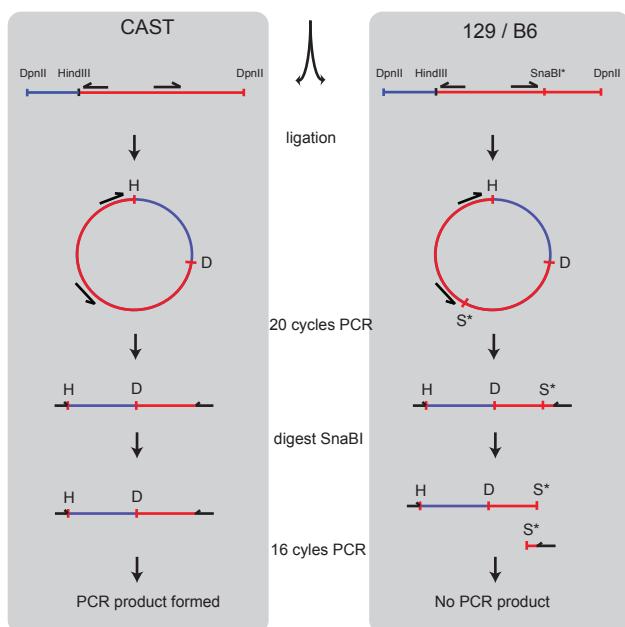
Supplemental Table S1: Active genes located in gene-dense areas share interactions identified on other chromosomes. Values indicated the percentage of overlapping region contacted by the different viewpoints comparing the different experiments. The first number in each cell indicates the percentage of interacting sequence shared between the 4C experiment on the left and the 4C experiment plotted on top. The second percentage represents the reciprocal analysis.

Supplemental Table S2: List of escaping genes. In total twenty X-linked genes, contacted by an escaping gene as measured by 4C, were either reported or confirmed to escape XCI. Escape from XCI was confirmed by allele-specific expression analysis, except for *Taf1* and *Ogt1*. For these two genes, gene

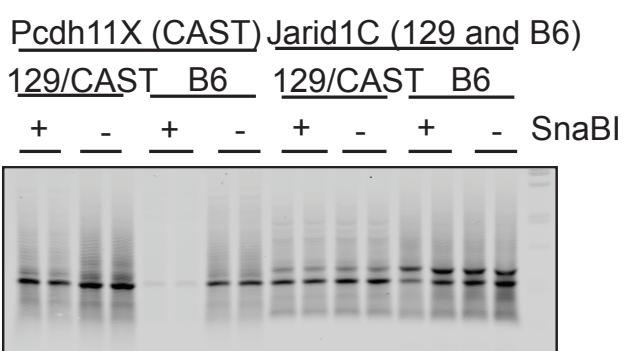
activity on the X_i was detected by RNA FISH in 82% of the cells, using a BAC probe (RP23-268G11) spanning both genes. Tissue specificity was determined based on data provided by <http://biogps.gnf.org/>.

Supplemental Table S3: List of RFLPs and primer sequences used in the different experiments.

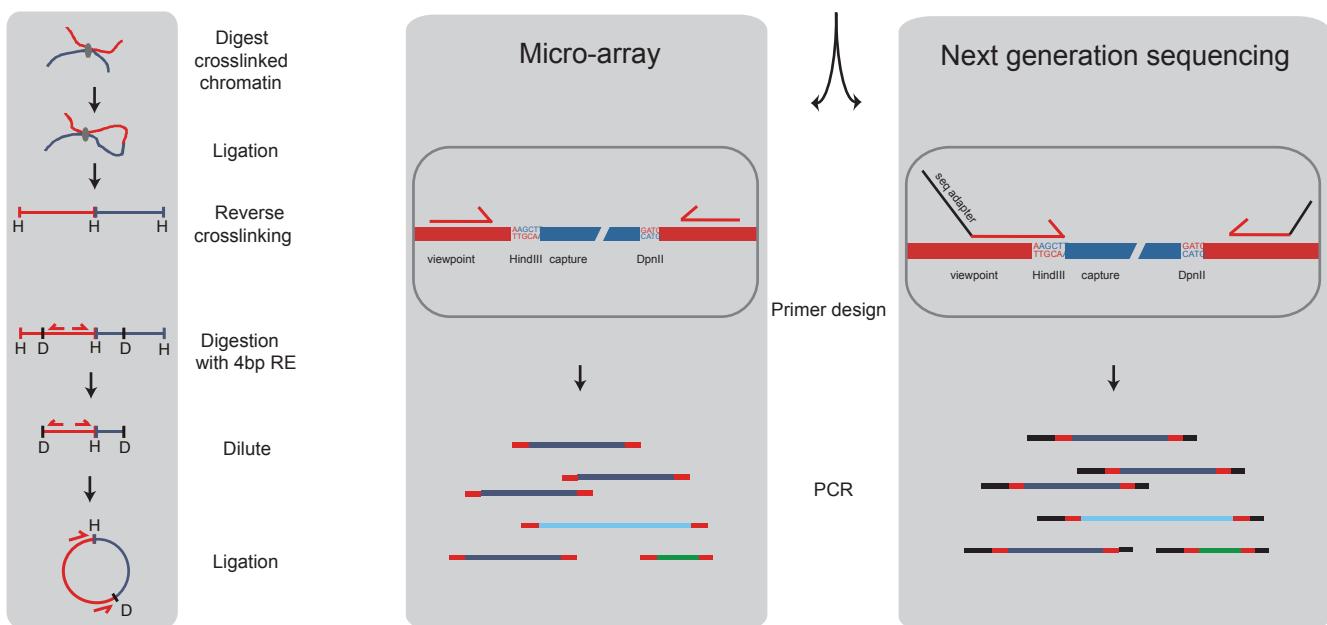
A



B

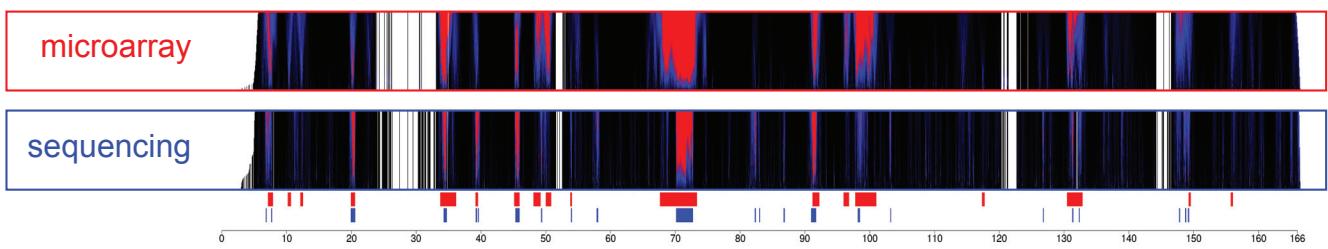


C



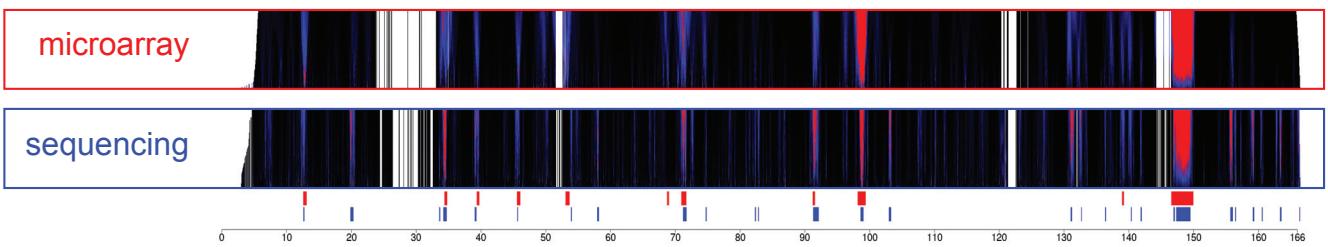
D

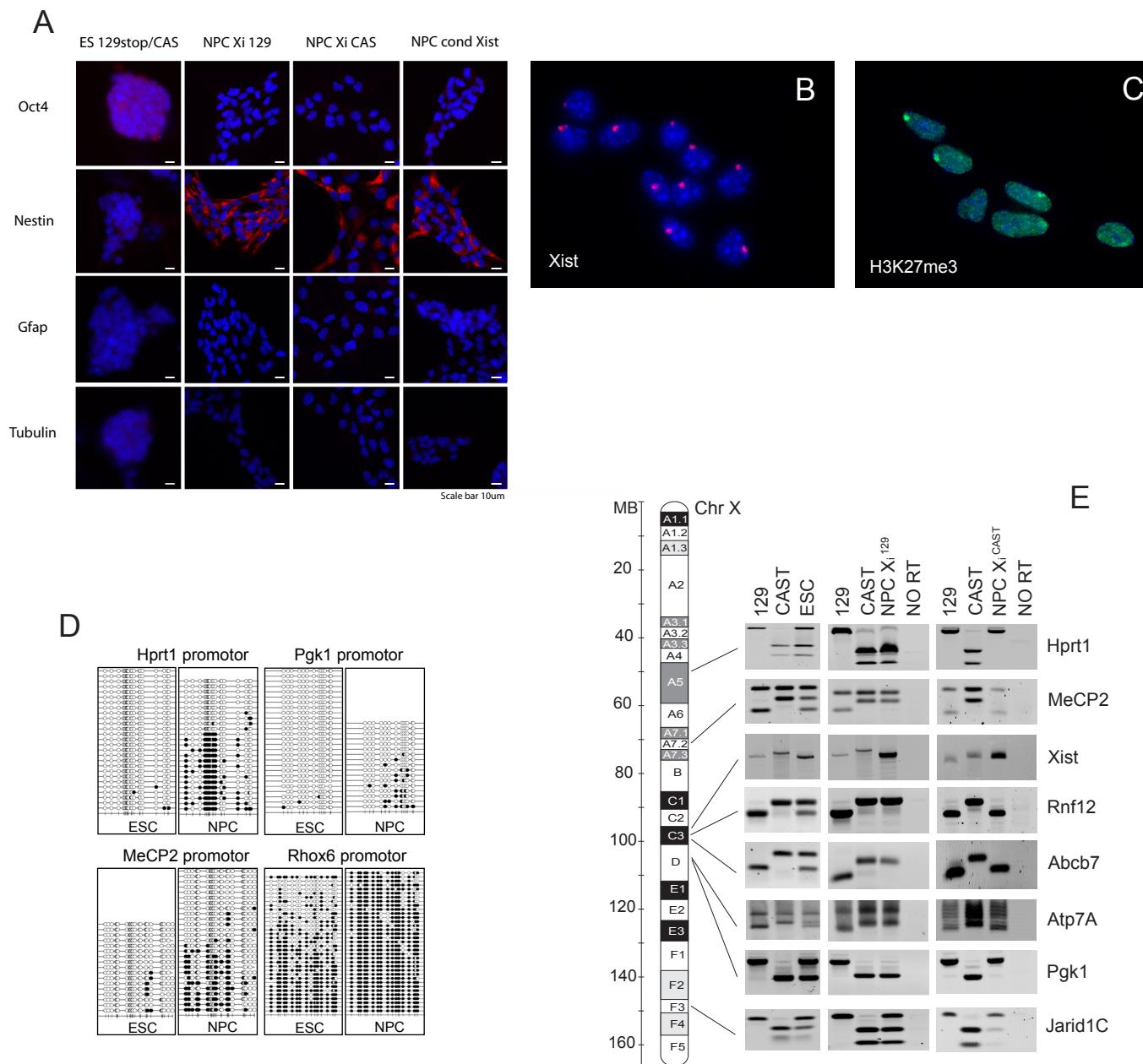
MeCP2



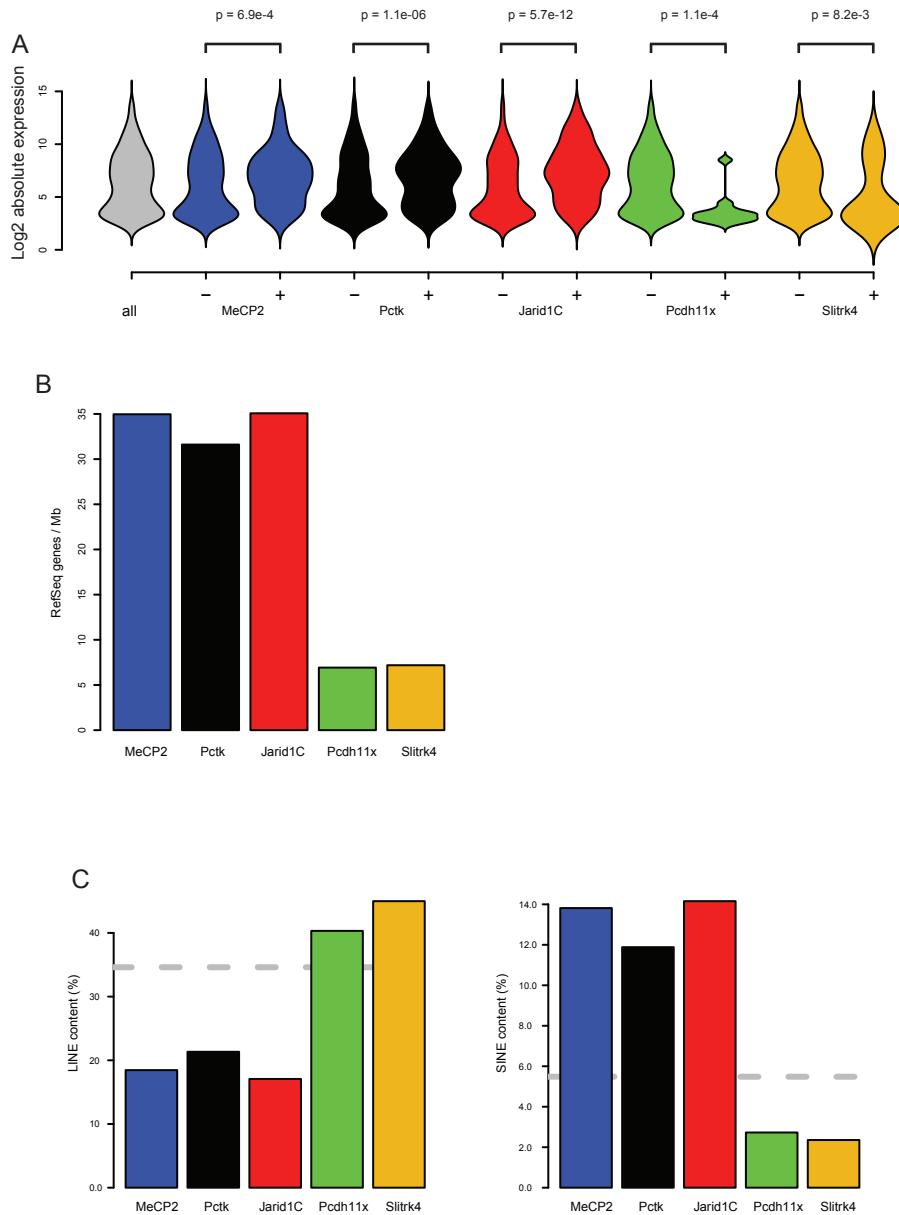
E

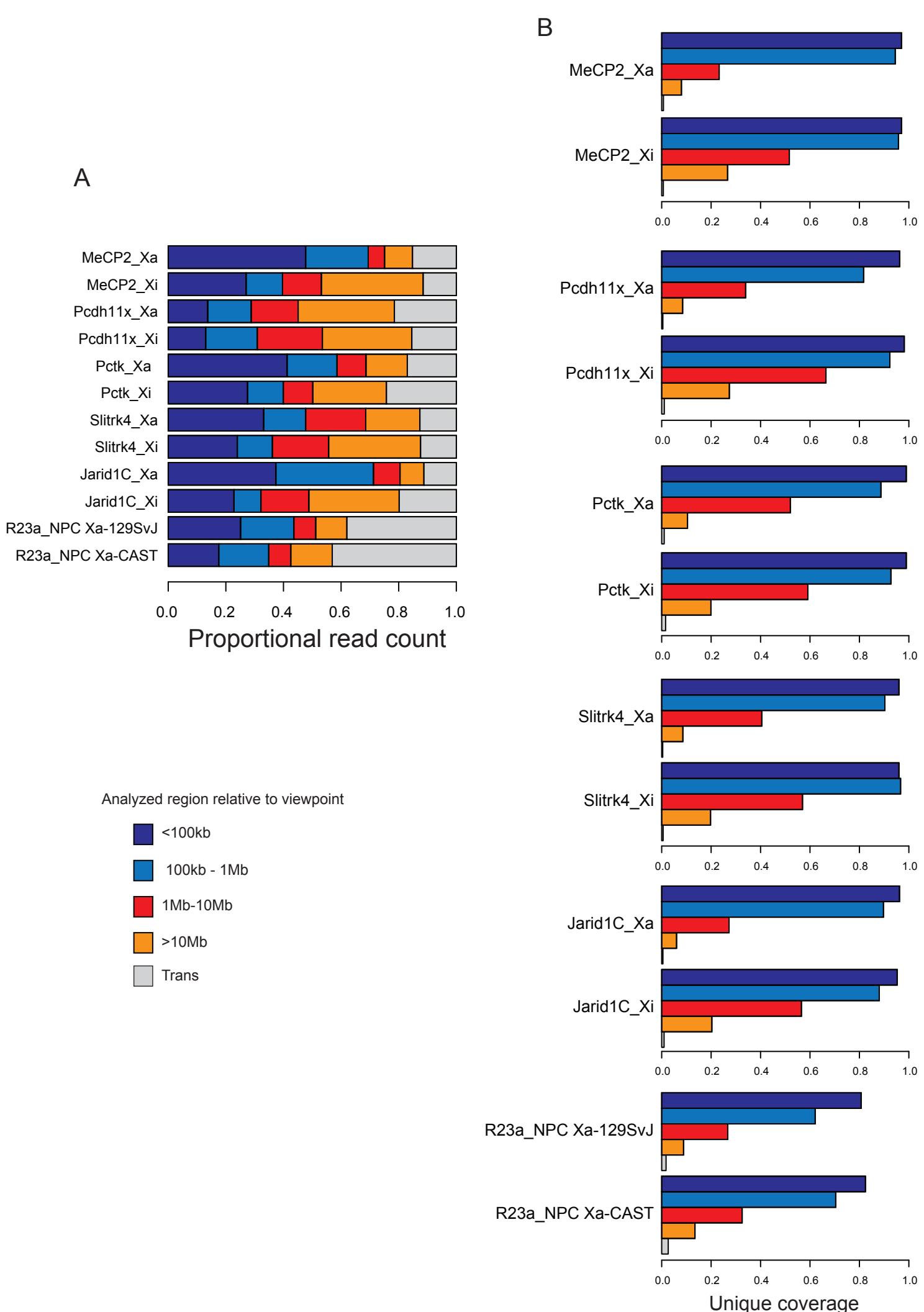
Jarid1C

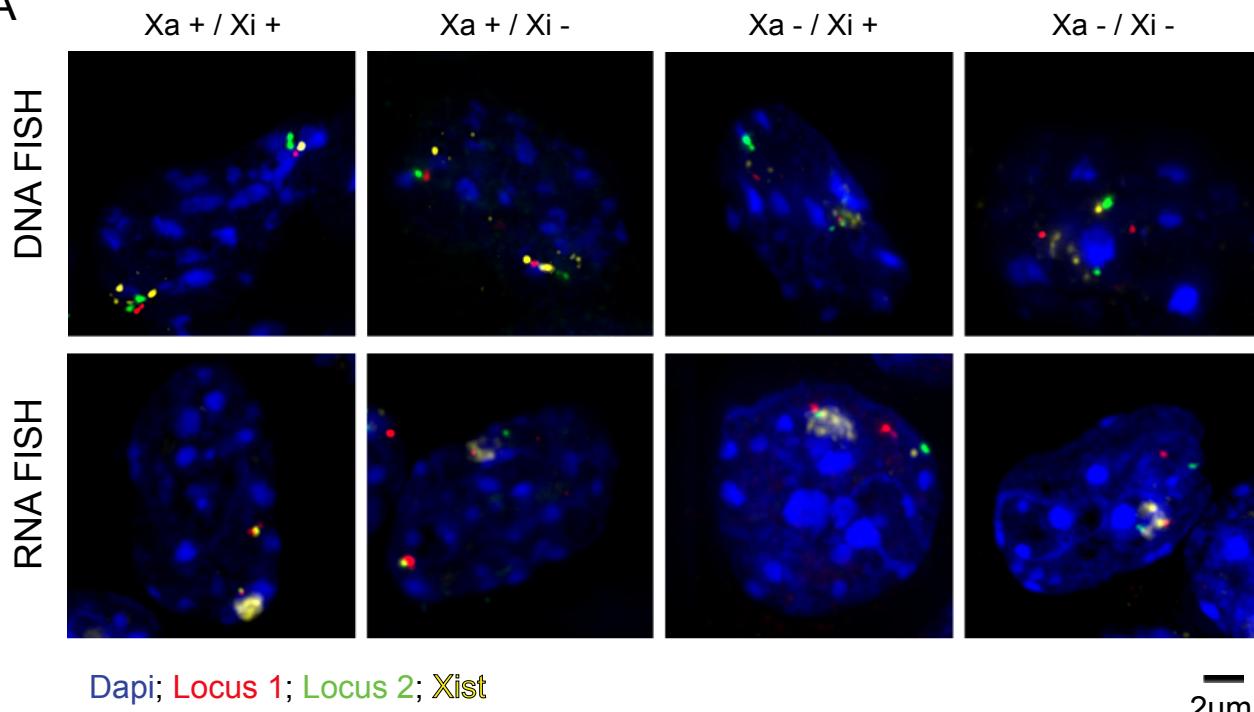
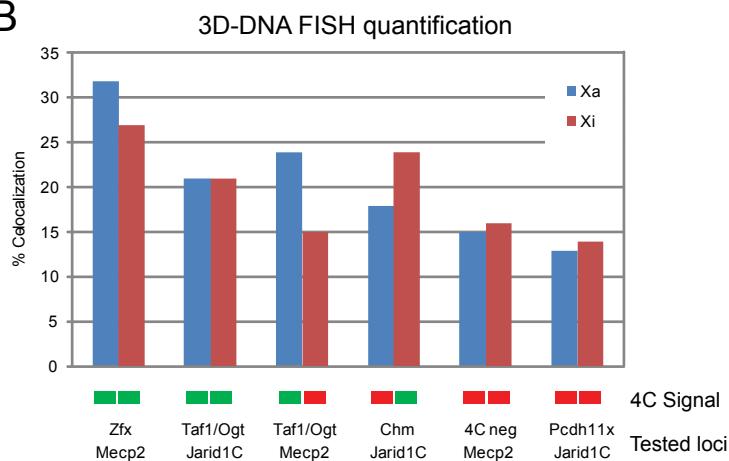
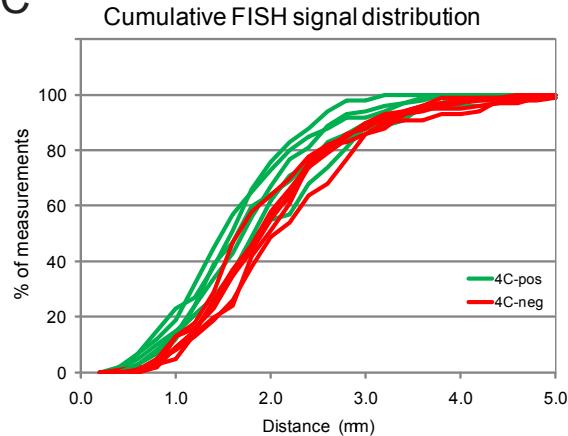
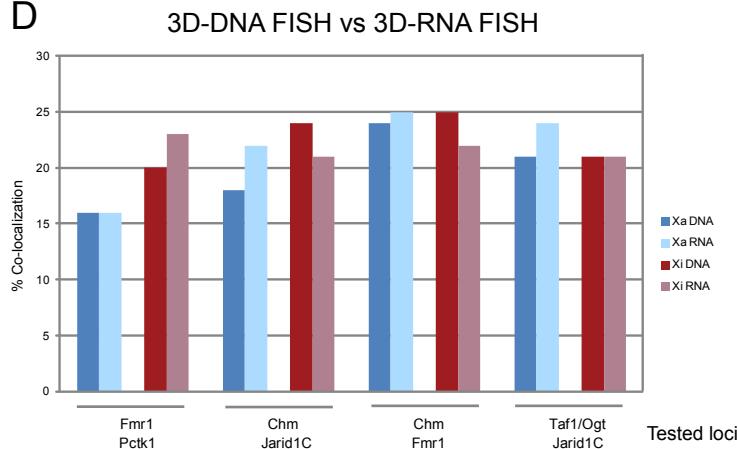
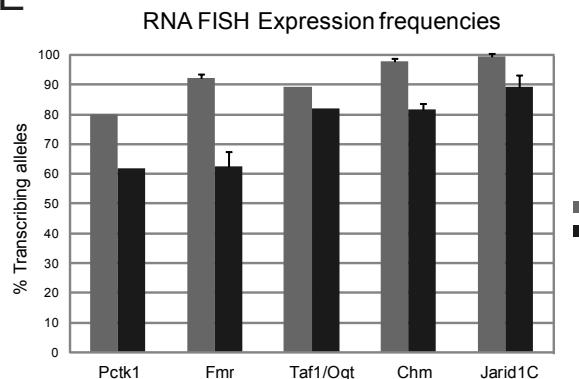




Splinter et al 2011, Supplemental Figure S3

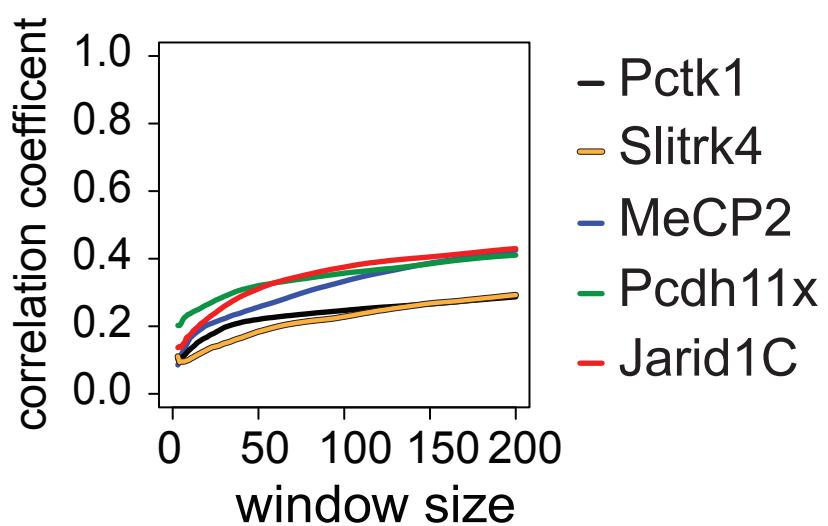


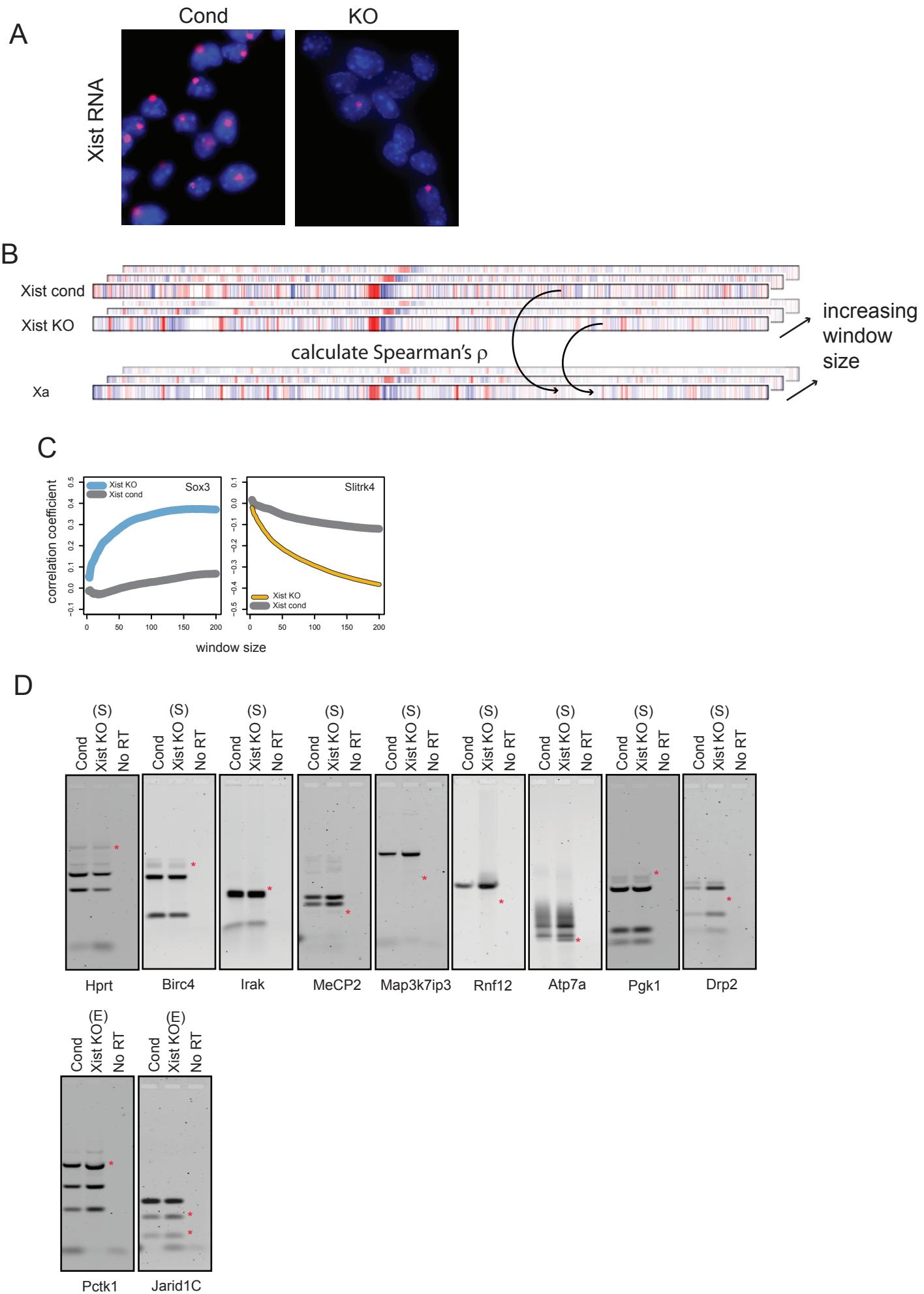


A**B****C****D****E**

Locus	BAC	Position
Pctk1	RP23-362P12	(20.2MB)
Fmr	RP24-183G11	(64.9MB)
MeCP2	RP23-77L16	(71.3MB)
4C neg	RP23-25P21	(87.2MB)
Zfx	RP23-269L6	(91.3MB)
Taf1/Ogt	RP23-268G11	(97.8MB)
Chm	RP23-118M16	(110.2MB)
Pcdh11x	RP23-20G23	(117.5MB)
Jarid1C	RP24-148H21	(148.7MB)

Supplemental Figure S6_Splinter et al. 2011





Splinter et al. 2011, Supplemental Table S1

	MeCP2_Xa_trans	Pctk1_Xa_trans	Pctk1_Xi_trans	Jarid1C_Xa_trans	Jarid1C_Xi_trans
MeCP2_Xa_trans	100.0% / 100.0%	47.1% / 49.0%	66.9% / 34.7%	29.1% / 53.5%	28.6% / 31.7%
Pctk1_Xa_trans	49.0% / 47.1%	100.0% / 100.0%	67.5% / 33.7%	28.0% / 49.6%	19.9% / 21.2%
Pctk1_Xi_trans	34.7% / 66.9%	33.7% / 67.5%	100.0% / 100.0%	16.7% / 59.3%	25.0% / 53.3%
Jarid1C_Xa_trans	53.5% / 29.1%	49.6% / 28.0%	59.3% / 16.7%	100.0% / 100.0%	39.1% / 23.5%
Jarid1C_Xi_trans	31.7% / 28.6%	21.2% / 19.9%	53.3% / 25.0%	23.5% / 39.1%	100.0% / 100.0%

Splinter et al 2011, Supplemental Table S2

gene ID	picked up by	gene position 1	gene position 2	domain start	domain end	gene name	source	tissue specific
ENSMUSG00000039231	Pctk1 and Jarid1C	7638297	7651886	7162709	7878720	Suv39h1	confirmed	-
ENSMUSG00000044148	Pctk1 and Jarid1C	12232005	12250794	12107199	13056309	810030o07Rik	Yang et al, 2010	-
ENSMUSG00000031010	Pctk1 and Jarid1C	12648624	12766815	12107199	13056309	Usp9x	confirmed	+
ENSMUSG0000000787	Pctk1 and Jarid1C	12858096	12871178	12107199	13056309	Ddx3x	Yang et al, 2010	-
ENSMUSG00000037369	Pctk1 and Jarid1C	17739701	17857062	17551677	17930443	Utx	Yang et al, 2010	-
ENSMUSG00000031065	Jarid1C	20265080	20277006	20082653	20518714	Pctk1	confirmed	+
ENSMUSG00000016409	Pctk1 and Jarid1C	34666790	34690741	34475296	35100569	Nkap	confirmed	-
ENSMUSG00000036551	Pctk1 and Jarid1C	34690693	34708837	34475296	35100569	Akap	confirmed	-
ENSMUSG00000085396	Pctk1	47908921	47988498	47642658	48182857	6720401G13Rik	Yang et al, 2010	
ENSMUSG0000000838	Pctk1 and Jarid1C	65931716	65971138	65726353	66151529	Fmr1	confirmed	+
ENSMUSG00000031386	Pctk1 and Jarid1C	71188131	71211696	71145255	71390304	Hcf1	confirmed	+/-
ENSMUSG00000031197	Pctk1	72759638	72780281	72623098	73138391	Vbp1	confirmed	-
ENSMUSG00000035150	Pctk1 and Jarid1C	91434046	91458201	91264822	91574717	Eif2s3x	Yang et al, 2010 and confirmed	-
ENSMUSG00000031314	Pctk1 and Jarid1C	98728073	98797128	98498880	99512449	Taf1		-
ENSMUSG00000034160	Pctk1 and Jarid1C	98835399	98879690	98498880	99512449	Ogt	confirmed by RNA FISH	-
ENSMUSG00000086503	Pctk1 and Jarid1C	100655714	100678556	100601273	100990400	Xist	Yang et al, 2010 and confirmed	-
ENSMUSG00000031226	Pctk1	102275095	102312429	102141536	102407292	2610029G23Rik	Yang et al, 2010	-
ENSMUSG00000025531	Pctk1 and Jarid1C	110154204	110299126	110088894	110414742	Chm	confirmed	-
ENSMUSG00000025332	Pctk1	148667563	148709078	148481293	149025730	Jarid1C	Yang et al, 2010 and confirmed	-
ENSMUSG00000035299	Pctk1 and Jarid1C	166123131	166428730	166166237	166525117	Midl	Yang et al, 2010	-

Splinter et al 2011, Supplemental Table S3

analysis	gene	SNP position	RFLP enzyme	4C digestion	allele specific strategy	digested allele	analyzed allele	primer 1	primer 2
4C	Pctk1	chrX: 20275611	Pvull	HindIII/Pvull-DpnII	I	CAST	129SVJ	3021-TCTTCTTTGAGAAGCTT	3023-ACCCAGACCCATTACAGTC
4C	Sox3	chrX: 58154777	Tsp509I	HindIII/Csp6I	II	CAST	129SVJ	3968-GGGTGATCAGGAAAGCTT	3971-ACATGGAGATGCTCAGATT
4C	Slitrk4	chrX: 61520329	Csp6I	HindIII/Csp6I	I	CAST	129SVJ	3957-TTACATCCTATAGCAAGCTT	3963-GCATGGAAAATTAAGTATTG
4C	MeCP2	chrX: 71283450	Csp6I	HindIII/Csp6I	I	CAST	129SVJ	2359-GGTTGGACACGAAAGCTT	2360-AGAAATAGCCCTCTATGGG
4C	Pcdh11x	chrX: 117542109	SnaBI	HindIII/Csp6I	II	129SVJ	CAST	3379-CAGTTGTGTCAAAAGCTT	3382-CCGTTAAATGACCTATA
4C	Pcdh11x	chrX: 117401474	DpnII	HindIII/DpnII	I	CAST	129SVJ	3583-GTTGGAAAACCCAAGCTT	3585-GAGGAATAATTGGTAGTTG
4C	Jarid1C	chrX: 148675298	DpnII	HindIII/DpnII	I	CAST	129SVJ	2361-TTGTGTTGTCAGACTT	3024-ACGTTAGGCAAATACAAACGG
4C	R23a	-	-	HindIII/DpnII	none	none	both	2720-GTAGTATTGATGAGACTT	2719-AAAATGGGTACATAGTTG
4C	Hbb-b1	-	-	HindIII/DpnII	none	none	both	2675-GAAATTGAGAAGCAAAGCTT	2397-GAGCATATAAGGTGAGGTAGG
analysis	gene	RFLP enzyme	digested allele	primer 1	primer 2	RNA type analyzed			
expression analysis	Fmr	Xhol	129	2823-GCAACTGCTCTTGAGTGG	2824-TGTTGTGTTCTCTGTCTCC	primairy transcripts			
expression analysis	Nkap	Hhal	129	2819-CTTAGGAGTTGGTCTCTCC	2820-TCCAGCAGTAAGAAGAAGG	primairy transcripts			
expression analysis	Usp	Ndel	129	2815-ATTGGTATTGCTGTGTC	2816-TGAGTTTCATCCAGAGACC	primairy transcripts			
expression analysis	Vpb	HaeIII	cas	2825-AGTACTGTATTGAGGAAGAGTGG	2826-CTACAAGCACCAGACATGC	primairy transcripts			
expression analysis	Xist	diff repeat length	-	1445-ACTGGGTCTCAGCGTGA	1447-GGAAATAGGTAAAGACACACTG	mature RNA			
expression analysis	Suv39h1	AflII	cast	1308-TGGGCACTCTACCTGTTG	1309-CTCCTCTGATTCTCCTGGG	primairy transcripts			
expression analysis	Pgk1	MseI	cas	2276-CGTGATGAGGGTGGACTAAC	2277-TAGTTGGACAGTGAGGCTCGG	mature RNA			
expression analysis	Hprt	SfaNI	cas	2294-TGGCAACATCAACAGGACTC	2299-AGTCCCAGCGTCGTGATTAG	mature RNA			
expression analysis	Pctk1	Pvull	cas	2264-CGACTTGTAGCGATGGAGC	2265-CAATGGTGGGTTGACAGGC	mature RNA			
expression analysis	Jarid1C	PmlI	cas	1440-AGGAGATGTTATTGGTGC	1441-GGTATTGATGTGAGAAAGGC	mature RNA			
expression analysis	Abc7l	NlaIII	129	2270-AAGCATTGGCAGTCTGACC	2271-TCTAGTATCACACATCTTAAACC	mature RNA			
expression analysis	Rlim/Rnf12	HaeIII	129	2278-GAGCCCCGATGAAAATAGAGC	2279-GGTGGCACTCTGTACTGC	mature RNA			
expression analysis	MeCP2	Ddel	129	2266-CATGGTAGCTGGGATGTTAGG	2267-GCAATCAATTCTACTTTAGAGCG	mature RNA			
expression analysis	Atp7a	HpyCH4III	cast	2274-GCCGCTTCATCTGTCTGTAG	2275-GCACACATTAGCAACTCTAAC	mature RNA			
expression analysis	Xiap/Birc4	mspI	129	3400-TCTTCAGAACCTCTATGGTG	3401-CTCGAAAGATCCAGAAATTG	mature RNA			
expression analysis	Map3k7ip3	Ecl136II	129	3402-TGAACCTCTCAAAGACCTG	3403-TGACTCTCTAAGCGTCTC	mature RNA			
expression analysis	Drp2	AluI	129	3404-AGATCCTGAAGACCAAC	3405-ACTAACATTGAGGGACAAGC	mature RNA			
expression analysis	Chm	Apol	129	3408-GGTGGAATCTATTGCTTCG	3049-CTGCTGATATCCGAAAGAGTC	mature RNA			
expression analysis	Irak	NlaIII	cast	3410-CTATACCTGCAAATTGACTG	3411-GCCACAGATAAAAGATCAGG	total RNA			
expression analysis	Hcfc	BamHI	129	3412-AACGCCACATTGACTATAC	3413-AGTCTGAGCTGTCCATTC	mature RNA			
expression analysis	Arhgap4	Apal	cast	3414-CATGATGGACTCCTACAAAC	3415-ATGGCCTGAGTCCTGAAC	mature RNA			
expression analysis	Lhfp1	Spel	129	3736-TATGTTGGAGAACATCA	3737-TGTGTTGTGAGGAAACAATA	total RNA			
expression analysis	Sox3	XmaI	cas	3740-CTTGTACCGAAGATGAGGAG	3741-GCGTACACGAAATAGCAAAT	total RNA			
expression analysis	Dmd	HpyCh4V	cas	3926-GTTCTTGCACCTCTAATG	3927-GAAGAGGGCAGAACAAATT	total RNA			
expression analysis	Slitrk4	Scal	129	3932-ATTAACCTCTCCAGCCACA	3933-CCTGAGTATCTTATTGGA	total RNA			
analysis	gene	primer 1	primer 2						
bisulfite analysis	Hprt	2832-GTGGGGATGTTTTAGTGAAGTT	2833-CAAAACCTAAACATCCCTCTCATAC						
bisulfite analysis	Pgk1	2834-TTATTTGGGAGTTGAGAAAAGTAGAG	2835-AAAAAAAACCTCCTAACTAAAAAAA						
bisulfite analysis	MeCP2	2836-TATTTAGGTTTAGTTGTGTAAG	2837-ATAAAAACTAAACCCCTTAAAC						
bisulfite analysis	RhoX6	2838-GAGTTTGAATTGTTTTGTG	2839-CCAAACTATTCTCCTTACCTCCT						
analysis	gene	floxed allele	primer 1	primer 2					
lox recombination	Xist	129	3373-AGCACTATCAGGCAGTTGAC	3374-CACTCTACCCCTGCCTTTTC					
normalization control	Hbb-b1	-	123-CCGAGAAATATGCTTGTATC	124-CAACTGATCCTACCTCACCT					