nature structural & molecular biology

SUPPLEMENTARY INFORMATION

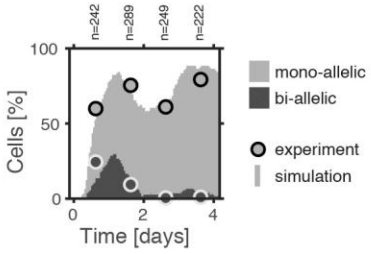
https://doi.org/10.1038/s41594-019-0214-1

In the format provided by the authors and unedited.

A symmetric toggle switch explains the onset of random X inactivation in different mammals

Verena Mutzel[1], Ikuhiro Okamoto[2,3], Ilona Dunkel[1], Mitinori Saitou[4,5,6], Luca Giorgetti[7], Edith Heard[8,9] and Edda G. Schulz[1]*

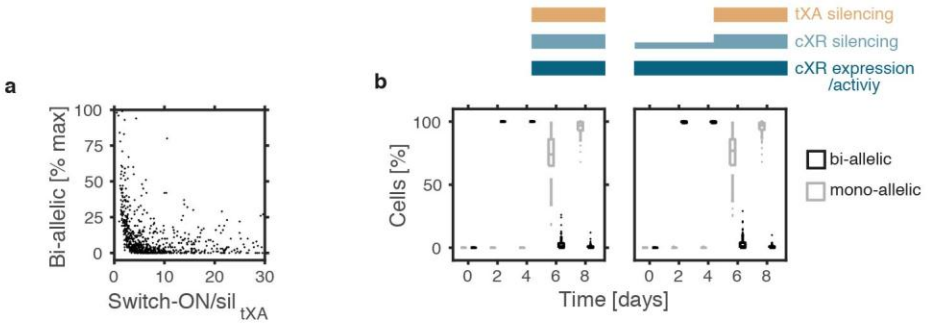
[1]Otto Warburg Laboratories, Max Planck Institute for Molecular Genetics, Berlin, Germany. [2]Department of Anatomy and Cell Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [3]Japan Science and Technology (JST), Exploratory Research for Advanced Technology (ERATO), Kyoto, Japan. [4]Institute for the Advanced Study of Human Biology (ASHBi), Kyoto University, Kyoto, Japan. [5]Department of Anatomy and Cell Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [6]Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto, Japan. [7]Friedrich Mieschner Institute for Biomedical Research, Basel, Switzerland. [8]Institut Curie, PSL Research University, CNRS UMR3215, INSERM U934, Paris, France. [9]European Molecular Biology Laboratory (EMBL), Directors' research unit, Heidelberg, Germany. *e-mail: edda.schulz@molgen.mpg.de



Supplementary Figure 1

The cXR-tXA model can reproduce up-regulation of *Xist* in differentiating mESCs.

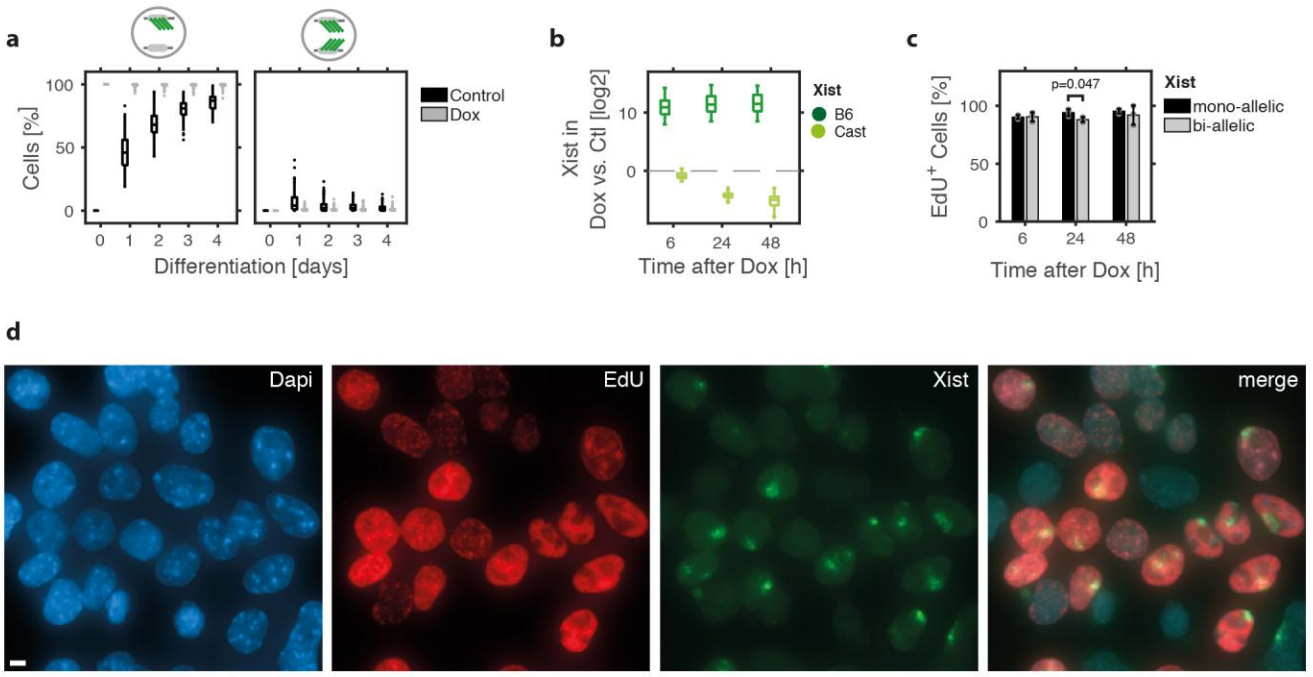
Fraction of cells exhibiting mono-allelic (light grey) and bi-allelic Xist expression (dark grey) during differentiation of mESC line TX1072. Experimental data (circles) is shown together with a simulation using the parameter set that best explains the data. The number of cells analyzed is given on top. The data was pooled from 3 independent experiments.



Supplementary Figure 2

Transient bi-allelic up-regulation of *Xist* in the cXR-tXA model.

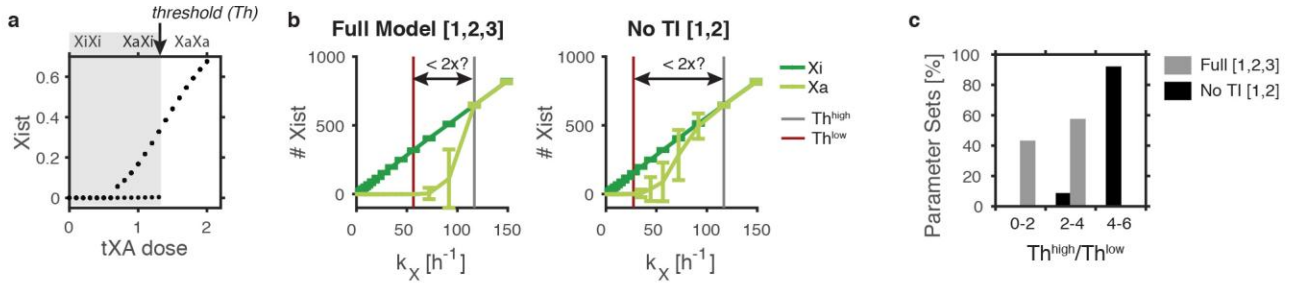
(a) For all parameter sets that reproduced mono-allelic *Xist* up-regulation in the cXR-tXA model, the maximal fraction of cells with bi-allelic *Xist* expression observed during the simulation is shown as a function of the the ratio of switch-ON time (first time point, when *Xist* levels reach 20% of the high steady state) and tXA silencing delay (sil_{tXA}). If *Xist* up-regulation is slow (high Switch-ON time), it will normally occur one allele at a time. Subsequent silencing will shift the system to the bistable regime (cp. Fig. 2e) and thereby lock in the mono-allelic state before *Xist* up-regulation from the other X chromosome occurs. This results in a low frequency of bi-allelically expressing cells as observed in mice. If *Xist* up-regulation is rapid and silencing is slow (long silencing delay sil_{tXA}), *Xist* will initially be expressed from two alleles as observed in rabbit embryos. In this scenario the choice of the inactive X can subsequently occur through mono-allelic silencing of tXA and cXR. Alternatively, silencing of both alleles might reverse *Xist* up-regulation completely as *Xist* expression is unstable if both tXA alleles are silenced such that the cell can undertake a second attempt to reach the mono-allelic state. **(b)** Simulation of bi-allelic expression upon reduced *Xist*-mediated silencing as observed in human embryos, assuming that in the first 4 days of the simulation either silencing and cXR expression is absent (left) or that cXR is silenced partially (dampening), while tXA is unaffected by *Xist* (right). Boxplots show the percentage of mono- and bi-allelically expressing cells for 100 randomly chosen parameter sets that can reproduce mono-allelic *Xist* up-regulation (center line, median; box limits, upper and lower quartiles; whiskers, most extreme data points not considered outliers; points, outliers).



Supplementary Figure 3

Bi-allelic up-regulation of Xist is reversible.

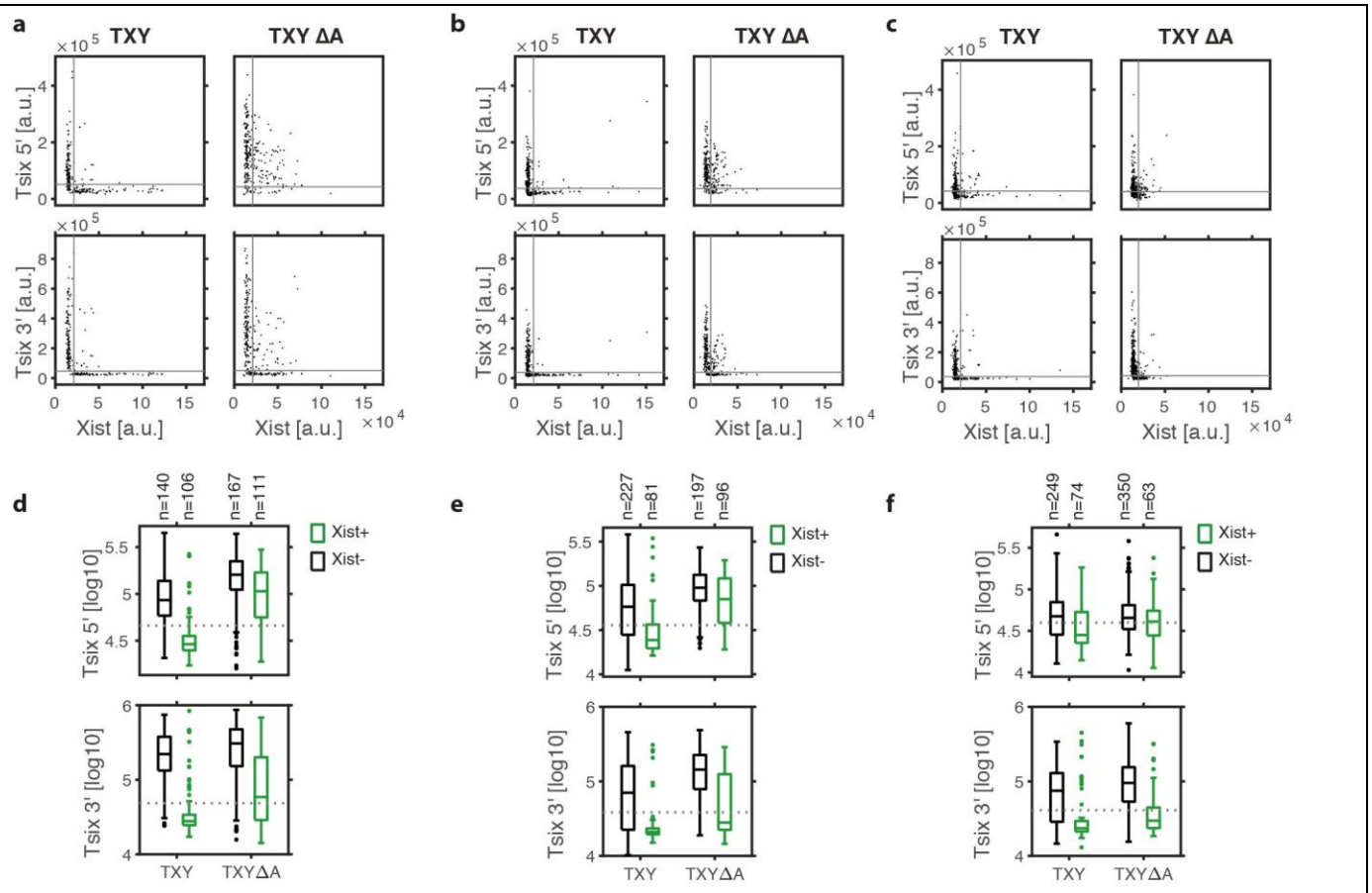
(a) Simulation of doxycycline treatment one day before the onset of differentiation (linked to Fig. 4a-c). Boxplots show the frequency of mono-allelic (left) and bi-allelic *Xist* expression (right) in dox-treated (grey) and control cells (black) for 100 parameter sets that could reproduce mono-allelic *Xist* up-regulation. **(b)** Boxplots show the simulation results for artificial bi-allelic *Xist* induction as described in Fig. 4e in the main text, using the same parameters sets as in (a). On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually. **(c-d)** Bi-allelic *Xist* up-regulation is artificially induced by treating TX1072dT cells with doxycycline after 48h of differentiation (cp. Fig. 4e). Cells were treated with EdU to assay proliferation through measuring its incorporation into DNA during replication. EdU was labeled fluorescently through Click-it chemistry and *Xist* was visualized by RNA FISH. (c) The EdU-positive fraction was quantified at the indicated time points within cells expressing *Xist* mono- (black) and bi-allelically (grey). Mean and s.d. of n=3 independent experiments are shown, in each replicate at least 50 cells were counted per group, except for bi-allelic cells at 48h. * paired two-sample two-sided T-test. Scale bar indicates 5µm.



Supplementary Figure 4

Transcriptional interference can generate a precise threshold, which is required for reliable mono-allelic *Xist* up-regulation

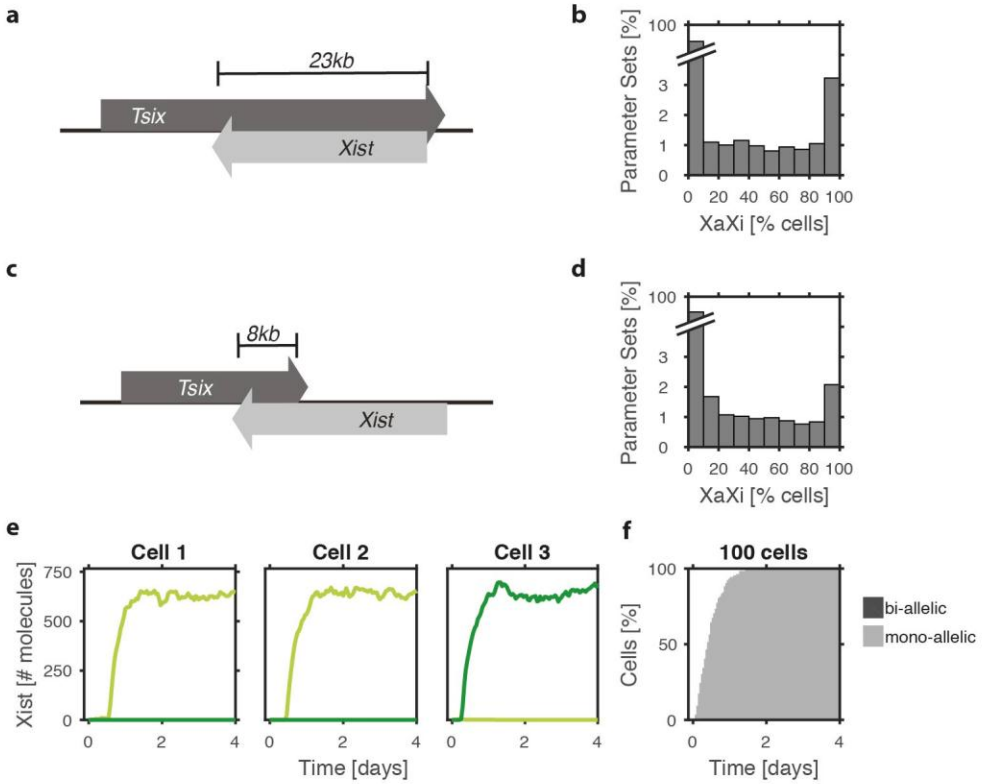
(a) Steady state *Xist* levels simulated deterministically (see Fig. 2e) to indicate that a sharp threshold is required between a single (1x) and a double (2x) tXA dose. **(b)** Maintenance of the XaXi state was simulated by initiating an allele either from the Xa (light green) or from the Xi state (dark green). For an example parameter set (k_T=113 h^{-1}, $t_{1/2}^{repr}$ =0.7 h) mean and standard deviation of *Xist* expression across 100 cells from the chromosomes that initiated as Xa (light green) and Xi (dark green), respectively, is shown for different values of k_X for the full Xist-Tsix model (left) and the reduced model without transcriptional interference (right). The vertical lines indicate the k_X threshold value, above which >1 (Th^{low}, red) or >99 (Th^{high}, grey) out of 100 cells up-regulate *Xist* from the Xa. **(c)** Distribution of the Th^{high}-to-Th^{low} ratio (red and grey in (b)) across all parameter sets of the Full model (grey) and the reduced model without transcriptional interference (black). Since tXA reduces k_X 2-fold upon *Xist* up-regualtion a threshold ratio of <2 is required to allow reliable *Xist* up-regulation with a double dose (2x) of tXA and stable maintenance of the XaXi state with a single dose (1x). This is only possible in the Full model with transcriptional interference.



Supplementary Figure 5

Transcriptional interference at the *Xist*–*Tsix* locus.

(a-f) TXY and TXYΔA ESCs were treated with doxycycline for 24 hours and nascent transcription of *Xist* and *Tsix* (5' and 3') was assessed by RNA FISH. (a-c) Quantification of 3 biological replicates, where each dot represents the measured signal intensities of a single allele. Grey lines indicate the detection threshold estimated from negative control regions. (d-f) Box plots of *Tsix* signal intensity at *Xist+* (green) and *Xist-* alleles (black) in the two cell lines as indicated for the data shown in (a-c); dotted lines indicate the detection threshold (center line, median; box limits, upper and lower quartiles; whiskers, most extreme data points not considered outliers; points, outliers).



Supplementary Figure 6

Also a reduced overlap between *Xist* and *Tsix* as in the human locus allows mono-allelic up-regulation of *Xist*.

(a,c) Schematic representation of the *Xist-Tsix* locus architecture in the mouse (a) and the human (c) genome, respectively. **(b,d)** Distribution of the mean frequency of mono-allelic *Xist* up-regulation across all parameter sets tested, in simulations assuming the locus architecture shown in (a) and (c), respectively. For details see supplemental note 1 (section 3.5). **(e-f)** Simulation of *Xist* up-regulation using the model with the human locus architecture in (c) for one example parameter set, showing three individual cells (e) and a population of 100 cells (f). Light and dark green in (e) represent *Xist* levels expressed from the two X chromosomes, light and dark grey in (f) represent mono- and bi-allelic *Xist* expression, as indicated.

SUPPLEMENTARY NOTES: MODEL DESCRIPTION

1 Supplementary Note 1: Model Comparison

1.1 Single regulator models

To investigate the requirements for mono-allelic and female-specific *Xist* expression, all X-linked *Xist* regulators were grouped into 8 different classes according to whether they activate (XA) or repress (XR) *Xist*, whether they act *in cis* (c) or *in trans* (t) and whether they are silenced by *Xist* or escape XCI. To simulate regulation of *Xist* by each regulator class, 8 different ordinary differential equation (ODE) models were constructed as follows.

	Chr. X_1	**Chr. X_2**	
Xist	$\frac{dx_1}{dt} = f(r) - x_1$	$\frac{dx_2}{dt} = f(r) - x_2$	(1)
- cis regulator r	$f(r) = a + b\frac{{r_1}^n}{{r_1}^n + K^n}$	$f(r) = a + b\frac{{r_2}^n}{{r_2}^n + K^n}$	
- trans regulator r	$f(r) = a + b\frac{\left(0.5(r_1 + r_2)\right)^n}{\left(0.5(r_1 + r_2)\right)^n + K^n}$ XA: a=0, b=1 XR: a=1, b=-1		
Regulator - silenced	$\frac{dr_1}{dt} = 1 - \frac{{x_1}^n}{{x_1}^n + K^n} - r_1$	$\frac{dr_2}{dt} = 1 - \frac{{x_2}^n}{{x_2}^n + K^n} - r_2$	(2)
- escaping	$\frac{dr_1}{dt} = 1 - r_1$	$\frac{dr_2}{dt} = 1 - r_2$	

Activation or repression of *Xist* by a regulator and silencing of the regulator by *Xist* are described as Hill functions with two parameters: a Hill coefficient *n* that describes the cooperativity or threshold behavior of the interaction and a threshold *K*, which is the regulator level where activation or repression is half-maximal. Degradation of *Xist* and its regulators is assumed to occur with first-order kinetics with a degradation rate of $1h^{-1}$. Since the production rates can maximally be 1 and degradation rates are set to 1, the levels of *Xist* and its regulators can vary between 0 and 1. Since each interaction is described by two parameters, a Hill coefficient *n* and a threshold *K*, each model has either 2 (escaping regulator) or 4 parameters (silenced regulator).

1.2 Simulating mono-allelic expression

To assess which network can maintain stable mono-allelic expression (simulation 1), cells with one active X (Xa), where *Xist* expression is low and with one inactive X (Xi), where *Xist* expression is high were simulated with the following initial conditions.

Xist (Xi) - x_1	**Xist (Xa) - x_2**	**Regulator (Xi) - r_1**	**Regulator (Xa) - r_2**
1	0.01	0.01 (silenced) 1 (escaping)	1

Each network was simulated with at least 10,000 parameter sets, where parameter values were randomly drawn from a uniform distribution between 1 and 5 for Hill coefficients (n) and from a logarithmic distribution between 0.01 and 10 for threshold parameters (K). Each parameter set was simulated for 100h using the ode23tb integrator in Matlab. The final state of this simulation was then used as initial conditions to solve the equation system for the steady state using the function fsolve in Matlab and the results were rounded to 6 digits. A parameter set was classified as mono-allelic, if Xist(Xi)>10*Xist(Xa) at the steady state. For the cXR (*cis*-acting *Xist* repressor) model 20% of parameter sets maintained mono-allelic expression, but none of the other models was able to do so.

1.3 Two-regulator models

To test whether other regulator classes than cXR would be able to maintain mono-allelic *Xist* expression, when combined with a second regulator, we built another 28 networks, each containing two different regulators A and B. Equation (1) in section 1.1 was thus modified as follows.

	Chr. X_1	**Chr. X_2**	
Xist	$\frac{dx_1}{dt} = f(r_A) \cdot f(r_B) - x_1$	$\frac{dx_2}{dt} = f(r_A) \cdot f(r_B) - x_2$	(3)

The resulting models have between 4-8 parameters depending on whether the regulators are silenced or escape.

1.4 Simulating mono-allelic expression in two-regulators models

For each of the 28 two-regulator models >10,000 randomly chosen parameter sets were simulated as described in section 1.2 As shown in table A1 only networks containing a cXR can maintain mono-allelic expression.

	cXA	**ecXA**	**tXA**	**etXA**	**cXR**	**ecXR**	**tXR**	**etXR**
cXA	0	0	0	0	10.2	0	0	0
ecXA		0	0	0	14.6	0	0	0
tXA			0	0	14.4	0	0	0
etXA				0	14.9	0	0	0
cXR					20.0	8.1	9.4	8.8
ecXR						0	0	0
tXR							0	0
etXR								0

Table A1. For each model containing either one regulator (diagonal) or two regulators the percentage of parameter sets that maintain mono-allelic expression is given.

1.5 Simulating male cells and bi-allelic expression in females

Since *Xist* is expressed from exactly one chromosome in each female cell, the underlying network should maintain mono-allelic expression, but destabilize the bi-allelic state in female cells and prevent *Xist* expression in male cells with only one X chromosome. To test which network would fulfill these criteria, three additional simulations were performed, where female cells initiated from an XiXi state (simulation 3), where male cells initiated from an Xi (simulation 4) or from an Xa state (simulation 2). To simulate male cells the models were modified as follows such that only a single X chromosome would be present.

	Chr. X_1	
Xist	$\frac{dx_1}{dt} = f(r) - x_1$	(4)
- cis regulator	$f(r) = a + b\frac{{r_1}^n}{{r_1}^n + K^n}$	
- trans regulator	$f(r) = a + b\frac{(0.5r_1)^n}{(0.5r_1)^n + K^n}$ XA: a=0, b=1 XR: a=1, b=-1	
Regulator - silenced	$\frac{dr_1}{dt} = 1 - \frac{{x_1}^n}{{x_1}^n + K^n} - r_1$	(5)
- escaping	$\frac{dr_1}{dt} = 1 - r_1$	

For each mono-allelic parameter set identified in sections 1.2 and 1.4 three additional simulations were performed with the following initial conditions.

	Model	**Xist-1**	**Xist-2**	**Regulator-1**	**Regulator-2**
XiXi	female	0.99	1	0.01 (silenced) 1 (escaping)	0.01 (silenced) 1 (escaping)
Xi	male	1		0.01 (silenced) 1 (escaping)	
Xa	male	0.01		1	

Based on the steady state of $\overline{x_1}$ (Xist^{high} state) in the mono-allelic simulation in section 1.2, parameter sets were classified as follows.

class	**simulation**	**rule**
XiXi unstable	XiXi	$\overline{x_1} < 10 \cdot Xist^{high}$ OR $\overline{x_2} < 10 \cdot Xist^{high}$
Xi unstable	Xi	$\overline{x_1} < 10 \cdot Xist^{high}$
Xa stable	Xa	$\overline{x_1} < 10 \cdot Xist^{high}$

As shown in table A2 only a single model, namely the cXR-tXA model could maintain the XaXi and Xa states in female and male cells, respectively, while destabilizing both the XiXi and Xi states. All parameter sets that destabilized the XiXi state also destabilized the Xi state in male cells.

	cXR-cXA	**cXR-ecXA**	**cXR-tXA**	**cXR-etXA**	**cXR**	**cXR-ecXR**	**cXR-tXR**	**cXR-etXR**
XiXi unstable	0	0	25.4	0	0	0	0	0
Xi unstable	0	0	38.6	16.4	0	0	0	0
Xa stable	100	100	100	100	100	100	97.3	94.8

Table A2. Percentage of parameter sets in each class among all parameter set that can maintain the XaXi state.

2 Supplementary Note 2: The cXR-tXA model

The cXR-tXA model contains a *cis*-acting positive feedback loop mediated by a *cis*-acting repressor (cXR) and a *trans*-acting negative feedback, mediated by a *trans*-acting activator (tXA). The analysis in section 1 shows that the cXR-tXA model can explain maintenance of the correct *Xist* expression pattern (post-XCI state) for a subset of parameter sets. In the next step it was tested whether and under which conditions the model could also reproduce mono-allelic up-regulation of *Xist*, where a transition from the XaXa state (pre-XCI) to the XaXi state (post-XCI) occurs. Since a symmetry-breaking event, initiated by stochastic fluctuations, is required to transition from a symmetric

XaXa state to an asymmetric XaXi state, *Xist* up-regulation was simulated in a stochastic manner.

2.1 Stochastic simulation of the cXR-tXA model

Because absolute levels do not affect the qualitative results of the deterministic simulations in section 1 the equations were formulated in a way that the levels of *Xist* and its regulators would be scaled between 0 and 1. In stochastic simulations by contrast absolute abundances can strongly affect the allelic variability. To perform the simulations in a realistic regime, scaling factors were added to the model in section 1 (p_{21}, p_{22}, p_{23}) such that the maximal levels of *Xist* and its regulators would vary between 50 and 500 molecules per allele. Additionally, the rate of Xist RNA degradation was adapted to experimental measurements which have estimated the Xist RNA half-life to lie between 2 and 6h[1,2]. We therefore use the mean (4h), which results in a degradation rate of 0.1733 h^{-1} ($\ln(2)/t_{1/2}$). Since the molecular identity and thus also the degradation kinetics of cXR and tXA are unknown their degradation rates remain set to 1 h^{-1} but the unknown degradation kinetics are indirectly accounted for by assuming that silencing of cXR and tXA might occur with variable kinetics as described below. The ODE formulation of the model was thus modified as follows.

	Chr. X_1	**Chr. X_2**	
Xist	$\frac{dx_1}{dt} = p_{21} \cdot f(cXR_1) \cdot f(tXA) - 0.1733 * x_1$	$\frac{dx_2}{dt} = p_{21} \cdot f(cXR_2) \cdot f(tXA) - 0.1733 * x_2$	(6)
	$f(cXR_1) = 1 - \frac{{cXR_1}^{p_{13}}}{{cXR_1}^{p_{13}} + (p_{22} \cdot p_{14})^{p_{13}}}$	$f(cXR_2) = 1 - \frac{{cXR_2}^{p_{13}}}{{cXR_2}^{p_{13}} + (p_{22} \cdot p_{14})^{p_{13}}}$	
	$f(tXA) = \frac{\left(0.5(tXA_1 + tXA_2)\right)^{p_{11}}}{\left(0.5(tXA_1 + tXA_2)\right)^{p_{11}} + (p_{23} \cdot p_{12})^{p_{11}}}$		
tXA	$\frac{dtXA_1}{dt} = p_{23}\left[1 - \frac{{x_1}^{p_3}}{{x_1}^{p_3} + (p_{21} \cdot p_4)^{p_3}}\right] - tXA_1$	$\frac{dtXA_2}{dt} = p_{23}\left[1 - \frac{{x_2}^{p_3}}{{x_2}^{p_3} + (p_{21} \cdot p_4)^{p_3}}\right] - tXA_2$	(7)
cXR	$\frac{dcXR_1}{dt} = p_{22}\left[1 - \frac{{x_1}^{p_5}}{{x_1}^{p_5} + (p_{21} \cdot p_6)^{p_5}}\right] - cXR_1$	$\frac{dcXR_2}{dt} = p_{22}\left[1 - \frac{{x_2}^{p_5}}{{x_2}^{p_5} + (p_{21} \cdot p_6)^{p_5}}\right] - cXR_2$	(8)

The kinetics of gene silencing do not effect the steady states of a system that was analyzed above in section 1, but will influence the ability to switch between states. Therefore two additional parameters were added to describe how long after *Xist* up-regulation cXR (p_7) and tXA (p_8) will be silenced. This is implemented by assuming that *Xist* will transition through p_7 or p_8 silencing intermediates x', x'', x'''... with a rate of 1h^{-1} before reaching the silencing competent states, x_1 or x_2 (Fig. A1).

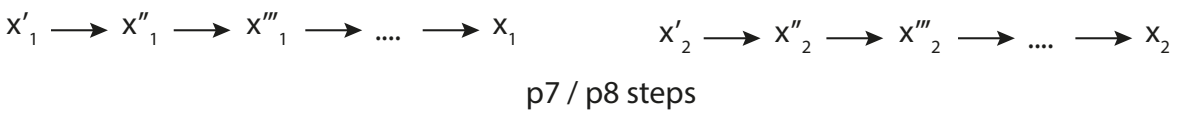


Figure A1. Schematic representation of the implementation of a silencing delay.

2.2 Simulating mono-allelic *Xist* up-regulation

To obtain more parameter sets that fulfill the requirements of XCI the ODE simulations of the cXR-tXA model described in section 1 were repeated for 300,000 parameter sets randomly drawn from the distributions described in section 1.2. Each parameter set that maintained the XaXi and Xa states, while destabilizing the XiXi and Xi states (3.8%) was combined with 10 sets of silencing delays p_7 and p_8 randomly drawn from a uniform distribution between 1 and 20 and scaling factors p_{21}, p_{22} and p_{23} randomly drawn from a logarithmic distribution between 50 and 500. As shown in table M3, the resulting model has 13 parameters that were varied between simulations.

Parameter	**Function**	**Value**
p_3	Xist --| tXA, Hill coefficient n	1 … 5
p_4	Xist --| tXA, threshold K	0.01 ... 10 (log distributed)
p_5	Xist --| cXR, Hill coefficient n	1 … 5
p_6	Xist --| cXR, threshold K	0.01 ... 10 (log distributed)
p_7	silencing delay cXR	1 ... 20 h
p_8	silencing delay tXA	1 ... 20 h
p_{11}	tXA --> Xist, Hill coefficient n	1… 5
p_{12}	tXA --> Xist, threshold K	0.01 ... 10 (log distributed)
p_{13}	cXR --| Xist, Hill coefficient n	1 … 5
p_{14}	cXR --| Xist, threshold K	0.01 ... 10 (log distributed)
p_{18}	transition rate between silencing intermediates	$1h^{-1}$
p_{21}	scaling factor Xist	50 … 500 (log distributed)
p_{22}	scaling factor cXR	50 … 500 (log distributed)
p_{23}	scaling factor tXA	50 … 500 (log distributed)

Table A3. Parameters in the tXA-cXR model

For each parameter set, 100 cells were simulated for 100h starting from an XaXa state (x_1=x_2=0, cXR_1=cXR_2=p_{22}, tXA_1=tXA_2=p_{23}). The simulations were performed in Julia using the Gillespie algorithm[3] and run on a computing cluster. Each chromosome was classified as *Xist* positive (Xist+) or negative (Xist-) for each hour of the simulation, depending on whether the mean *Xist* level exceeded 20% of the $Xist^{high}$ state estimated from the ODE simulations above ($0.2*Xist^{high}*p_{21}$). Based on this classification the mean fraction of cells exhibiting mono- and bi-allelic expression was calculated. A small percentage of parameter sets tested (0.71%) could recapitulate efficient mono-allelic *Xist* up-regulation (>80% cells in the Xist+Xist- state during the last 20h of the simulation) as shown in Figure A2.

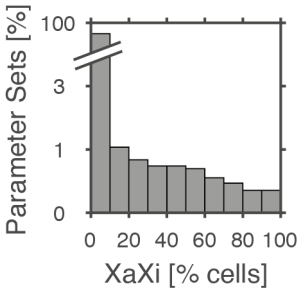


Figure A2. Distribution of the percentage of cells that reproduce mono-allelic *Xist* up-regulation for all simulated parameter sets

2.3 Steady state analysis

To understand the prerequisites for mono-allelic *Xist* up-regulation in the cXR-tXA model we analyzed the steady states of the system, both locally at the allele level and globally at the cell level. The steady state levels of *Xist* for an individual allele were identified through an ODE simulation (see section 1) starting from different initial conditions (x=0, 0.1, 0.2...1, cXR=1, 0.9, 0.8,...0) and for different tXA doses, which remained constant during the simulation. The steady states reached after 100h of simulation are shown in Fig. 2e (top) in the main text for an example parameter set (p_3=3.6, p_4=0.20, p_5=1.5, p_6=0.018, p_{11}=3.4, p_{12}=0.80, p_{13}=4.9, p_{14}=0.37). tXA doses corresponding to 0, 1 or 2 active X-chromosomes (XiXi, XaXi, XaXa) are indicated, as calculated from the $Xist^{high}$ and $Xist^{low}$ steady state as follows.

$$tXA(Xi) = 1 - \frac{(Xist^{high})^{p3}}{(Xist^{high})^{p3} + p4^{p3}} \qquad tXA(Xa) = 1 - \frac{(Xist^{low})^{p3}}{(Xist^{low})^{p3} + p4^{p3}}$$

To identify the global steady states of an entire cell with two X chromosomes, the cXR-tXA ODE model was simulated for all combinations of initial values for x_1=0, 0.1, 0.2, ...1 and x_2=0, 0.1, 0.2, ...1. Initial values for cXR_1, cXR_2, tXA_1 and tXA_2 were set to the steady state associated with x_1 or x_2 as calculated from equation (2). The steady states reached after 100h of simulation for the example parameter set given above are shown in Fig. 2e (bottom) in the main text. Steady states that can only be reached from a symmetric initial condition (x1=x2) are indicated as unstable (open circles).

To investigate the roles played by the two regulatory modules mediated by tXA and cXR, the effect of perturbing each module on the system's steady states was analyzed (Fig.2f-g in the main text). To perturb the positive feedback the threshold level of cXR required to repress *Xist* was set to a high value such that cXR is not any more involved in *Xist* regulation (p_{14}=1000). To perturb the tXA modules, tXA was set to a constant value corresponding to a single dose present in the XaXi state as described above.

Surprisingly, we found a subset of parameter sets that could reproduce mono-allelic *Xist* up-regulation, where the system exhibited local and global bistability also in the presence of a double tXA dose. In these cases, the low steady state of *Xist* was rather unstable such that small fluctuations would allow a transition to the high state.

2.4 Parameter rules

To understand, which parameter sets could reproduce mono-allelic *Xist* up-regulation, we analyzed the distribution of the parameter values among parameter sets that could maintain the correct *Xist* expression state in section 1 and that could simulate mono-allelic up-regulation in section 2.2 (Fig. A3).

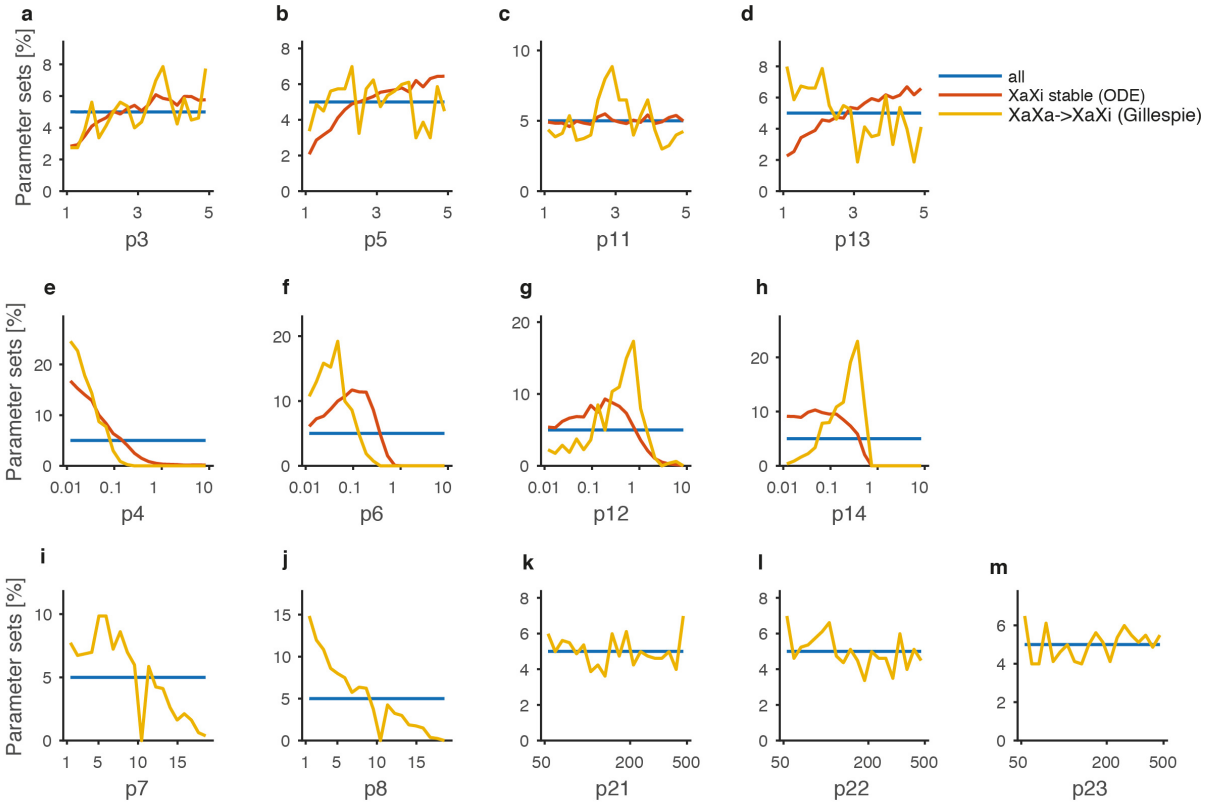


Fig. A3. Distributions of parameter values across all tested parameter sets (blue), across all parameter sets that could maintain the XaXi state in the ODE simulation (red) and across all parameter sets that could reproduce mono-allelic *Xist* up-regulation (yellow). Silencing delays (p_7 and p_8) and scaling factors (p_{21}, p_{22}, p_{23}) are only present in the stochastic model, but not in the ODE model.

For the Hill coefficients p_3, p_5, p_{11}, p_{13} no strong trends are observed (Fig. A3 a-d). The parameters p_4 and p_6 represent the silencing threshold for tXA and cXR, respectively. They are equal to the *Xist* level where tXA or cXR will be reduced to half of their maximal levels. Their values must be rather low to ensure efficient silencing upon mono-allelic *Xist* up-regulation (Fig. A3 e+f). The parameters p_{12} and p_{14} describe how tXA and cXR, respectively, control *Xist*. They represent the tXA and cXR levels where *Xist* activation

and repression will be half-maximal. Both are required at intermediate levels for mono-allelic *Xist* up-regulation to allow *Xist* to respond sensitively to changes in tXA and cXR (Fig. A3 g+h). The effect of p_{12} and p_{14} are discussed in more detail below. The silencing delays p_7 and p_8 should be short (Fig. A3 i+j), such that the $Xist^{high}$ state can be locked in efficiently through cXR silencing and that the system can move to the bi-stable regime through tXA silencing. For the scaling factors p_{21}, p_{22} and p_{23} that determine the absolute levels of Xist, cXR and tXA, respectively, all values tested between 50 and 500 are compatible with Xist up-regulation (Fig. A3 k-m).

Both p_{12} and p_{14} must assume intermediate levels to allow *Xist* up-regulation (Fig. A3 g+h). In particular p_{14} must lie in a precise parameter range. To better understand these parameter requirements, we analyzed how these parameters affect the steady states of the system by visualizing the steady states in the Xist-cXR phase space. To this end, equations (6) and (8) were solved for steady state conditions ($\frac{dy}{dt} = 0$) to

$$\overline{x_1} = f(tXA)\left(1 - \frac{cXR_1^{p_{13}}}{cXR_1^{p_{13}} + p_{14}{}^{p_{13}}}\right)$$

$$\overline{cXR_1} = 1 - \frac{x_1^{p_5}}{x_1^{p_5} + p_6{}^{p_5}}$$

For one example parameter set that can recapitulate the XaXa-XaXi transition, we plot Xist (x1) vs cXR for both equations for 3 different doses of tXA, which would correspond to the XaXa (2x), XaXi (1x) and XiXi (0x) states (Fig. A4).

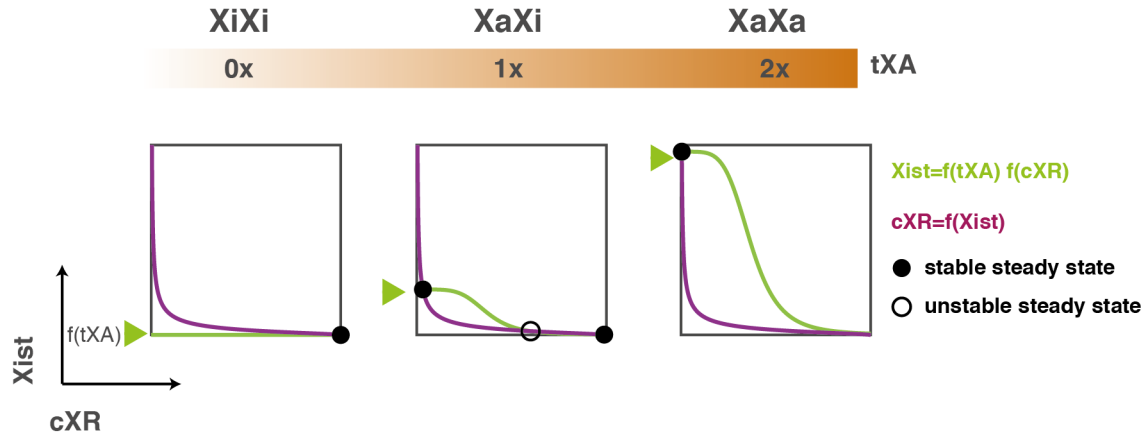


Fig. A4. Xist-cXR phase space for one example parameter set for different tXA doses as indicated on top.

The two curves intersect at the system's steady states. At the mono-allelic XaXi state (middle) the system exhibits bistability at the allele level. Because both a low and a high steady state exist for *Xist*, the two chromosomes can be stably maintained in two alternative states. In a putative XiXi state (left), where tXA is silenced on both chromosomes only the low state is stable. As a consequence *Xist* would be down-regulated if both chromosomes silence tXA, such that an XiXi state cannot be maintained

in the system. Prior to the onset of X inactivation, both copies of tXA are active (XaXa) and only a single high steady state for *Xist* is present (right). This disappearance of the $Xist^{low}$ state is the prerequisite for *Xist* up-regulation in the presence of a double tXA dose. In male cells that have only a single tXA dose (middle) the low state exists, thus preventing *Xist* up-regulation.

To understand how p_{12} and p_{14} would affect the system's steady states, we plotted the phase diagram again for a two-fold increase or decrease in p_{12} or p_{14}, respectively (Fig. A5)

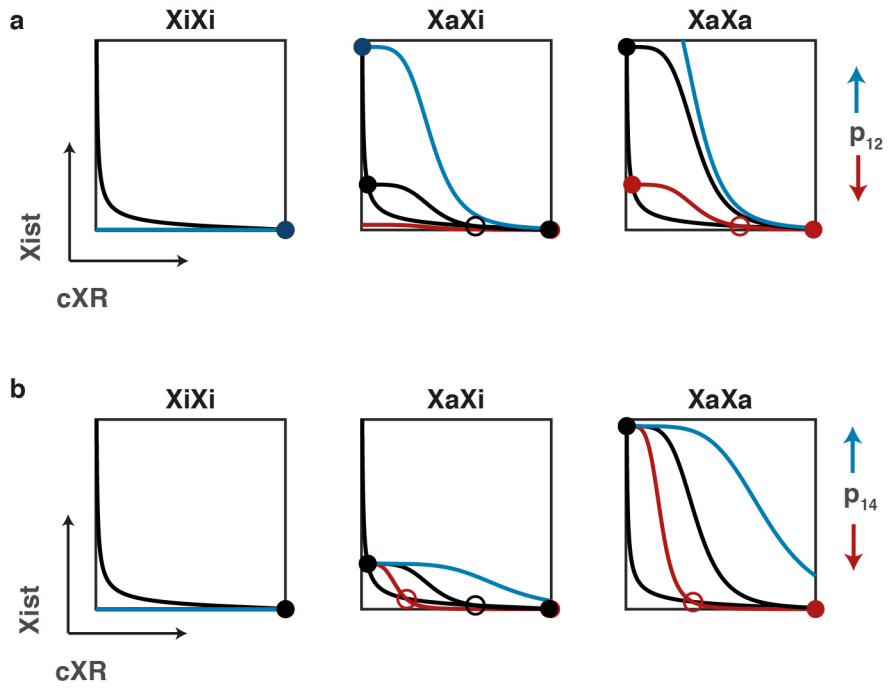


Fig. A5. Xist-cXR phase space in response to changes in p_{12} (a) and p_{14} (b) for the same parameter set shown in Fig. A4. Blue indicates a two-fold increase of the respective parameter value and red a two-fold decrease.

An increase of p_{12} will result in disappearance of the $Xist^{low}$ state in the XaXi scenario (Fig. A5a, middle, blue), while a decrease will result in loss of the $Xist^{high}$ state (red). Therefore, p_{12} is required at intermediate levels to maintain and establish the mono-allelic state (cp. Fig. A3g). For p_{14}, an increase will lead to loss of the $Xist^{low}$ state (Fig. A5b, middle, blue). As a consequence, p_{14} must be below a certain level to allow bistability and thus maintenance of the XaXi state (cp. Fig. A3h, red). A decrease of p_{14} by contrast results in the appearance of an $Xist^{low}$ state in the XaXa scenario (Fig. A5b right, red), thus preventing *Xist* up-regulation in the pre-XCI state. Therefore, mono-allelic *Xist* up-regulation requires p_{14} to lie in a precise window (Fig. A3h).

2.5 Transient bi-allelic *Xist* up-regulation

To investigate, whether the cXR-tXA model could reproduce transient bi-allelic *Xist* up-regulation as observed in rabbit embryos, we quantified the frequency of bi-allelic *Xist* expression across time for each parameter set that reproduced mono-allelic *Xist* up-

regulation (in the last 20h of the simulation) in >80% cells in the simulations in section 2.2. The maximal level of bi-allelic expression observed was highly variable across parameter sets (Fig. A6).

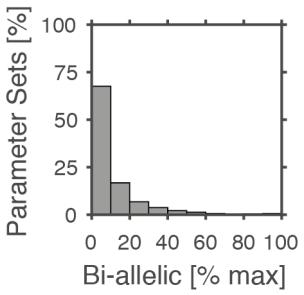


Figure A6. Distribution of the maximal percentage of cells that exhibit bi-allelic *Xist* expression during the simulation, for all parameter sets, where *Xist* is up-regulated mono-allelically in >80% of cells.

The level of transient bi-allelic *Xist* expression in the simulation depended on how fast *Xist* was up-regulated and how fast tXA was silenced. To quantify the speed of *Xist* up-regulation, we defined the "Switch-ON time" as the time point when one X chromosome had reached the Xist+ state (as defined in section 2.2). If switch-ON did not occur before the end of the simulation it was set to the total simulation time (100h). The ratio of the average switch-ON time and the tXA silencing delay (p_8) is inversely correlated with the maximal level of bi-allelic expression observed (Fig. S2a).

2.6 Reproducing experimental measurements of *Xist* up-regulation

To investigate whether the model could also reproduce quantitative *Xist* expression data we compared the simulations to RNA FISH data from differentiating female mESCs and to *in vivo* data in mouse and rabbit[4,5]. To temporally align simulation and experiment, different values for the time point when random XCI initiates were tested (offset) within a 24h time window before the first time point when Xist clouds were observed. To identify the parameter set - offset combination that explains the data best we used a maximum likelihood estimate (MLE) approach. Both data and simulations were modeled as a multinomial distribution (mnpdf). The log Likelihood was then calculated as follows

$$MLE = \sum_t log \sum_i mnpdf(D_t|p_i) * mnpdf(S_t|p_i) \qquad with\ i = 0, 0.01, \ldots, 1$$

where t represents the measured time points and D_t and S_t are vectors containing the number of cells with no, mono-allelic and bi-allelic Xist expression in the data and the simulation, respectively. The parameter - offset combinations that maximized the MLE were selected. A summary of the data sets used and the window in which the offset was tested as well as the selected offset are given in table A4. The best parameter sets that were used in Fig. 3d and S1a are given in table A5.

Data set	Time points	Offset window tested	Offset selected
mESCs	0, 1, 2, 3, 4 days	0...24h	9 h
mouse embryos	E5.0, E5.5[4], E6.5[4], E7.5[4]	96 ... 120 h	106 h (~E4.4)
rabbit embryos	Morula (67), 96, 120 h.p.c.[5]	43...67 h	57 h

Table A4. Summary of the data sets that were compared to model simulation.

Data set	p_3	p_4	p_5	p_6	p_7	p_8	p_{11}	p_{12}	p_{13}	p_{14}	p_{21}	p_{22}	p_{23}
mESC	1.5	0.023	3.7	0.026	1	16	2.5	0.64	1.9	0.10	75	167	151
mouse	3.4	0.017	2.2	0.019	1	5	2.7	1.03	2.6	0.20	347	76	79
rabbit	2.2	0.048	3.0	0.047	6	12	1.7	1.4	2.3	0.35	451	207	220

Table A5. Parameter sets that best explain the experimental data.

The 100 parameter sets that could best explain the mouse *in vivo* data were used to simulate aneuploid and polyploid cells in section 2.7, while the 100 parameter sets that best fit the mESC data were used for the simulations in section 2.9.

2.7 Simulating aneuploid and polyploid cells

To simulate cells that are mono-, tri- or tetrasomic for the X chromosome, *Xist* up-regulation was simulated as described in section 2.2 except that each cell contained one, three or four X chromosomes, each contributing a single tXA dose. For the 100 parameter sets that could best explain the mouse *in vivo* data (see section 2.6) 100 cells were simulated. Each cell was classified as no Xi, Xi, XiXi, XiXiXi or XiXiXiXi depending on how many Xist+ alleles were present during the last 20 h of the simulation (Fig. 2h in the main text).

To simulate triploid (3n3X) and tetraploid cells (4n4X) we assumed either that tXA is repressed by autosomal transcription factors (Fig. 2i in the main text) or that an increase in nuclear size would result in an effective dilution of tXA compared to diploid cells (Fig. 2j). To simulate autosomal tXA repression, equation (7) was substituted by

$$\frac{dtXA}{dt} = p_{23} \cdot \frac{2}{p_{33}} \left[1 - \frac{x^{\,p_3}}{x^{\,p_3} + (p_{21} \cdot p_4)^{p_3}} \right] - tXA \qquad (9)$$

where p_{33} indicates the autosomal ploidy (3 for triploid cells, 4 for tetraploid cells). To simulate tXA dilution, equation (6) was substituted by

$$\frac{dx}{dt} = p_{21} \cdot f(cXR^{\,}) \cdot \frac{\left(\frac{1}{p_{33}} (\sum_{i=1}^{n} tXA_i) \right)^{p_{11}}}{\left(\frac{1}{p_{33}} (\sum_{i=1}^{n} tXA_i) \right)^{p_{11}} + (p_{23} \cdot p_{12})^{p_{11}}} - x \qquad with\ n = \#\ X\ chromosomes \quad (10)$$

The four X chromosomes in a tetraploid cell and the three X chromosomes in a triploid cell would therefore produce the same effective tXA concentration as the two X's in a diploid cell. The simulations were performed as described for polysomic cells above.

2.8 Simulating *Xist* up-regulation in human embryos

To understand whether cXR-tXA model could also explain extended bi-allelic *Xist* expression in human embryos, we tested whether and under which conditions absent or reduced silencing ability of *Xist* would lead to bi-allelic *Xist* expression. We identified two scenarios that could reproduce sustained bi-allelic expression: (1) cXR up-regulation or activation and (2) cXR dampening. To simulate these scenarios equations (7) and (8) were modified as follows and 100 parameter sets that could reproduce mono-allelic *Xist* expression were simulated as described in section 2.2.

	(1) cXR upregulation/activation	**(2) cXR dampening**
tXA	$for\ t \leq t1$: $\frac{dtXA_i}{dt} = p_{23} - tXA_i$	$for\ t \leq t1$: $\frac{dtXA_i}{dt} = p_{23} - tXA_i$
	$for\ t > t1$: $\frac{dtXA_i}{dt} = p_{23}\left[1 - \frac{{x_i}^{p_3}}{{x_i}^{p_3} + (p_{21} \cdot p_4)^{p_3}}\right] - tXA_i$	$for\ t > t1$: $\frac{dtXA_i}{dt} = p_{23}\left[1 - \frac{{x_i}^{p_3}}{{x_i}^{p_3} + (p_{21} \cdot p_4)^{p_3}}\right] - tXA_i$
cXR	$for\ t < t1$: $\frac{dcXR_i}{dt} = 0 - cXR_i$	$for\ t < t1$ $\frac{dcXR_i}{dt} = p_{22}\left[1 - p_{34} * \frac{{x_i}^{p_5}}{{x_i}^{p_5} + (p_{21} \cdot p_6)^{p_5}}\right] - cXR_i$
	$for\ t \geq t1$: $\frac{dcXR_i}{dt} = p_{22}\left[1 - \frac{{x_i}^{p_5}}{{x_i}^{p_5} + (p_{21} \cdot p_6)^{p_5}}\right] - cXR_i$	$for\ t \geq t1$: $\frac{dcXR_i}{dt} = p_{22}\left[1 - \frac{{x_i}^{p_5}}{{x_i}^{p_5} + (p_{21} \cdot p_6)^{p_5}}\right] - cXR_i$

In scenario 1 it was assumed that cXR was not expressed before t1 (=day 4) and that tXA could not be silenced by *Xist* during that period. In scenario 2, we assumed that cXR is subject to dampening by *Xist* RNA while tXA is unaffected by *Xist*. To simulate dampening of cXR by *Xist* we introduced an additional parameter p_{34} that describes maximal reduction of cXR by *Xist*. For each parameter set, p_{34} was randomly drawn from a uniform distribution between 0.01 and 0.99. Both scenarios could reproduce extended bi-allelic expression and the transition to the mono-allelic state (Fig. 3e in the main text and supplemental Fig. S2b).

2.9 Simulating experimental modulation of bi-allelic *Xist* up-regulation

Our model analysis showed that a single network architecture can generate different degrees of transient biallelic *Xist* expression. The decision as to whether a cell will undergo direct mono-allelic up-regulation or follow a route via transient bi-allelic expression depends on which reaction occurs first: Silencing or up-regulation of *Xist* from the second allele. If one could artificially accelerate *Xist* up-regulation from one allele, this would prolong the time before *Xist* is up-regulated from the other allele. As a result, the switch-ON-to-silencing ratio would be increased and the extent of transient bi-allelic expression reduced (cp. Fig S2a).
To simulate induction of *Xist* with doxycycline equation (6) was modified by introducing two additional parameters per chromosome as follows.

	Chr. X_1 (B6)	**Chr. X_2 (Cast)**
Xist	$\frac{dx_1}{dt} = p_{21} \cdot (p_{24} + p_{29} \cdot f(cXR_1) \cdot f(tXA)) - x_1$	$\frac{dx_2}{dt} = p_{21} \cdot (p_{30} \cdot f(cXR_2) \cdot f(tXA)) - x_2$

P_{24} describes *Xist* regulation by doxycycline on the B6 chromosome and is set to 10 if Doxycycline is present or to 0 if doxycycline is absent. P_{29} and p_{30} control regulation of *Xist* by tXA and cXR and are set to 1 on an allele that is not induced with doxycycline but to 0 if induced with doxycycline, since *Xist* expression then becomes independent of regulation by tXA and cXR. Before the start of differentiation p_{29} and p_{30} are also set to 0 representing the action of stem cell specific factors that prevent up-regulation of *Xist* in undifferentiated cells by repressing *Xist*. Parameter settings to simulate doxycycline treatment are shown below.

	+ Dox		**Control**	
	$t < 0$	$t \geq 0$	$t < 0$	$t \geq 0$
p_{24}	10	10	0	0
p_{29}	0	0	0	1
p_{30}	0	1	0	1

To simulate skewing in the TX1072dT line due to a heterozygous *Tsix* deletion on the Cast chromosome p_{29} was set to 33% of p_{29}, a summary of the parameter setting is shown below.

	$t < 48h$	$t \geq 48h$
p_{24}	0	10
p_{29}	0.33	0
p_{30}	1	1

The simulation was performed for the 100 parameter sets that best described the mESC data in section 2.6.

3 Supplementary Note 3: The *Tsix-Xist* model

3.1 Model description

Reactions describing transcription initiation, transcription elongation and RNA degradation of *Xist* and *Tsix* were combined into a mathematical model (Fig. A7).

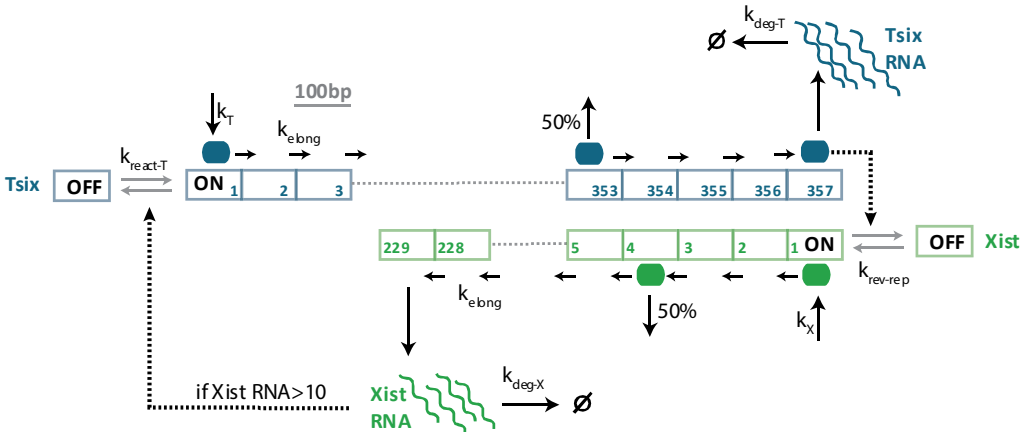


Fig. A7. Schematic representation of antisense simulation.

Both promoters were assumed to exist in an 'OFF' state, where no transcription occurs, and an ‚ON' state, where transcription is initiated with constant transcription rates k_X and k_T. The *Tsix* promoter is turned off by *Xist* RNA-mediated silencing, the *Xist* promoter is turned off by passing *Tsix* polymerases. The OFF state is reverted back to the ON state with rate $k_{rev\text{-}rep}$. To describe transcriptional elongation, the *Xist* and *Tsix* gene bodies were divided into segments of 100nt and RNA polymerases move along the gene body with a constant rate (k_{elong}). Fully elongated transcripts produce one RNA molecule. Degradation of the RNA obeys first-order reaction kinetics with the rates $k_{deg\text{-}X}$ and $k_{deg\text{-}T}$. Transcription of *Xist* and *Tsix* mutually interfere by the following mechanisms:

(i) Transcriptional interference (TI): TI was essentially modeled as in Sneppen et al.[6] : Passing polymerases prevent binding of a polymerase to the promoter (occlusion) while collisions between sense- and antisense elongating polymerases result in dislodgement from the gene. While Sneppen et al. assumed that both polymerases will dislodge upon collision, here one of the two polymerases is randomly chosen and removed from the gene. Additionally, in our model promoter bound and elongating polymerases have the same chance of being dislodged upon collisions.

(ii) *Tsix*-dependent repression of the *Xist* promoter: *Tsix* polymerases transcribing through the *Xist* promoter region induce a transition to the OFF state of the *Xist* promoter that can be reverted to the ON state with $k_{rev\text{-}rep}$ (the half-life of the repressed state is given by $t_{1/2}^{repr} = \frac{ln2}{k_{rev-rep}}$).

(iii) Silencing of the *Tsix* promoter by *Xist* RNA: If *Xist* RNA is present above a threshold level of 10 RNA molecules it induces a transition of the *Tsix* promoter to the OFF state.

To account for regulation of Xist by tXA, the tXA dosage was assumed to modulate the effective *Xist* initiation rate k_X^{eff} as follows:

$$k_X^{eff} = q_{tXA} \cdot k_X$$

where k_X is the *Xist* initiation rate in the presence of a single tXA dose and q_{tXA}=0,1,2 depending on whether no, one or two tXA loci are active. The tXA concentration was modeled as a step function with the value 1 if the respective tXA allele is active and the value 0 if the respective tXA allele has been silenced by *Xist* RNA. The kinetics of RNA and protein decay of tXA were not accounted for explicitly but were instead assumed to modulate the tXA silencing kinetics (see below).

To reproduce coupling of XCI to development, k_X^{eff} must be influenced by the differentiation timing, representing the action of stem cell specific factors that prevent *Xist* up-regulation in undifferentiated cells by repressing *Xist*. The differentiation dependency was formulated as a step function that changes its value at the point of induction of differentiation such that k_X^{eff} prior to differentiation was 10% of the k_X value after the onset of differentiation.

$$\text{Before differentiation:} \quad k_X^{eff} = 0.1 \cdot q_{tXA} \cdot k_X$$

$$\text{After onset of differentiation:} \quad k_X^{eff} = q_{tXA} \cdot k_X$$

Since *Xist*-dependent silencing is known to occur with a delay of hours or days after *Xist* has been up-regulated[7,8], we implemented a silencing delay described by the parameters sil_{Tsix} or sil_{tXA}. To this end, each chromosome passes stochastically through a number of intermediate states S1, S2...Sn once *Xist* expression from that chromosome has exceeded a certain threshold (10 molecules) and gene silencing occurs once the final silencing state Sn has been reached. If the level of *Xist* RNA molecules drops below the threshold before Sn has been reached, the chromosome immediately passes back to the unsilenced state S0. The transitions through the intermediate states occur with rate $1h^{-1}$ such that the number of intermediate states given by the model parameters sil_{Tsix} or sil_{tXA} is equal to the mean silencing delay. Silencing is assumed to be reversed, if the *Xist* level drops below the threshold of 10 molecules. Reactivation of *Tsix* and tXA will then occur with a single stochastic reaction with the rates $k_{react\text{-}T}$ and $k_{react\text{-}tXA}$ respectively (Fig. A8).

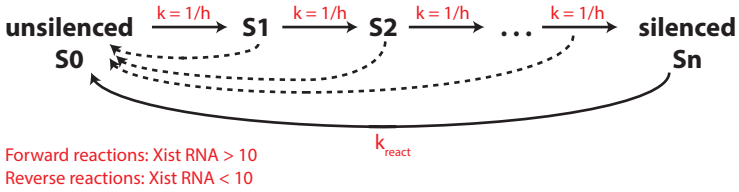


Fig. A8. Schematic representation silencing and reactivation in the antisense model.

3.2 Simulating maintenance of the XaXi state

To find parameter values that could reproduce mono-allelic *Xist* up-regulation, we first scanned a large parameter space for sets that would maintain the XaXi state. Subsequently we tested, whether those sets can reproduce the transition from the XaXa to the XaXi state. Degradation and elongation rates were set to fixed values based on previous experimental estimates (Table A6). For the *Xist* half-life previous studies have attempted an experimental estimation, resulting in values of 2 and 6h respectively[1,2]. We therefore use the mean (4h), which results in a degradation rate of 0.1733 h^{-1} ($\ln(2)/t_{1/2}$). Since silencing has already occurred in the XaXi state and silencing kinetics should therefore not affect the outcome of the simulation, silencing and reactivation rates were set to constant values ($1h^{-1}$). Similarly, tXA is present at a constant single dose in post-XCI cells with a single Xa and is therefore set to a constant value of 1. All other parameters were varied within realistic parameter ranges and systematically combined resulting in 8000 parameter sets in total (Table A6).

Description	Parameter	Parameter value(s)
Xist transcription rate [h^{-1}]	k_X	5, 6.35, 8.1, 10.35, 13.2, 16.8, 21.4, 27.3, 34.75, 44.3, 56.45, 71.9, 91.65, 116.8, 148.8, 189.65, 241.65, 307.95, 392.4, 500
Tsix transcription rate [h^{-1}]	k_T	as k_X
Xist degradation rate [h^{-1}]	k_{deg-X}	0.1733 [1,2]
Tsix degradation rate [h^{-1}]	k_{deg-T}	1.3868 [1]
Elongation rate [bp/sec]	k_{elong}	40 [9]
Reversal rate of *Xist* promoter repression [h^{-1}]	$k_{rev-rep}$	0.1, 0.1438, 0.2069, 0.2976, 0.4281, 0.6159, 0.8859, 1.2743, 1.8330, 2.6367, 3.7927, 5.4556, 7.8476, 11.288, 16.238, 23.357, 33.598, 48.329, 69.519, 100

Table A6. Parameter values in the antisense model

To investigate Xa and Xi stability, each state was simulated from initial conditions where RNA levels of transcribed genes were set to their maximal steady state value and the polymerase complexes were randomly distributed along the gene body (Table A7).

	Xi	**Xa**
***Xist* RNA**	$k_X/k_{deg\text{-}X}$	0
***Tsix* RNA**	0	$k_T/k_{deg\text{-}T}$
***Xist* promoter**	ON	OFF
***Tsix* promoter**	OFF	ON
# *Xist* polymerases	$\frac{L}{k_{elong}} * k_X$ (L=22 900 bp)	0
# *Tsix* polymerases	0	$\frac{L}{k_{elong}} * k_T$ (L=35 700 bp)

Table A7. Initial conditions on Xa and Xi in the antisense model

In the simulation, transcription elongation occurs at fixed time intervals of 2.5 seconds inferred from measurements of polymerase speed (elongation of one 100bp interval at k_{elong}=40bp/sec). Between elongation steps, all other reactions are simulated using the stochastic Gillespie algorithm[3]. For each parameter set 100 Xi/Xa pairs were simulated for 500h to reach the steady state.

A simulation was classified as stably maintaining the XaXi state if *Xist* was on average present with >10 molecules at the Xi and with <10 molecules at the Xa during the last 50h of the simulation. Parameter sets, where >99% of Xa/Xi pairs stably maintained the XaXi state were classified as mono-allelic.

3.3 Simulating mono-allelic *Xist* up-regulation

To simulate the onset of X inactivation, both chromosomes were initiated from the Xa state (see Table A7 in 3.2) in undifferentiated cells with double tXA dosage present (q_{tXA}=2). *Xist* up-regulation was simulated for all parameter sets that could stably maintain the XaXi state in the previous section (4001 sets). Each parameter set was combined with 500 combinations of randomly sampled values for sil_{tXA}, sil_{Tsix}, $k_{react\text{-}tXA}$ and $k_{react\text{-}T}$ (see following table).

Description	**Parameter**	**Parameter values**
Silencing delay of tXA [h]	sil_{tXA}	1 – 48 (log distributed)
Silencing delay of *Tsix* [h]	sil_{Tsix}	1 – 48 (log distributed)
Reactivation rate of tXA [h^{-1}]	$k_{react\text{-}tXA}$	0.1-100 (log distributed)
Reactivation rate of *Tsix* [h^{-1}]	$k_{react\text{-}T}$	0.1-100 (log distributed)

For each parameter set 100 cells were simulated. To reach the steady state prior to differentiation, the cells were simulated for 10h in an undifferentiated state, then 100 hours of differentiation were simulated as this is the relevant time scale of XCI. Each cell was classified as mono-allelic, if during the last 20h of the simulation >10 molecules of *Xist* RNA were present on average at one chromosome (Xi) and <10 molecules on the other (Xa). 1.17% of the tested parameter sets up-regulated *Xist* mono-allelically in

>99% cells and were thus classified as mono-allelic (XaXa->XaXi in Fig. 5d in the main text).

3.4 Model simplification

To analyze which of the repressive mechanisms are required for establishment and maintenance of the XaXi state, we systematically investigated reduced model structures with a combination of two or only a single repressive mechanism. These simplifications were implemented as follows.

- In all models without *Xist* promoter repression, passing *Tsix* polymerases do not affect the *Xist* promoter state (parameter $k_{rev\text{-}rep}$ removed).
- In all models without *Tsix* promoter silencing, the *Xist* RNA does not affect the *Tsix* promoter state.
- In all models without transcriptional interference (TI), *Xist* and *Tsix* polymerases were assumed to be able to bypass each other.

First, maintenance of the XaXi state was assessed for all six reduced models as described in section 3.2. For all parameter sets that could maintain the XaXi state, *Xist* up-regulation was simulated as described in section 3.3. A summary of the simulations of the full model and of all reduced models is given in table A8. Only the reduced model, where mutual inhibition of *Xist* and *Tsix* was mediated by silencing of the *Tsix* promoter and by TI could reproduce mono-allelic *Xist* up-regulation. This model was termed the "antisense model" and used for all subsequent simulations.

Model	**Parameters varied**	**XaXi -> XaXi**		**XaXa->XaXi**	
		# Parameter sets tested	**XaXi stable**	**# Parameter sets tested**	**XaXi established**
(1) *Tsix* silencing (2) *Xist* repression (3) TI	k_X, k_T, $k_{rev\text{-}rep}$	8000	50%	2 000 500	1.17%
(2) *Xist* repression (3) TI	k_X, k_T, $k_{rev\text{-}rep}$	8000	0%	-	-
(1) *Tsix* silencing (3) TI	k_X, k_T	400	45%	90 000	1.53%
(1) *Tsix* silencing (2) *Xist* repression	k_X, k_T, $k_{rev\text{-}rep}$	8000	26.4%	1 057 500	0%
(1) *Tsix* silencing	k_X, k_T	400	0%	-	-
(2) *Xist* repression	k_X, k_T, $k_{rev\text{-}rep}$	8000	0%	-	-
(3) TI	k_X, k_T	400	0%	-	-

Table A8. Summary of model reduction

3.5 Reduced overlap between *Xist* and *Tsix*

To investigate whether the antisense model could in principle still ensure stable XaXi maintenance and robust mono-allelic *Xist* up-regulation if the overlap between *Xist* and *Tsix* was reduced, we repeated the analysis described in sections 3.2 and 3.3 for the human *XIST/TSIX* locus architecture with 8kb overlap (Fig. S6). First, maintenance of the XaXi state was assessed using the same parameter sets as in the antisense model with the mouse overlap (section 3.2). For all parameters that could maintain the XaXi state, *Xist* upregulation was simulated as described in section 3.3. A simulation of *Xist* up-regulation for one example parameter set is shown in Fig S6e-f. A summary of the simulations is given in table A9 and in Fig S6d.

Model	**parameters varied**	**XaXi -> XaXi**		**XaXa->XaXi**	
		# parameter sets tested	**XaXi stable**	**# parameter sets tested**	**XaXi established**
23kb overlap	k_X, k_T	400	45%	90 000	1.53%
8kb overlap	k_X, k_T	400	38%	76 000	0.85%

Table A9. Summary of simulations with mouse and human locus architecture

3.6 Simulations of *Xist* and *Tsix* mutant cell lines

To simulate experimental data we selected a subset of parameter sets from the simulation in 3.4 that robustly led to mono-allelic *Xist* up-regulation (>99% cells) and were in agreement with experimental observations according to the following constraints:

(i) The maximal percentage of bi-allelically expressing cells over the simulated time course should be below 20%

(ii) The mean *Xist* expression level must lie between 200 and 600 RNA molecules[1].

(iii) All cells up-regulate *Xist* within 48h after induction of differentiation.

Since only 34 parameter sets with unique k_X, k_T, sil_{tXA} and sil_{Tsix} combinations fulfilled these criteria, we performed another simulation to identify more parameter sets that could potentially simulate experimental data. To this end, the simulation in 3.4 was repeated with additional, randomly sampled values for k_X and k_T. The parameter ranges for k_X were set such that the steady state expression level of *Xist* ($k_X/k_{X\text{-}deg}$) was between 200 and 600 molecules. Since an analysis of the parameter sets identified in section 3.4 had revealed that mono-allelic *Xist* up-regulation requires a k_X-to-k_T ratio between 0.4 and 0.8, k_T was sampled within this range. A total of 500,000 parameter sets were simulated. All parameters were sampled randomly within the ranges given in the following table.

Description	Parameter	Parameter values
Xist initiation rate [h^{-1}]	k_X	34-104 (log distributed)
Tsix initiation rate [h^{-1}]	k_T	$k_X/0.8$-$k_X/0.4$ (lin distributed)
Silencing delay of tXA [h]	sil_{tXA}	0 – 48 (log distributed)
Silencing delay of *Tsix* [h]	sil_{Tsix}	0 – 48 (log distributed)
Reactivation rate of tXA [h^{-1}]	$k_{react\text{-}tXA}$	0.1-100 (log distributed)
Reactivation rate of *Tsix* [h^{-1}]	$k_{react\text{-}T}$	0.1-100 (log distributed)

From these simulations 100 sets fulfilling the above requirements were randomly selected and were used in the following simulations. *Xist* and *Tsix* mutations were simulated as described above with the following modifications.

(i) *Tsix*$^{+/-}$: *Tsix* initiation rate k_T=0 on one allele

(ii) *Tsix*$^{-/-}$: *Tsix* initiation rate k_T=0 on both alleles.

(iii) *Xist*$^{+/-}$: *Xist* initiation rate k_X=0 on one allele

For each of the selected parameter sets 100 cells were simulated for each mutant. Fig. 7b-d in the main text shows the trajectory of one representative cell (top), the final expression state of all cells for one parameter set (middle) and the distributions of expression patterns observed across all tested parameter sets (bottom). To estimate the half time of *Xist* up-regulation $T_{1/2}$ shown in Fig. 7e (main text) we determined the earliest time point where 50% of simulated cells had up-regulated *Xist* (>10 molecules).

4 References

1. Sun, B. K., Deaton, A. M. & Lee, J. T. A Transient Heterochromatic State in Xist Preempts X Inactivation Choice without RNA Stabilization. *Molecular Cell* **21,** 617–628 (2006).
2. Yamada, N. *et al.* Xist Exon 7 Contributes to the Stable Localization of Xist RNA on the Inactive X-Chromosome. *PLoS Genet.* **11,** e1005430 (2015).
3. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* **81,** 2340–2361 (1977).
4. Sakata, Y. *et al.* Defects in dosage compensation impact global gene regulation in the mouse trophoblast. *Development* **144,** 2784–2797 (2017).
5. Okamoto, I. *et al.* Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* **472,** 370–374 (2011).
6. Sneppen, K. *et al.* A mathematical model for transcriptional interference by RNA polymerase traffic in Escherichia coli. *J. Mol. Biol.* **346,** 399–409 (2005).
7. Chow, J. C. *et al.* LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. *Cell* **141,** 956–969 (2010).
8. Borensztein, M. *et al.* Xist-dependent imprinted X inactivation and the early developmental consequences of its failure. *Nat Struct Mol Biol* **24,** 226–233 (2017).
9. Jonkers, I., Kwak, H., Lis, J. T. & Struhl, K. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife* **3,** e02407 (2014).