nature portfolio

https://doi.org/10.1038/s41586-024-07841-y

Supplementary information

Embryonic genome instability upon DNA replication timing program emergence

In the format provided by the authors and unedited

SUPPLEMENTARY INFORMATION

Supplementary Note 1 | On the resolution of scRepli-seq data

For reliable binarization of scRepli-seq data derived from single S-phase cells, we empirically reasoned that (1) the NGS tag density profile must exhibit bimodal distribution that reflect the presence of replicated and unreplicated bins and (2) the average read count of unreplicated bins should be more than 20–30 to be distinguished from noise. When single mid-S mESCs with 4-6 million (M) scRepli-seq reads were analyzed, we found that 80-kb was the minimum bin size that met the above two conditions (Extended Data Fig. 1a and 1b, top panels in both). Of course, one can obtain more NGS read counts per bin by reading more. However, because single-cell-derived NGS library complexity is limited due to the scarcity of the starting material, valid read count (the number of reads with MAPQ>=10, excluding duplicate reads) does not increase in a manner proportional to the number of the total read count and reaches a limit at some point. For example, even when we increased the total number of reads per cell from 4.6 M to 18.4 M, the number of valid reads only went from 2.0 M to 5.2 M, and bimodal distribution could barely be confirmed at 40-kb resolution (Extended Data Fig 1a and 1b, bottom panels in both). These results indicate that it is difficult to dramatically improve NGS data resolution simply by increasing the total read count. From these observations, we reasoned that a total read count of 4–6 M per cell and a bin size of 80 kb were cost-effective and adopted them for non-haplotype-resolved scRepli-seq data analysis and binarization in this study. For haplotype-resolved scRepli-seq data binarization, we used 400-kb windows because we could recover only ~20% of all reads for each haplotype after haplotype-specific mapping[16].

Supplementary Note 2 | A detailed description of replication fork speed estimation

For the measurement and estimation of replication fork speed, we used a standard DNA fiber spreading assay using CldU/IdU incorporation. We noticed that the incorporation of CldU and IdU in cultured mammalian cells is frequently interrupted in a given DNA fiber, giving a 'dotty' pattern along the fiber (see Fig. 2h, 'mobile'), which has been observed by multiple laboratories[69]. What is the cause of imperfect IdU/CldU detection? In our DNA fiber spreading assay, we performed immunofluorescence using an anti-single-stranded (ss) DNA antibody in addition to IdU/CldU detection, which provided an answer to this question.

Interestingly, ssDNA staining also showed 'dotty' patterns with the signal frequently interrupted along the DNA fiber (Extended Data Fig. 5a). Intriguingly, almost all IdU/CldU dot signals colocalized with ssDNA dots, and IdU/CldU dots without ssDNA dots being colocalized were very rare (Extended Data Fig. 5a). Thus, gaps in ssDNA signals do not occur randomly; rather, they occur in specific regions of the genome and exhibit lack of antibody staining. It is possible that stretched DNA fibers cannot be fully denatured because the duplex DNA is anchored at multiple positions on the glass slide, which become the ssDNA signal gaps, while the ssDNA signal dots could represent stretches of denatured DNA. We could also assume that ssDNA dots without IdU/CldU signals are truly negative for IdU/CldU incorporation. With these observations, we have a much better understanding of how DNA replication is regulated during the first several cell cycles of mouse embryos after fertilization.

Our statistics based on the analysis of DNA fibers immediately after sequential 30-min IdU and CldU labeling can be summarized as follows:

(1) In 1-cell embryos, ~82% of DNA fibers contained the 'immobile' class forks, with single dots (signals) of IdU+CldU or CldU only that colocalized with ssDNA dot signals, which were on average ~22-kb apart from each other in early S-phase. Among all fibers, <11% contained the 'mobile' class forks with a median speed of ~0.38 kb/min.

(2) In 2- and 4-cell embryos, >63% of DNA fibers contained the 'immobile' class forks, with single dots of IdU+CldU or CldU only that colocalized with ssDNA dot signals, which were on average 12–14-kb apart from each other in early S-phase. Among all fibers, 19–24% contained the 'mobile' class forks with a median speed of 0.29–0.35 kb/min.

(3) In 8-cell embryos, <8% of DNA fibers contained the 'immobile' class forks and instead, >60% of fibers contained the 'mobile' class forks with a median speed of ~0.76 kb/min,

which is slower but comparable to that of mESCs (~1.2 kb/min). Our measurement is comparable to the reported median fork speed of 0.68 kb/min for 8-cell mouse embryos[9,27].

Based on the above statistics, at 1-, 2-, and 4-cell stages, 63–82% of the replication forks move so slowly ('immobile') that it is difficult to accurately measure their fork speed with DNA combing or DNA fiber spreading assays. Nonetheless, we can estimate their fork speed from the DNA fiber spreading assay since the microscopic resolution (0.3 μm) is about the length of a 1.3-kb DNA segment (1 μm = ~4.4-kb DNA, see Methods), and the gaps between ssDNA signal dots were on average ~1.0 μm (=~4.4 kb). As the total labeling time was 60 min (30 min + 30 min) and replication can start at any time during this period, we assumed the average labeling time of these dots of 'immobile' class forks to be 30 min. Based on these numbers, if we assume bi-directional forks, their speed must be 0.65–4.4 kb/30 min, i.e., 22–147 bp/min or less. If we assume uni-directional forks, their speed must be <1.3–4.4 kb/30 min, i.e., <43–147 bp/min. Taken together, the speed of the 'immobile' and 'mobile' classes of forks are <22–147 bp/min (or <43–147 bp/min for unidirectional forks) and 300–400 bp/min, respectively, based on our DNA fiber spreading assay. If we adopt the labeling time of the entire 60 min for the calculation, 'immobile' class bi-directional forks will be <11–73 bp/min (or <22–73 bp/min assuming uni-directional forks).

In addition, we can also take advantage of an orthogonal fork speed estimation method based on the measurements of (i) IOD, (ii) S-phase length based on EdU incorporation, and (iii) binarized whole-S scRepli-seq:
(i) IOD median values are 21.9, ~13.7, and ~12.3 kb for 1-, 2-, and 4-cell stages, respectively.
(ii) S-phase length is ~5, ~4, and ~11 hours for 1-, 2-, and 4-cell stages, respectively.
(iii) Binarized whole-S scRepli-seq data showed that at 80-kb resolution, there is virtually no bin-to-bin difference in the DNA synthesis rate in 1- and 2-cell embryos while the 4-cell embryos have acquired a somatic cell-like RT program.

Based on the values obtained from (i), (ii), and (iii), each 80-kb bin on the genome must contain 3–4 replication origins that are 22-kb apart from each other on average in 1-cell embryos. Thus, a given bi-directional fork must travel 11 kb on average to complete the task of genome duplication in 5 hours (=300 min), which translates to an average fork speed of 37 bp/min (=11 kb/300 min). Likewise, in 2-cell embryos with an IOD of ~14-kb, the average fork speed should be 29 bp/min (7 kb/240 min). Because the fork speed estimates were

remarkably similar between the 1- and 2-cell stages, and because the ratio between the types of forks does not change drastically until the 4-cell stage (Fig. 2i), we assume that the speed of the 'immobile' bi-directional forks in the 4-cell embryos is also on the order of ~33 bp/min [=(29+37)/2]. Importantly, the IOD was measured during the earliest 1 hour of S-phase (Fig. 2j). Therefore, if additional origins fire later in S-phase, the length of DNA that needs to be replicated by the earliest forks will naturally become shorter, meaning that these values are maximum values and forks could be slower depending on the later origin firing dynamics. If we assume uni-directional forks, the speed will be <~66 bp/min as they must travel double the distance of bi-directional forks.

Taken together, 1-, 2-, and 4-cell 'immobile' fork speed estimate is <33 bp/min based on a combination of IOD, S-phase length, and binarized whole-S scRepli-seq information. As this number is similar to <22–147 bp/min, which is the fork speed estimate of the 'immobile' class based on the DNA fiber spreading assay, the average fork speed of the 'immobile' class (which constitutes 63–82% of the fibers) is on the order of 33 bp/min or less in 1-, 2-, 4-cell embryos assuming bi-directional forks (or ~66 bp/min or less assuming uni-directional forks), while the IOD is on the order of 12–22 kb at these stages. The fastest forks ('mobile' class; 10–24% of all fibers) are on the order of 300–400 bp/min. The remaining class (7–15% of all fibers) must contain 'intermediate' velocity bi-directional forks that are on the order of 33–300 bp/min (or 66–300 bp/min assuming uni-directional forks). Importantly, these final numbers will not change even if we adopt the labeling time of 60 min for the fork speed calculation (after combining the two independent calculations to estimate the fork speed).

Palmerola[9] et al. and Nakatani et al.[27] reported that forks are on the order of 0.25 kb/min and 0.33 kb/min in 1-cell and 2-cell mouse embryos, respectively. As described above, our observation is that 60–80% of the forks ('immobile' class) are <33 bp/min in 1-, 2-, 4-cell embryos. So why is our number so different from the earlier study? It is possible that the 'immobile' class forks were overlooked as 'noise' (because they have never been observed in cultured cells) and only the fastest forks (i.e., the 'mobile' class) were measured within the cell population, which is only 10–24% of the total population in 1-, 2-, and 4-cell embryos. We believe that this could have been the case, because our 'mobile' class forks showed an average speed of 0.3–0.4 kb/min (median values of fork speed in 1, 2, and 4-cell stages are 0.38, 0.29, and 0.35 kb/min, respectively), which is almost identical to 0.25 or 0.33 kb/min from the earlier studies[9,27]. In addition, our 4-cell scRepli-seq NGS data was consistent with

the slow fork movement of < 33 bp/min (Extended Data Fig. 5c). If the average fork speed was 0.3–0.4 kb/min or more (4-cell average fork speed was 0.53 kb/min in Nakatani et al.[27]) with ~12-kb IOD, a given Mb-sized replication domain would complete its replication within an hour and we would not have seen the less prominent bimodal distribution of NGS tag density of 30–40% S-phase cells in 4-cell embryos (Extended Data Fig. 5c). This observation also supports our conclusion that the majority of replication forks in 1-, 2-, and 4-cell mouse embryos are much slower than previously reported.

Supplementary Note 3 | Types of chromosomal abnormalities observed and the interpretation of the nucleoside supplementation experiment

Chromosome abnormalities can generally be classified into two main types: 1) the 'break' type (structural error), often originating from problems in chromosome arms, and 2) the 'whole' type (numerical error), often originating from the merotelic attachment of centromeres to spindle microtubules. Both types are frequently preceded by anaphase problems such as chromosome lags or bridges. However, whereas anaphase chromosome lags or bridges leading to the 'break' type are often due to unresolved chromosome arms, those leading to the 'whole' type are often due to chromosomes with centromeres remaining near the center of the spindle. Among the abnormal anaphases observed during live imaging of embryonic 2-to-4-cell, 4-to-8-cell, and 8-to-16-cell divisions, we found no cases where centromeres remained in the middle of the spindle (Fig. 3c, 'merotelic'). Moreover, we found no cases where abnormal anaphase was preceded by chromosome congression defects at late metaphase (n=24, 59, and 20 cell divisions at 2-to-4-cell, 4-to-8-cell, and 8-to-16-cell divisions, respectively). Instead, we predominantly found anaphase problems with unresolved chromosome arms, which eventually resulted in breakage into unbalanced chromosome masses (Fig. 3c, 'unbalanced') or breakage leaving an acentric fragment (Fig. 3c, 'acentric fragment') by late anaphase. These live imaging observations are consistent with scRepli-seq results, which detected the 'break' type errors as the majority of chromosomal abnormalities (Fig. 3b).

Interestingly, nucleoside supplementation from 4-cell G1 greatly reduced chromosome aberrations at 4-to-8-cell anaphase (from 21.1% to 4.26%, Fig. 4j), while significantly but only modestly increasing replication fork speed (from 0.39 to 0.53 kb/min on average, Fig. 4g). We favor the idea that the 4-cell replication fork speed is slightly below the threshold required to complete S-phase without replication stress, and thus the modest increase in fork speed by nucleoside supplementation resulted in a great reduction in anaphase chromosome aberrations, although the possibility that other effects of nucleoside supplementation also contributed to the reduction in chromosome aberrations is not formally excluded.

Supplementary Table 1 | List of chromosome aberrations detected by scRepli-seq

Embryos at different stages were dissected into single cells and analyzed by scRepli-seq. Chromosome aberrations were categorized into gain/loss of whole chromosomes ('whole') or chromosome fragments ('break'). When a pair of cells within an embryo showed gain and loss of a whole or fragmented chromosome in a reciprocal manner, the chromosome aberration originated from the last cell division. When multiple pairs of cells within an embryo showed gain and loss of a whole or fragmented chromosome in a reciprocal manner, the chromosome aberration originated from a cell division preceding the last cell division. Non-reciprocal chromosome aberrations were not observed.

All reciprocal errors originating from the last cell division

Sample_name	stage	Embryo_ID	Error Chr No. (Whole gain)	Error Chr No. (Whole loss)	Error Chr No. (Break gain)	Error Chr No. (Break loss)
DEP-01-031-13	8-cell	8A				
DEP-01-031-14	8-cell	8A				16
DEP-01-031-15	8-cell	8A				
DEP-01-031-16	8-cell	8A				
DEP-01-031-17	8-cell	8A				
DEP-01-031-18	8-cell	8A				
DEP-01-031-19	8-cell	8A			16	
DEP-01-031-20	8-cell	8A				
DEP-01-031-37	8-cell	8B				
DEP-01-031-38	8-cell	8B				
DEP-01-031-39	8-cell	8B			12	
DEP-01-031-40	8-cell	8B				10
DEP-01-031-41	8-cell	8B				12
DEP-01-031-42	8-cell	8B				
DEP-01-031-43	8-cell	8B			10	
DEP-01-031-44	8-cell	8B				
DEP-01-031-69	8-cell	8E				
DEP-01-031-70	8-cell	8E				8
DEP-01-031-71	8-cell	8E				
DEP-01-031-72	8-cell	8E				
DEP-01-031-73	8-cell	8E				
DEP-01-031-74	8-cell	8E				

DEP-01-031-75	8-cell	8E			8	
DEP-01-031-76	8-cell	8E				
DEP-01-025-09	8-cell	8F				2
DEP-01-025-10	8-cell	8F				
DEP-01-025-11	8-cell	8F				
DEP-01-025-12	8-cell	8F				
DEP-01-025-13	8-cell	8F				
DEP-01-025-14	8-cell	8F				
DEP-01-025-15	8-cell	8F			2	
DEP-01-025-16	8-cell	8F				
DEP-01-025-32	8-cell	8G				
DEP-01-025-33	8-cell	8G			8	
DEP-01-025-34	8-cell	8G				
DEP-01-025-35	8-cell	8G				
DEP-01-025-36	8-cell	8G				8
DEP-01-025-37	8-cell	8G				
DEP-01-025-38	8-cell	8G				
DEP-01-025-39	8-cell	8G				
DEP-01-025-56	8-cell	8H				
DEP-01-025-57	8-cell	8H				7
DEP-01-025-58	8-cell	8H				
DEP-01-025-59	8-cell	8H			1	
DEP-01-025-60	8-cell	8H				1
DEP-01-025-61	8-cell	8H			7	
DEP-01-025-62	8-cell	8H				
DEP-01-025-63	8-cell	8H				
DEP-01-025-64	8-cell	8I				
DEP-01-025-65	8-cell	8I				
DEP-01-025-66	8-cell	8I			14	
DEP-01-025-67	8-cell	8I				
DEP-01-025-68	8-cell	8I				
DEP-01-025-69	8-cell	8I				14
DEP-01-025-70	8-cell	8I				
DEP-01-025-71	8-cell	8I				
DEP-01-025-87	8-cell	8J				
DEP-01-025-88	8-cell	8J				

DEP-01-025-89	8-cell	8J				
DEP-01-025-90	8-cell	8J				
DEP-01-025-91	8-cell	8J			1	
DEP-01-025-92	8-cell	8J				
DEP-01-025-93	8-cell	8J				
DEP-01-025-94	8-cell	8J				1
P445-9	8-cell	8K				
P445-10	8-cell	8K				
P445-11	8-cell	8K				
P445-12	8-cell	8K				
P445-13	8-cell	8K				
P445-14	8-cell	8K			3	
P445-15	8-cell	8K				3
P445-16	8-cell	8K				
P445-17	8-cell	8L				
P445-18	8-cell	8L				
P445-19	8-cell	8L				
P445-20	8-cell	8L				
P445-21	8-cell	8L				1
P445-22	8-cell	8L				
P445-23	8-cell	8L				
P445-24	8-cell	8L			1	
DEP-01-025-79	8-cell	8M				
DEP-01-025-80	8-cell	8M				
DEP-01-025-81	8-cell	8M				
DEP-01-025-82	8-cell	8M	8			
DEP-01-025-83	8-cell	8M				
DEP-01-025-84	8-cell	8M		8		
DEP-01-025-85	8-cell	8M				
DEP-01-025-86	8-cell	8M				
LCS_01_002_009	4-cell	4A				
LCS_01_002_010	4-cell	4A	6			1
LCS_01_002_011	4-cell	4A		6	1	
LCS_01_002_012	4-cell	4A				
LCS_01_001_041	8-cell	8N			16	
LCS_01_001_042	8-cell	8N				16

LCS_01_001_043	8-cell	8N				
LCS_01_001_044	8-cell	8N				
LCS_01_001_045	8-cell	8N				
LCS_01_001_046	8-cell	8N				
LCS_01_001_047	8-cell	8N				
LCS_01_001_048	8-cell	8N				
LCS_01_001_073	8-cell	8O				
LCS_01_001_074	8-cell	8O				
LCS_01_001_075	8-cell	8O		12		
LCS_01_001_076	8-cell	8O	12			
LCS_01_001_077	8-cell	8O				15
LCS_01_001_078	8-cell	8O				
LCS_01_001_079	8-cell	8O				
LCS_01_001_080	8-cell	8O			15	
LCS_01_002_029	8-cell	8P				1
LCS_01_002_030	8-cell	8P				
LCS_01_002_031	8-cell	8P				
LCS_01_002_032	8-cell	8P				
LCS_01_002_033	8-cell	8P				
LCS_01_002_034	8-cell	8P			1	
LCS_01_002_035	8-cell	8P				
LCS_01_002_036	8-cell	8P				
LCS_01_002_037	8-cell	8Q				
LCS_01_002_038	8-cell	8Q				X
LCS_01_002_039	8-cell	8Q			X	
LCS_01_002_040	8-cell	8Q				
LCS_01_002_041	8-cell	8Q				
LCS_01_002_042	8-cell	8Q				
LCS_01_002_043	8-cell	8Q				
LCS_01_002_044	8-cell	8Q				
LCS_01_002_077	8-cell	8R				
LCS_01_002_078	8-cell	8R				
LCS_01_002_079	8-cell	8R				X
LCS_01_002_080	8-cell	8R				
LCS_01_002_081	8-cell	8R				
LCS_01_002_082	8-cell	8R				

LCS_01_002_083	8-cell	8R				
LCS_01_002_084	8-cell	8R			X	
LCS_01_002_085	8-cell	8S			1	
LCS_01_002_086	8-cell	8S				
LCS_01_002_087	8-cell	8S				
LCS_01_002_088	8-cell	8S				
LCS_01_002_089	8-cell	8S				
LCS_01_002_090	8-cell	8S				
LCS_01_002_091	8-cell	8S				
LCS_01_002_092	8-cell	8S				1
LCS_01_002_085	8-cell	8T				
LCS_01_002_086	8-cell	8T			X	
LCS_01_002_087	8-cell	8T				
LCS_01_002_088	8-cell	8T				
LCS_01_002_089	8-cell	8T				X
LCS_01_002_090	8-cell	8T				
LCS_01_002_091	8-cell	8T				
LCS_01_002_092	8-cell	8T				
LCS_01_002_045	8-cell	8U				
LCS_01_002_046	8-cell	8U	8			
LCS_01_002_047	8-cell	8U				
LCS_01_002_048	8-cell	8U				
LCS_01_002_049	8-cell	8U				
LCS_01_002_050	8-cell	8U		8		
LCS_01_002_051	8-cell	8U				
LCS_01_002_052	8-cell	8U				
LCS_01_005_033	16-cell	16A				1
LCS_01_005_034	16-cell	16A				
LCS_01_005_035	16-cell	16A				
LCS_01_005_036	16-cell	16A				
LCS_01_005_037	16-cell	16A				
LCS_01_005_038	16-cell	16A				
LCS_01_005_039	16-cell	16A				
LCS_01_005_040	16-cell	16A				
LCS_01_005_041	16-cell	16A				
LCS_01_005_042	16-cell	16A				

LCS_01_005_043	16-cell	16A				
LCS_01_005_044	16-cell	16A				
LCS_01_005_045	16-cell	16A				
LCS_01_005_046	16-cell	16A				
LCS_01_005_047	16-cell	16A			1	
LCS_01_005_048	16-cell	16A				

All reciprocal errors originating from a cell division preceding the last cell division

Sample_name	stage	Embryo_ID	Error Chr No. (Whole gain)	Error Chr No. (Whole loss)	Error Chr No. (Break gain)	Error Chr No. (Break loss)
DEP_01_036_01	4-cell	4AA		13		
DEP_01_036_02	4-cell	4AA	13			
DEP_01_036_03	4-cell	4AA		13		
DEP_01_036_04	4-cell	4AA	13			
DEP_01_036_041	4-cell	4BB		10		10
DEP_01_036_042	4-cell	4BB		10		10
DEP_01_036_043	4-cell	4BB	10		10	
DEP_01_036_044	4-cell	4BB	10		10	
LCS_01_001_017	8-cell	8AA				
LCS_01_001_018	8-cell	8AA	1			
LCS_01_001_019	8-cell	8AA				
LCS_01_001_020	8-cell	8AA		1		
LCS_01_001_021	8-cell	8AA	1			
LCS_01_001_022	8-cell	8AA				
LCS_01_001_023	8-cell	8AA		1		
LCS_01_001_024	8-cell	8AA				
LCS_01_001_065	8-cell	8BB			X	
LCS_01_001_066	8-cell	8BB				
LCS_01_001_067	8-cell	8BB				
LCS_01_001_068	8-cell	8BB				X
LCS_01_001_069	8-cell	8BB				
LCS_01_001_070	8-cell	8BB				
LCS_01_001_071	8-cell	8BB				
LCS_01_001_072	8-cell	8BB			X	X
DEP-01-031-53	8-cell	8CC			X	X
DEP-01-031-54	8-cell	8CC			X	

DEP-01-031-55	8-cell	8CC				X
DEP-01-031-56	8-cell	8CC			X	
DEP-01-031-57	8-cell	8CC			X	
DEP-01-031-58	8-cell	8CC			X	
DEP-01-031-59	8-cell	8CC				X
DEP-01-031-60	8-cell	8CC		X		

Supplementary Video 1 | Live imaging of chromosome segregation at each cell division
Maximum intensity z-projection images of Major-satellite-mClover (centromeres, green) and H2B-mCherry (chromosomes, magenta) during mitosis of each cell division are shown. Time after nuclear envelope breakdown is shown (mm:sec). Scale bar is 5 μm.

Supplementary Video 2 | Live imaging of cell cycle progression from 1- to 8-cell stage
Live imaging of cell cycle progression from 1- to 8-cell stages. 3D-reconstructed images of mEGFP-PCNA (cell cycle marker, green) and H2B-mCherry (chromosomes, magenta). The cell lineage that exhibited chromosome segregation error during the 4-to-8-cell division is indicated by asterisks. Note that PCNA foci mark late S-phase (arrows). Arrowheads show micronuclei. Time after the first M-phase is shown (hr:mm). Signals are interpolated in z for visualization.

Supplementary References

69. Tuduri, S., Tourrière, H. & Pasero, P. Defining replication origin efficiency using DNA fiber assays. *Chromosome Res.* **18**, 91–102 (2010).