iScience, Volume *24*

Supplemental information

Semicoordinated allelic-bursting shape dynamic random monoallelic expression in pregastrulation embryos

Hemant Chandru Naik, Kishore Hari, Deepshikha Chandel, Susmita Mandal, Mohit Kumar Jolly, and Srimonta Gayen

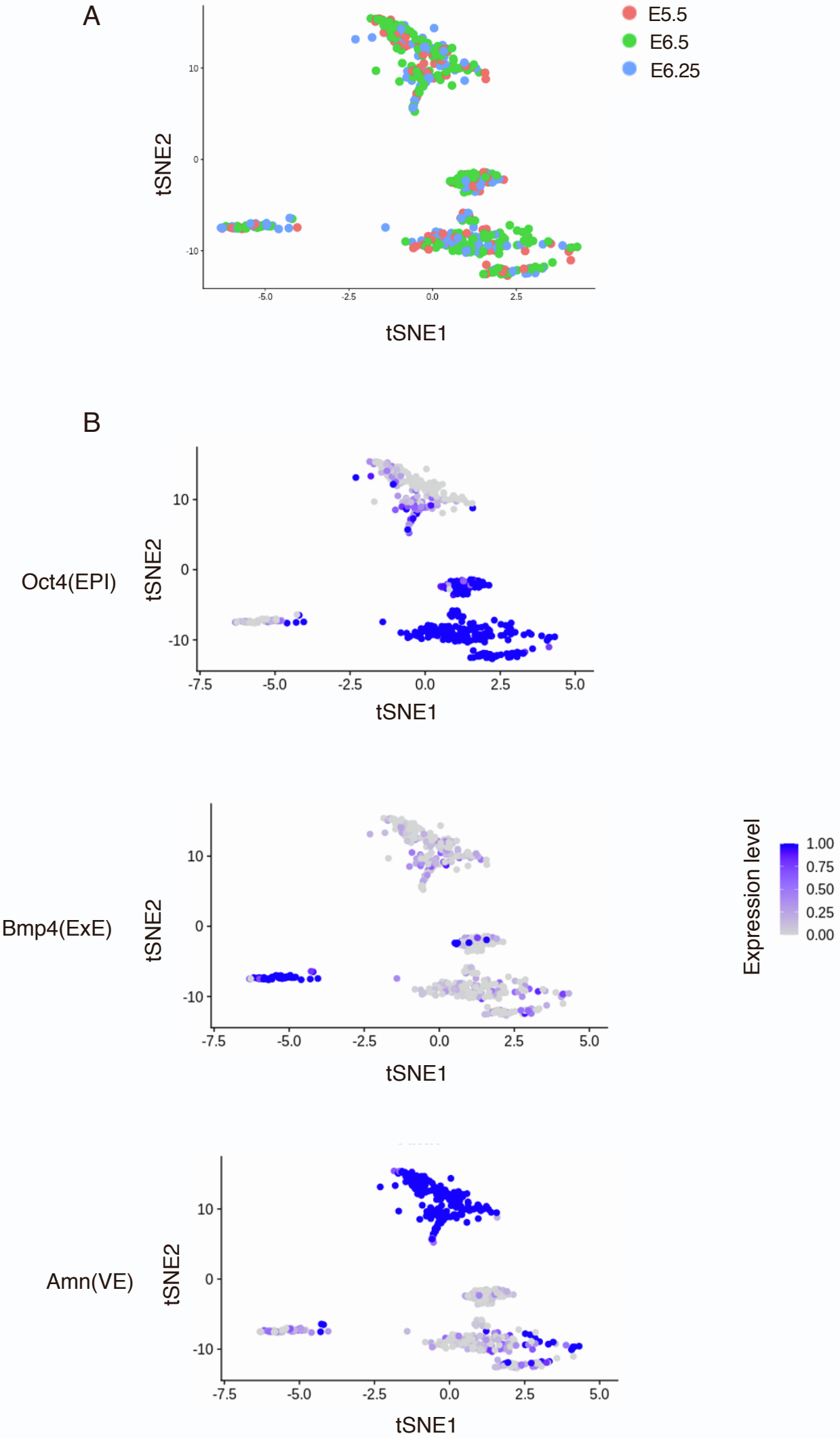
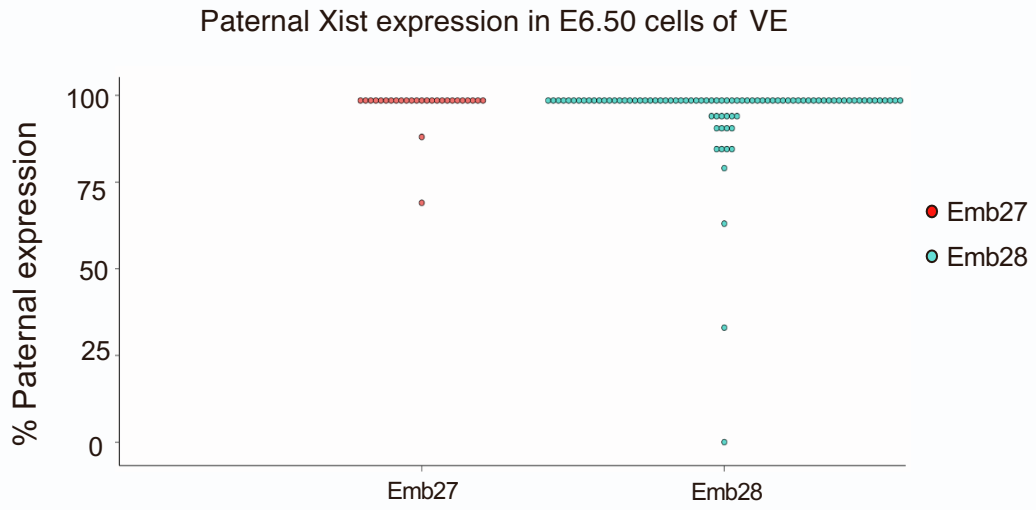

A
E5.5
E6.5
E6.25
tSNE2
tSNE1
B
Oct4(EPI)
Bmp4(ExE)
Amn(VE)
Expression level
1.00
0.75
0.50
0.25
0.00

Figure S1: Lineage profiling of pre-gastrulation mouse embryos based on single cell transcriptomics. Related to Figure 1. (A) Clustering of all cells (n=510) from the three different stages (E5.5, E6.25 and E6.50) into two principal dimensions using t-SNE analysis based on 3000 most variable genes. (B) Representation of lineage specific marker expression of the clustered cells generated in t-SNE plot: Pou5f1 for EPI, Bmp4 for ExE and Amn for VE.

A


Paternal Xist expression in E6.50 cells of VE
% Paternal expression
100
75
50
25
0
Emb27
Emb28
Emb27
Emb28

B

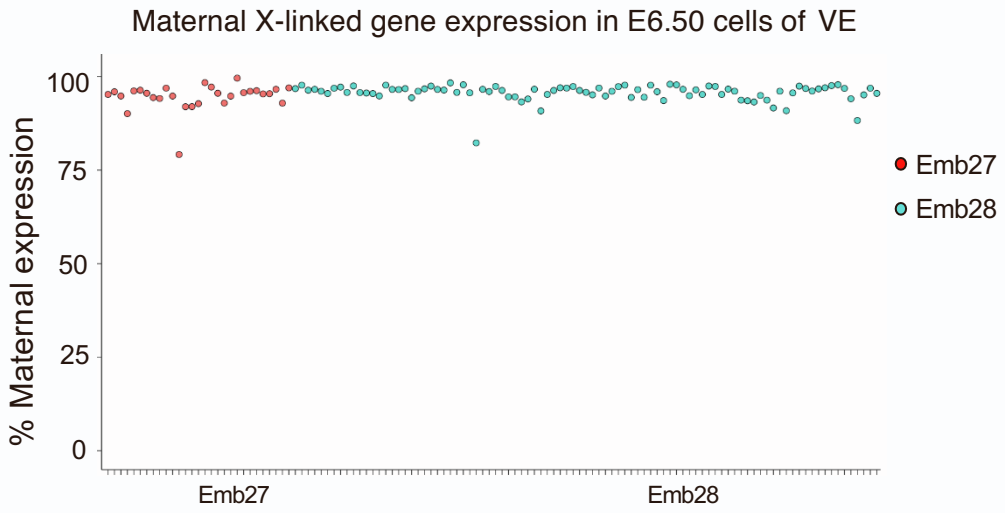

Maternal X-linked gene expression in E6.50 cells of VE
% Maternal expression
100
75
50
25
0
Emb27
Emb28
Emb27
Emb28

Figure S2: Expression of *XIST* from paternal X-chromosome and X-linked genes from maternal X-chromosome in E6.5 female VE cells. Related to Figure 1. Female VE cells undergo imprinted X-inactivation and therefore paternal X-chromosome is chosen as the inactive-X chromosome. *XIST* long noncoding RNA exclusively expresses from the inactive-X chromosome. (A) As expected, we found in almost all cells except few, allelic expression of *XIST* originated from paternal-X chromosome. (B) Profiling allelic expression of X-linked genes from maternal allele, showed >90% of expression from the active maternal-X chromosome almost in all cells and thus validating the accuracy of the allelic expression analysis method.

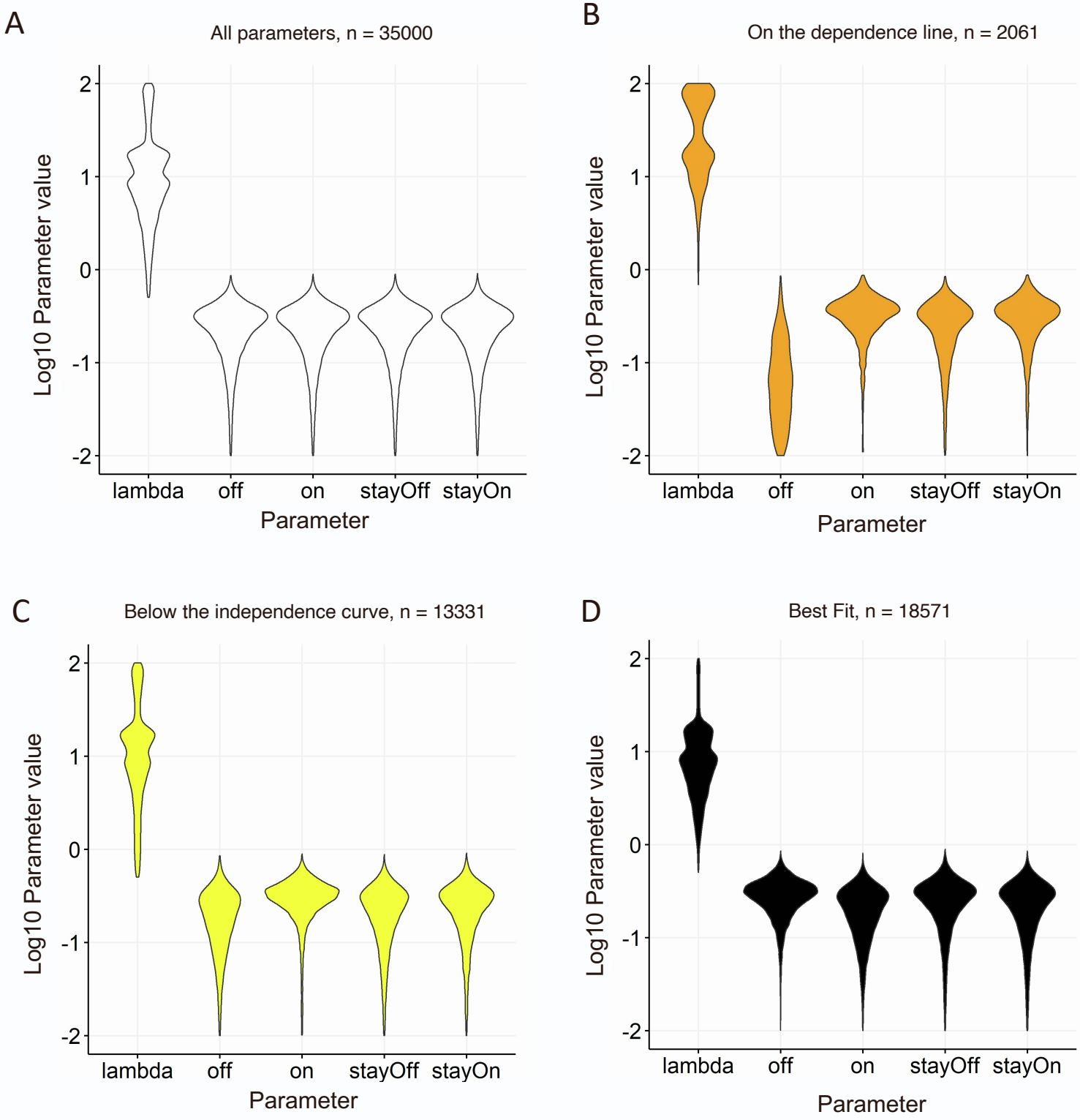

A
All parameters, n = 35000
B
On the dependence line, n = 2061
C
Below the independence curve, n = 13331
D
Best Fit, n = 18571
Log10 Parameter value
2
1
0
-1
-2
lambda
off
on
stayOff
stayOn
Parameter

Figure S3, Related to Figure 4: (A) Violin plots showing distributions from which each of the five parameters were sampled. (B) Parameter distribution for the genes represented with orange dots (strong dependence) in Figure 4B. (C) Parameter distribution for the genes represented with yellow dots (below the theoretical independence curve) in Figure 4B. (D) Parameter distribution for the genes represented with black dots (genes lying between the two theoretical curves) in Figure 4B.

Fig. S4

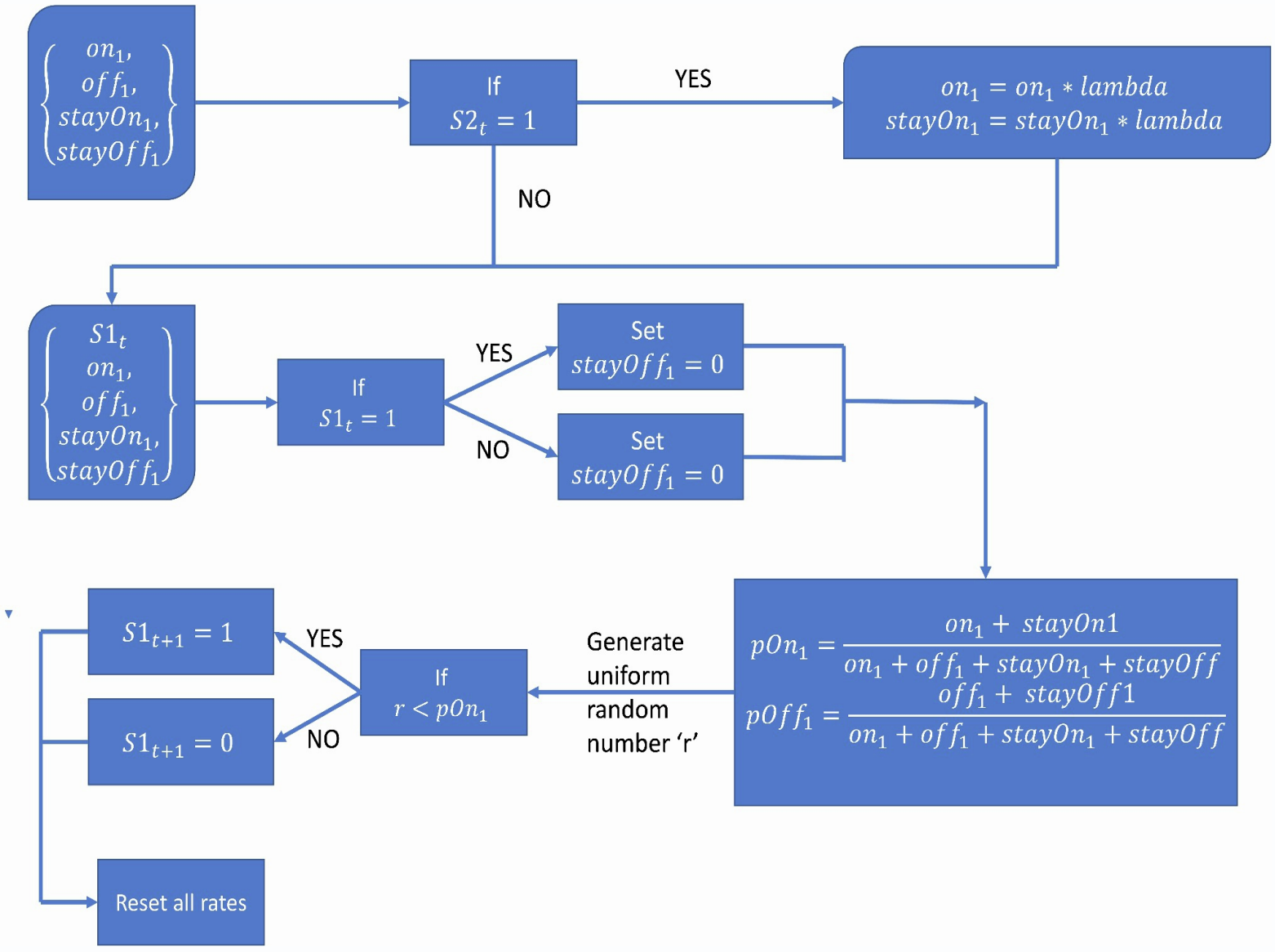

on_1, off_1, $stayOn_1$, $stayOff_1$
If $S2_t = 1$
YES
$on_1 = on_1 * lambda$
$stayOn_1 = stayOn_1 * lambda$
NO
$S1_t$ on_1, off_1, $stayOn_1$, $stayOff_1$
If $S1_t = 1$
YES
Set $stayOff_1 = 0$
NO
Set $stayOff_1 = 0$
$pOn_1 = \frac{on_1 + stayOn1}{on_1 + off_1 + stayOn_1 + stayOff}$
$pOff_1 = \frac{off_1 + stayOff1}{on_1 + off_1 + stayOn_1 + stayOff}$
Generate uniform random number 'r'
If $r < pOn_1$
YES
NO
$S1_{t+1} = 1$
$S1_{t+1} = 0$
Reset all rates

Figure S4, related to Figure 4: The model describes a gene with two alleles, with identical parameters. The simulation algorithm is described in the flow chart for one allele. The allele is described by the parameters {on1, off1, stayOn1, stayOff1}. At the beginning of the simulation, each of these parameters are sampled from a uniform distribution ranging from 0 to 1. The on/off state of the alleles at any given time t is given by {S1t, S2t}. The dependence parameter lambda is sampled from a uniform distribution ranging from 0.01 to 100.