Article

Not peer-reviewed version

The Fragile Site Landscape of Induced Pluripotent Stem Cells: Hierarchy, Variability, Tissue Specificity, and Links to Culture Acquired Rearrangements

Pozhitnova V.O. , Zheglo D.G. * , Kislova A.V. , Kiselev D.S. , Voronina E.S.

Posted Date: 22 July 2026

doi: 10.20944/preprints202607.1652.v1

Keywords: induced pluripotent stem cells (IPSc); replication stress; fragile sites; aphidicolin; caffeine; mitotic DNA synthesis (MiDAS); chromosome aberration

 doi:10.20944/preprints202607.1652.v1

Article

The Fragile Site Landscape of Induced Pluripotent Stem Cells: Hierarchy, Variability, Tissue Specificity, and Links to Culture Acquired Rearrangements

Pozhitnova V.O. †, Zheglo D.G. *,†, Kislova A.V. , Kiselev D.S. and Voronina E.S.

Research Centre for Medical Genetics (RCMG)

* Correspondence: dianazheglo@gmail.com

† These authors contributed equally to this work.

Abstract

Induced pluripotent stem cells (iPSCs) are prone to genomic instability during prolonged culture, with recurrent chromosomal aberrations conferring selective advantages. Replication stress is a major driver of this instability, yet the repertoire of replication stress-sensitive loci in iPSCs remains largely unexplored. Here, we mapped aphidicolin-sensitive fragile sites (asFS) in three independent iPSC lines using classical cytogenetic break analysis combined with Monte Carlo simulation and MiDAS mapping directly on banded metaphase chromosomes. We identified 28 asFS, which segregated into a highly active Major cluster (8 sites, accounting for 59% of breaks among asFS) and a less active Minor cluster (20 sites). Five universal asFS (9p21, 6q25-26, 20p11-12, 10q22, Xq25) were present in all three lines, representing a tissue-specific fragility signature of pluripotent stem cells, with Xq25 shifting into the Major cluster after correction for X chromosome dosage. Minor asFS showed preferential co-localization with physical breakpoints or minimal overlapping regions of recurrent culture-acquired aberrations, including 20q11.21 (*BCL2L1*), 1q32 (*MDM4*), 8q24 (*MYC*), 17q21 (*WNT3-WNT9B*), and 18q21 (*DCC/FRA18B*). MiDAS mapping validated most asFS and revealed additional replication stress-sensitive loci in pericentromeric and subtelomeric regions that are difficult to score by conventional G-banding. Comparison with fragile site maps from other cell types confirmed tissue specificity, showing that the iPSC asFS repertoire is distinct in rank order and relative activity. Collectively, our findings indicate that the asFS repertoire in iPSCs is hierarchically organized into a stable universal core and a variable peripheral component, and suggest that Minor asFS may serve as initial substrates for adaptive culture-acquired rearrangements. This work provides a framework for understanding how replication stress and clonal selection shape the mutational landscape of pluripotent stem cells

Keywords: induced pluripotent stem cells (IPSc); replication stress; fragile sites; aphidicolin; caffeine; mitotic DNA synthesis (MiDAS); chromosome aberration

1. Introduction

Induced pluripotent stem cells (iPSCs) hold great promise for regenerative medicine, disease modeling, and drug discovery due to their unlimited self-renewal and differentiation capacity. However, their clinical application is hampered by a high propensity for genomic instability, which emerges during reprogramming and prolonged in vitro culture. Approximately 22% of iPSC lines acquire karyotype abnormalities during cultivation [1]. These include recurrent gains of chromosomes 1q, 20q, 12, 17, and X, as well as losses of 18q, which are preferentially selected during passaging and can affect pluripotency, differentiation potential, and potentially confer tumorigenic transformation [2–6]. For several of these recurrent aberrations, minimal overlapping regions (MORs) have been defined, pointing to specific genomic segments that are consistently gained or lost across independent iPSC lines and often contain genes implicated in cell survival, proliferation, or

 doi:10.20944/preprints202607.1652.v1

pluripotency (e.g., *BCL2L1* at 20q11.21, *MYC* at 8q24, *MDM4* at 1q32, *NANOG* at 12p13.33, *TP53* at 17p13, and *SALL3* in the minimal overlapping region of 18q loss) [7]. Understanding the molecular mechanisms that drive this instability is therefore of fundamental importance.

Replication stress (RS) is inherently linked to the pluripotent cell state and is considered one of the major factors driving both numerical and structural chromosomal instability in iPSCs [1,8–10]. The accelerated cell cycle characteristic of pluripotency, achieved primarily by shortening the G1 phase, result in elevated basal levels of replication stress markers, including stalled replication forks, single-stranded DNA accumulation, and γH2AX phosphorylation [8,11]. Under chronic replication stress, this leads to increased rates of lagging chromosomes in anaphase, atypical mitoses, and micronuclei formation, which are well documented precursors of chromosomal rearrangements [12,13].

RS does not affect the genome uniformly. Replication-sensitive sites manifest as non-random gaps, breaks, and exchange figures on metaphase chromosomes and include rare fragile sites (RFS), common fragile sites (CFSs), early-replicating fragile sites (ERFSs) [14], as well as fragile centromeres [15] and telomeres [16]. At the molecular level, these regions are intrinsically difficult to replicate due to repetitive sequences, secondary structures, transcription-replication collisions, and large replicons that delay fork progression. Consequently, such loci are prone to entering mitosis with under-replicated DNA. This gives rise to ultrafine anaphase bridges, which can lead to chromosome missegregation and segmental loss [17]. To resolve these structures, cells engage mitotic DNA synthesis (MiDAS), which is initiated by the structure-specific nuclease MUS81-EME1 and proceeds through RAD52-mediated strand annealing and POLD3-dependent synthesis [18,19].

The repertoire of RS-sensitive fragile sites is cell-type specific and shows inter-individual variability [20,21]. Historically, such sites have been mapped in lymphocytes and fibroblasts using classical cytogenetics [22,23]. Only a limited number of studies have addressed other histological origins [20,24,25] and the landscape of RS-sensitive sites in pluripotent stem cells remains largely unexplored [26]. Later, alternative markers of RS-sensitive regions have been applied for high-resolution global mapping of fragile sites, such as aphidicolin-induced FANCD2 ChIP-seq [27,28], Repli-seq [29,30] and MiDAS-seq [31], which showed high concordance with sites of chromosome fragility. Recently, the ERFS repertoire in embryonic pluripotent stem cells was mapped by colocalization of γH2AX, RPA, BRCA1, and SMC5 in early replicating regions under HU-induced stress [32]. Many of these loci fall within promoters and enhancers of genes controlling pluripotency, proliferation, and genome integrity, and overlap with cancer-associated CNVs and SNVs.

In this study, we combined classical cytogenetic break mapping with Monte Carlo simulation to identify loci of non-random breakage in three independent iPSC lines exposed to aphidicolin-induced RS. To functionally validate the identified loci, we performed MiDAS mapping directly on banded metaphase chromosomes, which detects RS–induced repair sites at cytogenetic resolution and is less dependent on optimal chromosome morphology. Finally, we compared the obtained map of aphidicolin-sensitive fragile sites (asFS) with breakpoints of recurrent chromosomal rearrangements that accumulate during long-term iPSC culture.

Our results reveal that the asFS repertoire in iPSCs is hierarchical, consisting of highly active (Major) and moderately active (Minor) clusters. Notably, several Minor asFS co localize with breakpoints of recurrent aberrations, suggesting their involvement in adaptive culture acquired rearrangements. Together, our findings indicate that the interplay between replication stress and subsequent clonal selection shapes the mutational landscape of iPSCs, providing a framework for improving culture conditions and monitoring genetic stability in pluripotent stem cells.

2. Materials and Methods

Cell lines and Culture Conditions

Three human induced pluripotent stem cell lines, P1L5 (RCMGi001-A), P12L3 (RCMGi017-A), and P16L1, were obtained from the Biobank “All-Russian Collection of Biological Samples of

 doi:10.20944/preprints202607.1652.v1

Hereditary Diseases" (Research Centre for Medical Genetics, Moscow, Russia). Cells were cultured on Vitronectin-coated plates (A14700, Gibco™) in Essential 8™ (A1517001, Gibco™) or Gibris 8 (C780E, Pan-eco-ltd, Russia) medium supplemented with 50 U/mL penicillin and 50 µg/mL streptomycin. Treatments were performed when cells reached approximately 50–60% confluency, at least two days after reseeding. Replication stress was induced by adding aphidicolin (A4487, Sigma Aldrich) at a final concentration of 0.1µM, together with 1 mM caffeine (C0750, ReagentPlus®), 21 hours before cell harvest. For MiDAS detection, 20 µM EdU was added 40 minutes prior to harvest.

Metaphase Spread Preparation and Chromosome Banding

For metaphase preparation, cells were synchronized with 0.1 µg/mL Demecolcine (D7385, Sigma Aldrich) for 1 hour, harvested by trypsinization (Pan-eco-ltd, Russia), hypotonized in prewarmed 0.55% KCl for 13–14 minutes, and fixed in three changes of ice-cold methanol/glacial acetic acid (3:1, Chimmed, Russia). The cell suspension was dropped onto microscope glass slides (Menzel Glaser, Germany) and air-dried. To obtain G-like banded chromosomes, slides were stained with Vectashield anti-fade mounting medium with DAPI (H-1200, Vector Labs) and premixed 0.3 mg/mL Actinomycin D (1071001, Serva) for band contrast. Metaphase images were acquired using an AxioImager 2 (Zeiss) microscope and processed using the Q-banding mode of Case Data Manager 6.0 software (Applied Spectral Imaging). Chromosomal break locations were analyzed according to the International System for Human Cytogenomic Nomenclature (ISCN 2024).

MiDAS Detection

EdU incorporation was detected using the Click-iT™ EdU Cell Proliferation Kit for Imaging (C10339, Invitrogen) according to the manufacturer's instructions. Briefly, slides were blocked in 1% BSA, permeabilized with 0.5% Triton X-100 in PBS, and incubated with Alexa-594-picolyl azide in supplied click buffer, followed by three washes with PBS/Triton. Slides were counterstained with Glycerol Mounting Medium with DAPI (ab188804, Abcam).

Monte Carlo Simulation for Threshold Determination

To distinguish genuine fragile sites from randomly distributed breaks, we performed Monte Carlo simulations under the null hypothesis that breaks are uniformly distributed across all cytogenetic bands. The human haploid karyotype contains approximately 400 bands at standard resolution (ISCN 2024); each band was treated as a single analytical unit. For each simulation, the total number of breaks observed (N) was randomly assigned to the 400 bands with equal probability ($p = 1/400$) following a multinomial distribution. In each of 100,000 iterations, the maximum number of breaks occurring in any single band was recorded. The 99th percentile of this distribution was chosen as the significance threshold, corresponding to a one-tailed false positive rate of 1%. For the pooled dataset (966 breaks), the threshold was 11 breaks per locus. For individual lines, the thresholds were 7 breaks for P1L5 (342 breaks), and 6 breaks for both P12L3 (318 breaks) and P16L1 (306 breaks). The 95th percentiles were also computed for comparison (pooled: 9 breaks; per line: 5 breaks) and are provided in Supplementary Table S1. Simulations were performed in Python 3.12 using NumPy 2.0.2 with a fixed random seed (42) for reproducibility.

Identification and Classification of asFS

Loci with 11 or more breaks across all three lines were designated as significant aphidicolin-sensitive fragile sites (asFS). Each asFS was then examined for its activity pattern across individual lines using per-line thresholds (≥7 for P1L5, ≥6 for P12L3 and P16L1). Based on these patterns, loci were classified as Universal (significant in all three lines), Repressed (significant in two lines), Line-specific (significant in only one line), or Pool-only (significant only by pooled threshold). The classification is summarized in Supplementary Table S1.

 doi:10.20944/preprints202607.1652.v1

Hierarchical Clustering by Break Frequency

To identify intrinsic groupings among the 28 significant asFS based solely on break counts, we performed k-means clustering validated by silhouette analysis and gap statistic. Both methods consistently supported a two-cluster solution (silhouette score = 0.72, gap statistic = 1.34), separated by a natural gap of 9 breaks between the 7th and 8th loci (34 vs. 25 breaks). These clusters were designated as Major (top 8 loci after X chromosome dosage correction) and Minor (remaining 20 loci). Details are provided in Supplementary Figure S2 and Supplementary Table S1.

3. Results

3.1. The Landscape of Aphidicolin Sensitive Fragile Sites in iPSCs

To induce chromatid breaks, we treated cells with a low dose of aphidicolin, a reversible inhibitor of replicative polymerases, combined with caffeine to attenuate ATR and ATM signaling and thereby override cell-cycle arrest and cell death. Metaphase chromosomes were harvested 21 hours after treatment initiation. This duration has been shown to permit completion of a full S-phase under replication stress conditions [33]. The experimental conditions therefore allow for the induction of both ERFSs [13] and canonical CFSs. Across three independent iPSC lines, we scored 966 chromatid breaks distributed over 110 distinct cytogenetic bands (Supplementary Table S1).

The traditional threshold for fragile site identification, empirically established at ≥1% of total breaks when scoring approximately 300 breaks, has been widely used to define non-randomly fragile loci [20]. In this study, to distinguish genuine asFS from random background breakage, we performed Monte Carlo simulations of random break distribution across the 400 bands resolvable at standard cytogenetic resolution. At the 99% confidence level, the pooled threshold for non-random break accumulation was 11 breaks per locus across all three lines (Supplementary Figure S1). This threshold is equivalent to 1.2% of breaks and thus slightly exceeds the traditionally used cutoff. For individual lines, the thresholds were 7 breaks for P1L5 and 6 breaks for P12L3 and P16L1, reflecting the higher stringency required to call a locus significantly fragile in a single line.

Applying the pooled threshold (≥11 breaks), we identified 28 loci with significant non-random break accumulation (Figure 1a, 1b). These 28 asFS accounted for 691 of 966 breaks (71.6%), demonstrating that they represent the principal targets of aphidicolin induced chromosomal damage in iPSCs. The remaining 305 breaks (28.4%) were distributed across other 82 bands at frequencies below the statistical threshold, consistent with either stochastic breakage or low activity asFS that fall below the detection limit of this analysis.

3.2. Conservation of the Fragility Pattern Across Lines

The activity of the 28 asFS identified in the pooled analysis varied considerably between the three lines. When we applied the more stringent per-line thresholds, 17 loci reached significance in P1L5, 13 in P12L3, and 13 in P16L1 (Figure 1b, Figure 1c, Supplementary Table S1). These inter-line differences allowed us to classify the 28 asFS into four categories based on their conservation across lines.

Five loci exceeded the per-line threshold in all three lines: 9p21, 6q25-26, 20p11-12, 10q22, and Xq25. These universal asFS are constitutively fragile and represent the minimal fragility signature of the iPSC state. Six loci showed a repressed pattern, being active in two lines but significantly suppressed in the third. Repression occurred predominantly in P16L1 (four loci), with two additional loci repressed in P1L5, while P12L3 showed no repressed loci. This asymmetric distribution suggests that P16L1 has a distinct epigenetic landscape or a higher propensity for context-dependent modulation. The remaining loci comprised line-specific (13 loci) and pool-only (4 loci) sites. In addition, three loci that did not meet the pooled threshold were significant in individual lines (8q21-22 and 18p11.2 in P16L1, and 8q23-24 in P1L5), indicating that the full repertoire of aphidicolin-sensitive fragility in iPSCs extends beyond the 28 loci captured by pooled analysis.

 doi:10.20944/preprints202607.1652.v1

Inter-individual variability of fragile site landscapes has been previously observed in primary lymphocytes and dermal fibroblasts, with the level of conservation estimated at 45–48% [20], which is comparable to our findings. The discrepancies between the pooled set and individual lines may arise from genuine donor-specific epigenetic variation, microdeletions within a fragile site, or stochastic sampling due to the finite number of metaphases scored. Consequently, the absence of a locus from an individual line's significant list does not necessarily imply biological absence, as it may be active but at a level below our detection limit.

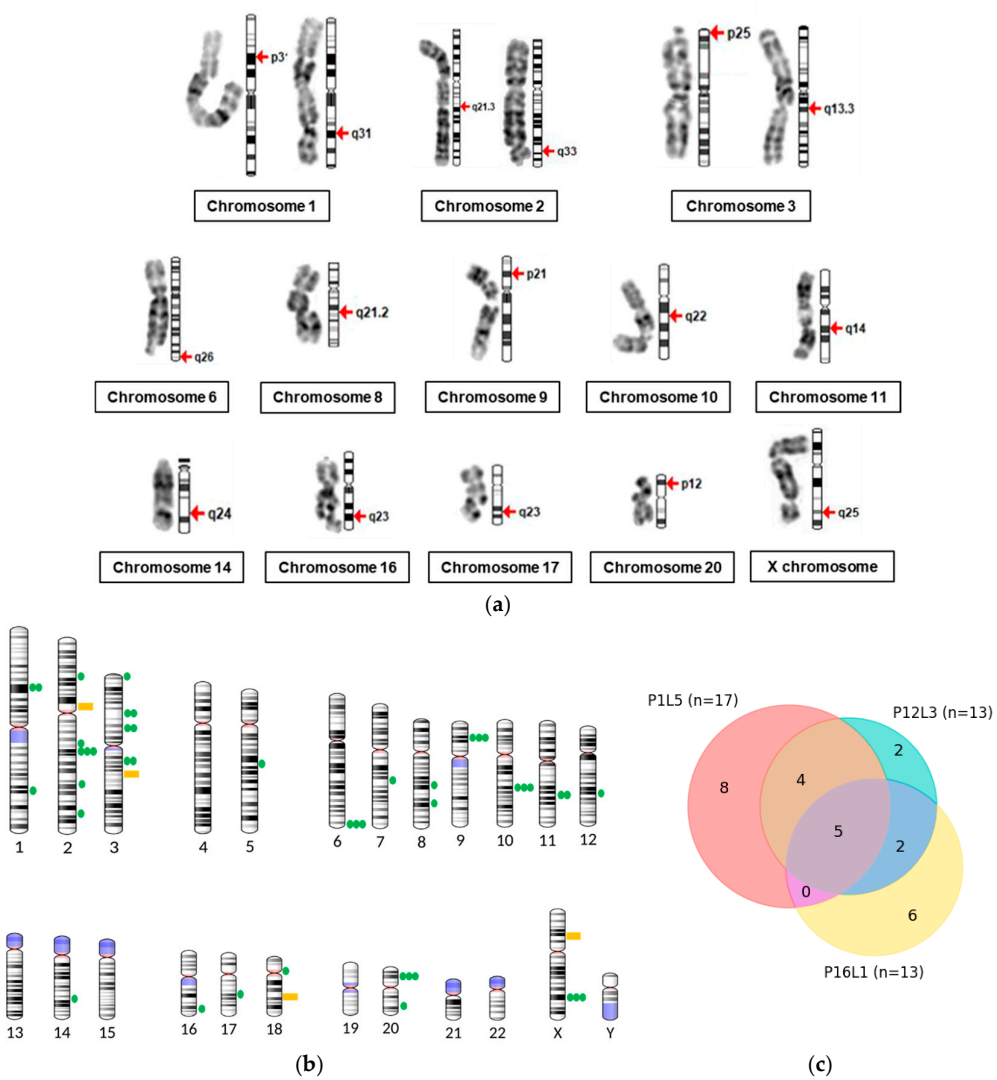
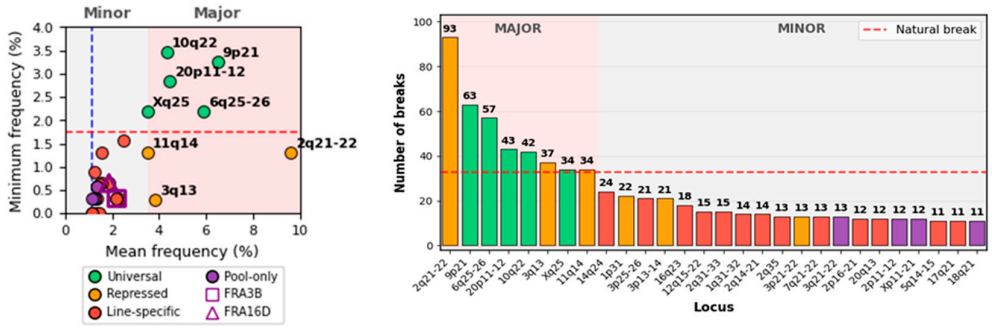


Figure 1. Mapping of aphidicolin-sensitive fragile sites (asFS) in iPSCs. (**a**) Representative images of chromatid breaks on banded metaphase chromosomes at the most active asFS loci. Arrows indicate breaks at the indicated chromosomal bands; (**b**) Ideogram of all 31 asFS mapped to human chromosomes. Each green dot represents an asFS locus; the number of dots per locus corresponds to the number of iPSC lines (P1L5, P12L3, P16L1) in which the locus is significantly fragile by the Monte-Carlo test (≥5 breaks per line). Orange boxes indicate loci that are significant only by the pooled threshold (≥9 breaks total) but do not reach significance in any individual line; (**c**) Venn diagram of loci exceeding per-line significance thresholds across three iPSC lines. Each circle represents the set of loci significant in P1L5 (red), P12L3 (teal), or P16L1 (yellow). Numbers indicate loci in each intersection. The central intersection of five loci (9p21, 6q25-26, 20p11-12, 10q22, and Xq25) represents the universal fragility signature. Four loci that were significant only in pooled analysis (3q21-22, 2p11-12, Xp11-21, 18q21) are not shown in the Venn diagram.

3.3. Major and Minor asFS: Two Tiers of Genomic Instability

 doi:10.20944/preprints202607.1652.v1

When we examined the distribution of break counts among the 28 asFS, a striking hierarchical structure emerged. K-means clustering, validated by silhouette analysis and gap statistic, consistently resolved the loci into two distinct groups separated by a natural gap between the 7th and 8th loci (34 vs. 25 breaks) (Supplementary Figure 2a). To account for the single X chromosome copy in the male line (P1L5), we corrected the break count of Xq25 by doubling its contribution from this line. This adjustment yielded a total of 34 breaks for Xq25, shifting it into the Major cluster and bringing all five Universal loci into the top tier (Figure 2a, 2b). The Major cluster thus comprised eight loci with break counts ranging from 34 to 93 (mean 49.4), accounting for 59% of all breaks occurring within asFS, despite representing only 29% of the loci. These eight loci were 2q21-22, 9p21, 6q25-26, 20p11-12, 10q22, 3q13, 11q14, and Xq25. The Minor cluster included the remaining 20 loci with break counts ranging from 11 to 25 (mean 15.6), accounting for the remaining 41% of breaks. This group contained all Repressed, Line-specific, and Pool-only loci, reflecting their more variable and context-dependent expression. If X inactivation in female lines additionally reduces fragile site expression, the true activity of Xq25 on the active X homolog would be even higher than estimated here.



(**a**) (**b**)

Figure 2. Break count distribution across 28 asFS in iPSCs. (**a**) Scatter plot of cytogenetic break distribution across 28 asFS in iPSCs. Each point represents a cytogenetic locus, plotted by its mean break frequency (x axis) versus its minimum frequency across the three lines (y axis). Loci are color coded according to their conservation pattern: Universal (green), Repressed (orange), Line specific (red), Pool only (purple). The dashed vertical line indicates the pooled Monte Carlo threshold (1.14%, equivalent to 11 breaks). The dashed horizontal line marks the per line significance threshold (1.75%, equivalent to 7 breaks in P1L5), below which a locus fails to reach significance in at least one individual line. Background shading separates the Major cluster (pink) from the Minor cluster (grey). FRA3B and FRA16D are highlighted. Universal loci and Major cluster Repressed loci are labeled with band names. (**b**) Loci are ranked by total break count across three lines. Bars are color coded according to conservation pattern: Universal (green), Repressed (orange), Line specific (red), Pool only (purple). Background shading separates the Major cluster (pink) from the Minor cluster (grey), with the boundary determined by k means clustering. The dashed red line marks the natural break between clusters at 33 breaks. Numbers above bars indicate exact break counts.

Our observation is in line with previous reports indicating that a relatively small number of fragile loci account for the majority of chromosomal breaks under replication stress conditions [20]. The Major cluster thus represents a small set of dominant fragility hotspots that collectively account for the majority of replication stress-induced chromosomal damage in iPSCs. Their universal conservation across independent lines suggests they are intrinsic to the pluripotent state. Conversely, the Minor cluster comprises a larger and more heterogeneous group of sites with lower baseline fragility that are more susceptible to line-specific modulation, whether by epigenetic differences, genetic background, or stochastic variation during reprogramming.

3.4. MiDAS Mapping Refines the Landscape of Aphidicolin Sensitive Loci

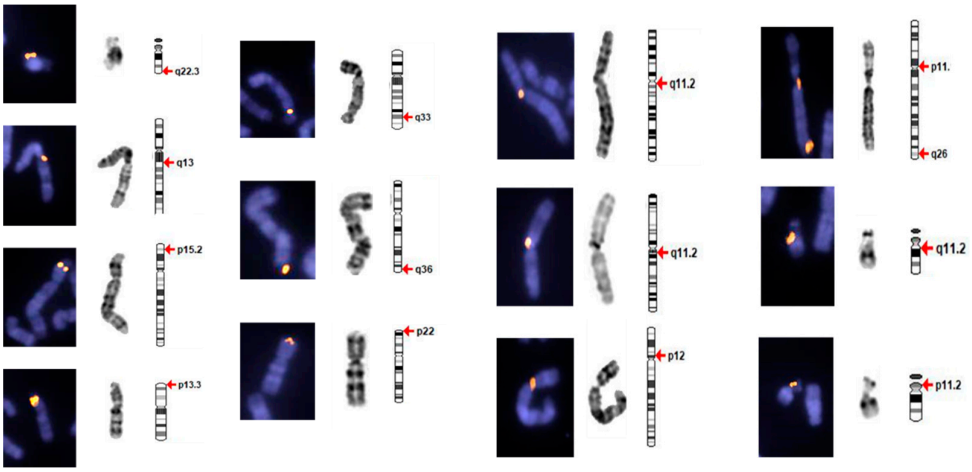
 doi:10.20944/preprints202607.1652.v1

To independently validate the 28 asFS identified by break counting and to capture loci potentially missed in cytogenetically challenging regions, we performed MiDAS site mapping in the same three iPSC lines. Following 21-hour aphidicolin/caffeine treatment and synchronization at the G2/M boundary, cells were released into mitosis and pulse-labeled with EdU to detect DNA synthesis. A total of 451 MiDAS signals were scored on metaphase chromosome spreads (147 in P1L5, 153 in P12L3, and 151 in P16L1) and mapped to cytogenetic bands (Figure 3a, 3b, Supplementary table S2).

We observed concordance between MiDAS- and break-based asFS for multiple loci (Figure 3c). For several of them, the interline distribution of MiDAS signals closely paralleled that of chromatid breaks, most notably at 2q21-22, where both methods showed highest activity in P12L3, intermediate in P1L5, and lowest in P16L1. This agreement across independent methods suggests that both report genuine biological fragility rather than method-specific artifacts. Of the 28 asFS, 22 (79%) showed MiDAS signals in at least one line. All eight Major cluster loci exhibited robust MiDAS signals across all three lines, confirming their status as genuine fragility hotspots. The six asFS without detectable MiDAS signals were exclusively Minor cluster loci with lower break counts, which may reflect technical limitations of the EdU pulse-labeling window or the relatively small number of MiDAS events scored per line, rather than the absence of mitotic DNA synthesis at these sites.

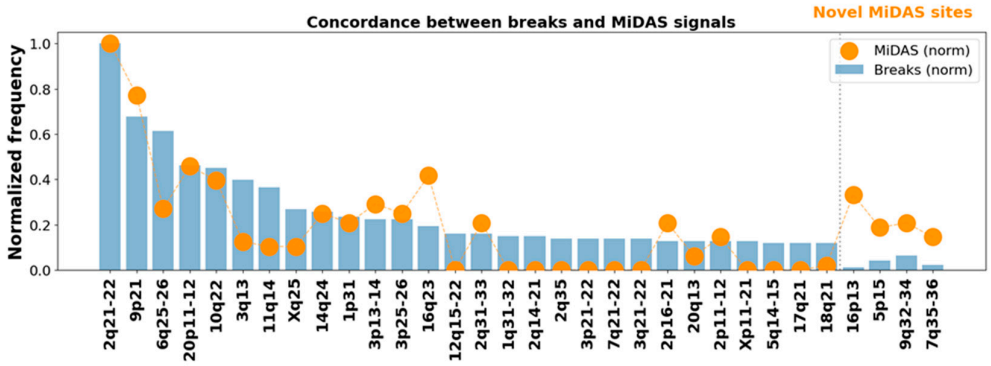
Comparison of the activity clusters between chromosomal breaks and MiDAS signals revealed a partial but not absolute concordance. While most Major asFS showed strong MiDAS signals, two (3q13 and 11q14) exhibited unexpectedly low MiDAS counts (6 and 5 signals, respectively), despite their high break frequencies. Xq25, which shifts to the Major cluster after X chromosome dosage correction in the male line, showed moderate MiDAS signals (7 signals), consistent with its position among the lower-activity Major sites. Conversely, several Minor asFS displayed MiDAS signals comparable to or even exceeding those of Major sites. Notably, 16q23, a Minor cluster locus by break count, showed prominent MiDAS signals across all three lines (Figure 3c, Supplementary table S2).

Beyond confirming known asFS, MiDAS mapping revealed novel MiDAS-positive loci that had not been identified by break counting. These include 16p13, 5p15, 9q32-34, and 7q35-36, all of which exhibited MiDAS signals comparable to those of Major sites despite having few or no breaks (Figure 3a. 3c). These loci map to cytogenetically challenging regions (pericentromeric, subtelomeric, and pale G-bands), where chromatid breaks are difficult to score by conventional G-banding. These results highlight the complementary nature of both approaches: break counting identifies direct chromosomal damage, while MiDAS captures repair-associated events and can reveal fragility in regions that are systematically underrepresented in cytogenetic analysis. Together, they provide a more complete picture of the aphidicolin-sensitive landscape in iPSCs.



(**a**) (**b**)

 doi:10.20944/preprints202607.1652.v1



(**c**)

Figure 3. Cytogenetic mapping of aphidicolin-induced MiDAS sites in iPSC (**a)** Representative images of MiDAS signals at cytogenetically challenging regions and newly identifies RS-sensitive loci; (**b**) Centromeric MiDAS signals **(c)** Concordance between chromatid break frequencies and MiDAS signals across aphidicolin-sensitive fragile sites (asFS) in iPSCs. Loci are ranked from left to right by decreasing total break counts. Blue bars represent normalized break frequencies; orange circles indicate normalized MiDAS signal counts. The dashed orange line connects MiDAS values to highlight deviations from the break-based ranking. Normalization was performed relative to the maximum value within each dataset to allow direct comparison of the two profiles. The overall ranking is largely consistent between the two methods, with most loci showing proportional break and MiDAS signals. Notable discrepancies are observed for 16p13, 5p15, 9q32-34, 7q35-36, which show higher MiDAS signals than expected from break counts, and for 3q13 and 11q14, which show lower MiDAS signals despite high break frequencies. Four loci not identified as asFS by break counting, 16p13, 5p15, 9q32-34 and 7q35-36, are shown separately on the right side of the plot as novel MiDAS-positive candidates with high signal counts.

3.5. Tissue-Specific Fragile Site Landscape in iPSCs

Comparison of the iPSC asFS repertoire with canonical fragile sites described in somatic cell types revealed substantial tissue-specific differences. Two of the most extensively characterized common fragile sites in human lymphocytes and fibroblasts, *FRA3B* (3p13-14) and *FRA16D* (16q23), were both identified within our iPSC dataset, but neither dominated the landscape. FRA3B (21 breaks) fell into the Minor cluster and showed repression in P16L1, while FRA16D (18 breaks) was line-specific to P1L5. Their relatively modest activity in iPSCs contrasts sharply with their well-documented prominence in lymphocytes, fibroblasts and epithelial cells [20,23], where they consistently rank among the most frequently expressed fragile sites. This reduced activity supports the notion that the fragile site landscape undergoes substantial remodeling upon reprogramming to pluripotency.

Conversely, *FRA2Ftel* (2q21-22), a fragile site that has received less attention in somatic cell studies, emerged as the most active locus in our iPSC dataset (93 breaks), with pronounced line-specific differences (highest in P12L3, intermediate in P1L5, lowest in P16L1). Its high activity across all three lines suggests that certain loci may acquire or maintain fragility in the pluripotent state, possibly reflecting the altered replication program and chromatin landscape characteristic of iPSCs.

Comparison with lymphocytes, fibroblasts, and epithelial cells shows that the iPSC asFS repertoire is distinct in rank order and relative activity, although some loci are retained across cell types [20,25]. This tissue-specific remodeling suggests that the regions most vulnerable to replication stress differ fundamentally between pluripotent and differentiated cells, reflecting the unique biology of the pluripotent state.

3.6. Co-Localization of Minor asFS with Recurrent Aberration Breakpoints

To assess whether the identified asFS may contribute to culture-acquired chromosomal rearrangements, we compared their genomic localization with known physical breakpoints and

 doi:10.20944/preprints202607.1652.v1

minimal overlapping regions (MORs) of recurrent aberrations documented in long-term iPSC cultures. The universal Major asFS (9p21, 6q25-26, 20p11-12, Xq25, 14q24) do not coincide with any known recurrent rearrangement hotspot.

1q Gain

Interstitial duplications of 1q or derivative chromosomes with various translocation partners are among the most frequent structural aberrations recurrently arising in cultured iPSCs and their derivatives. The minimal overlapping region (MOR) encompasses the MDM4 oncogene at 1q32. Varying breakpoints have been reported, including centromeric regions, 1q21, 1q23, 1q25, and other bands [1,34,35]. At the molecular level, SNP array analysis mapped breakpoints of two independent aberrations to q23.3–q42.12 and q21.2 [36]. In our study, two Minor asFS were identified on chromosome 1: at 1p31 (rank 10) and 1q31-32 (rank 16). The latter falls within the MOR of recurrent 1q gain. In addition, several non-significant breakage sites (1q42, 6 breaks; 1q24-25, 3 breaks) and MiDAS-positive loci (1q22-31, 4 signals; 1q12, 1 signal; 1q32, 1 signal) were detected, suggesting that this region may harbour additional low-level fragility.

Amplification of 20q11.21

Recurrent gain of the 20q11.21 region occurs through tandem duplications, triplications, isochromosome formation, and other gross rearrangements with varying breakpoints [37–39]. The MOR of 20q gain centers on BCL2L1, an anti-apoptotic gene that confers a selective growth advantage. In our study, a fragile and MiDAS-positive band was identified at 20q13 (FRA20E), which lies distal to the core 20q11.21 amplification region. Although 20q13 does not overlap the main MOR, its involvement in the formation of 20q amplifications cannot be excluded, as breakpoints may extend to the distal q-arm in some rearrangements.

18q Loss

Loss of 18q is a recurrent event in cultured iPSCs, frequently associated with reduced differentiation potential. The breakpoint has been mapped within the DCC gene at 18q21.2 [5,36], while the MOR contains *SALL3*, a gene implicated in pluripotency regulation. In our dataset, we identified FRA18B at 18q21 as a Minor asFS (18q21, rank 27), which directly overlaps the physical breakpoint within *DCC*. This suggests that the fragility at *FRA18B* may contribute to the initiation of 18q deletions.

8q Gain

Gain of 8q, particularly at the *MYC* locus (8q24), is a well-documented recurrent aberration in iPSCs. The MOR centres on *MYC*, a key oncogene driving proliferation. Although physical breakpoints for 8q gain have not been precisely mapped, we identified a Minor asFS at 8q23-24 (*FRA8C*/*D*), which lies within the MOR of 8q gain. This localization raises the possibility that fragility at this site may contribute to 8q amplification events, though further molecular characterization is needed.

12p Gain

Amplification of 12p13.33, encompassing the pluripotency factor *NANOG*, is a common adaptive change in iPSC cultures. No asFS were identified on the short arm of chromosome 12; the only asFS on this chromosome is located on the opposite arm at 12q15-22 (*FRA12L*, *ANKS1B* [34]), which does not overlap with the 12p gain region.

17q Gain, 17p Loss

The short arm of chromosome 17, particularly the 17p13.1 region containing the *TP53* gene, is a site of sporadic deletions that can confer a selective growth advantage. Additionally, the 17q21.31/*WNT3-WNT9B* amplification has been reported as a “hot spot” for genetic variation altering cell phenotype [41]. In our study, two fragile loci were identified on chromosome 17. The locus 17q21 (11 breaks) exceeded the pooled significance threshold, while 17q22-23 (6 breaks) did not.

Overall, our findings support a model in which the recurrent aberration landscape of cultured iPSCs arises from the interplay between replication-stress-induced fragility and subsequent clonal

 doi:10.20944/preprints202607.1652.v1

selection for growth-promoting changes. Minor asFS and even non-significant break sites appear to be preferentially captured in culture-acquired rearrangements, whereas the most active Major loci are conspicuously absent from known recurrent breakpoints. This suggests that regions of extreme fragility may produce lesions that are either lethal or lack selective advantage, whereas moderately fragile sites provide viable substrates for clonal expansion. Nevertheless, it remains possible that Major asFS contribute to submicroscopic or low-level mosaicism that is rarely documented in cytogenetic surveys.

4. Discussion

In this study, we provide the first systematic cytogenetic characterisation of aphidicolin-sensitive fragile sites (asFS) in human pluripotent stem cells. The resulting map offers a valuable reference for interpreting genomic instability events in iPSCs and for targeted screening of recurrent genetic aberrations that arise during culture. The asFS repertoire we identified is tissue-specific and hierarchically organised, comprising a stable universal core and a variable peripheral component. Despite inter-line variability, five loci (9p21, 6q25-26, 20p11-12, 10q22, and Xq25) were consistently fragile across all three lines, accounting for nearly 40% of all breaks at statistically significant sites. This suggests that the fragility of pluripotent cells is not random but reflects a defined, cell-type-specific genomic architecture.

This architecture is further resolved into two hierarchical tiers. The Major cluster, comprising eight highly active loci, accounts for the majority of replication-stress-induced damage despite representing only a quarter of the asFS repertoire. These sites are predominantly universal, linking high activity with cross-line conservation. In contrast, the Minor cluster contains most repressed and line-specific loci, which are more susceptible to modulation by epigenetic or stochastic factors.

The functional relevance of Minor asFS is underscored by their preferential co-localisation with breakpoints of recurrent culture-acquired aberrations, including 20q11.21, 1q32, 8q24, 17q21 and 18q21. The asFS that overlap with recurrent aberrations are predominantly Minor sites, whereas the most active Major loci are conspicuously absent from known rearrangement hotspots. This suggests that extreme fragility may produce lesions that are either lethal or lack selective advantage, whereas moderately fragile sites provide viable substrates for clonal dominance.

The integration of MiDAS mapping further refines the landscape by revealing additional replication-stress-sensitive loci in cytogenetically challenging regions (pericentromeric, subtelomeric, and pale G-bands) that are systematically underrepresented in break-based analyses. Novel candidates such as 16p13, 5p15, 9q32-34, and 7q35-36 illustrate how repair-based mapping can complement cytogenetic break counting, particularly in regions where chromosome morphology limits detection. This dual approach captures both the damage and the repair response, providing a more complete picture of replication-stress sensitivity in iPSCs.

In summary, our findings indicate that the asFS repertoire in iPSCs forms a structured, cell-type-specific map of genomic vulnerability. The distinction between Major and Minor clusters may have functional implications: Minor asFS, despite their lower individual activity, show preferential co-localisation with breakpoints or minimal overlapping regions of recurrent culture-acquired rearrangements. This suggests, but does not prove, that moderately fragile sites could serve as substrates for clonal selection, while the most active Major loci remain conspicuously absent from known rearrangement hotspots. Whether this reflects lethality, lack of selective advantage, or simply the limits of detection remains an open question. Nevertheless, the observed co-localisation patterns provide a rationale for further investigation into the relationship between intrinsic fragility and the recurrent karyotypic landscape of cultured iPSCs.

Supplementary Materials: The following supporting information can be downloaded at website of this paper posted on Preprints.org. Figure S1: Monte Carlo simulation results for threshold determination; Figure S2: Statistical validation of two-cluster solution for 28 fragile sites; Table S1: Identification of aphidicolin sensitive fragile sites in three iPSC lines; Table S2: Cytogenetic mapping of MiDAS loci in three iPSC lines.

 doi:10.20944/preprints202607.1652.v1

Author Contributions: Conceptualization, D.Z. and E.V.; methodology, D.Z. and V.P.; formal analysis, D.Z.; investigation, V.P., D.Z., A.K., D.K.; writing—original draft preparation, D.Z. and V.P.; writing—review and editing, E.V., D.Z., V.P., A.K., D.K.; visualization, D.Z. and V.P.; supervision, E.V. and D.Z..; funding acquisition, E.V. All authors have read and agreed to the published version of the manuscript

Funding: The research was carried out within the state assignment of the Ministry of Science and Higher Education of the Russian Federation for the Research Centre for Medical Genetics.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: The original contributions presented in this study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

iPSCs	Induced pluripotent stem cells
asFS	Aphidicolin-sensitive fragile sites
MiDAS	Mitotic DNA synthesis
MORs	Minimal overlapping regions
RS	Replication stress
RFS	Rare fragile site
CFS	Common fragile site
ERFS	Early-replicating fragile site
ISCN	International System for Human Cytogenetic Nomenclature
EdU	5-ethynyl-2′-deoxyuridine

References

1. Stavish D, Price CJ, Gelezauskaite G, Alsehli H, Leonhard KA, Taapken SM, et al. Feeder-free culture of human pluripotent stem cells drives MDM4-mediated gain of chromosome 1q. Stem Cell Rep. 2024;19:1217–32. https://doi.org/10.1016/j.stemcr.2024.06.003.
2. Assou S, Girault N, Plinet M, Bouckenheimer J, Sansac C, Combe M, et al. Recurrent Genetic Abnormalities in Human Pluripotent Stem Cells: Definition and Routine Detection in Culture Supernatant by Targeted Droplet Digital PCR. Stem Cell Rep. 2020;14:1–8. https://doi.org/10.1016/j.stemcr.2019.12.004.
3. Jo H-Y, Lee Y, Ahn H, Han H-J, Kwon A, Kim B-Y, et al. Functional in vivo and in vitro effects of 20q11.21 genetic aberrations on hPSC differentiation. Sci Rep. 2020;10:18582. https://doi.org/10.1038/s41598-020-75657-7.
4. Krivec N, De Deckersberg EC, Lei Y, Delbany DA, Regin M, Verhulst S, et al. Gain of 1q confers an MDM4-driven growth advantage to undifferentiated and differentiating hESC while altering their differentiation capacity. Cell Death Dis. 2024;15:852. https://doi.org/10.1038/s41419-024-07236-x.
5. Lei Y, Duong MC, Krivec N, Janssens C, Regin M, Huyghebaert A, et al. Loss of 18q Alters TGFB Signalling Affecting Anteroposterior Neuroectodermal Fate in Human Embryonic Stem Cells. Cell Prolif. 2025;:e13813. https://doi.org/10.1111/cpr.13813.
6. Markouli C, Couvreu De Deckersberg E, Regin M, Nguyen HT, Zambelli F, Keller A, et al. Gain of 20q11.21 in Human Pluripotent Stem Cells Impairs TGF-β-Dependent Neuroectodermal Commitment. Stem Cell Rep. 2019;13:163–76. https://doi.org/10.1016/j.stemcr.2019.05.005.
7. Vales JP, Barbaric I. Culture-acquired genetic variation in human pluripotent stem cells: Twenty years on. BioEssays. 2024;46:2400062. https://doi.org/10.1002/bies.202400062.
8. Ahuja AK, Jodkowska K, Teloni F, Bizard AH, Zellweger R, Herrador R, et al. A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells. Nat Commun. 2016;7:10660. https://doi.org/10.1038/ncomms10660.

 doi:10.20944/preprints202607.1652.v1

9. Halliwell JA, Frith TJR, Laing O, Price CJ, Bower OJ, Stavish D, et al. Nucleosides Rescue Replication-Mediated Genome Instability of Human Pluripotent Stem Cells. Stem Cell Rep. 2020;14:1009–17. https://doi.org/10.1016/j.stemcr.2020.04.004.
10. Ruiz S, Lopez-Contreras AJ, Gabut M, Marion RM, Gutierrez-Martinez P, Bua S, et al. Limiting replication stress during somatic cell reprogramming reduces genomic instability in induced pluripotent stem cells. Nat Commun. 2015;6:8036. https://doi.org/10.1038/ncomms9036.
11. Vallabhaneni H, Lynch PJ, Chen G, Park K, Liu Y, Goehe R, et al. High Basal Levels of γH2AX in Human Induced Pluripotent Stem Cells Are Linked to Replication-Associated DNA Damage and Repair. Stem Cells. 2018;36:1501–13. https://doi.org/10.1002/stem.2861.
12. Milagre I, Pereira C, Oliveira RA. Compromised Mitotic Fidelity in Human Pluripotent Stem Cells. Int J Mol Sci. 2023;24:11933. https://doi.org/10.3390/ijms241511933.
13. Waisertreiger I, Popovich K, Block M, Anderson KR, Barlow JH. Visualizing locus-specific sister chromatid exchange reveals differential patterns of replication stress-induced fragile site breakage. Oncogene. 2020;39:1260–72. https://doi.org/10.1038/s41388-019-1054-5.
14. Barlow JH, Faryabi RB, Callén E, Wong N, Malhowski A, Chen HT, et al. Identification of Early Replicating Fragile Sites that Contribute to Genome Instability. Cell. 2013;152:620–32. https://doi.org/10.1016/j.cell.2013.01.006.
15. Scelfo A, Angrisani A, Grillo M, Barnes BM, Muyas F, Sauer CM, et al. Specialized replication mechanisms maintain genome stability at human centromeres. Mol Cell. 2024;84:1003-1020.e10. https://doi.org/10.1016/j.molcel.2024.01.018.
16. Sfeir A, Kosiyatrakul ST, Hockemeyer D, MacRae SL, Karlseder J, Schildkraut CL, et al. Mammalian Telomeres Resemble Fragile Sites and Require TRF1 for Efficient Replication. Cell. 2009;138:90–103. https://doi.org/10.1016/j.cell.2009.06.021.
17. Debatisse M, Rosselli F. A journey with common fragile sites: From S phase to telophase. Genes Chromosomes Cancer. 2019;58:305–16. https://doi.org/10.1002/gcc.22704.
18. Bhowmick R, Minocherhomji S, Hickson ID. RAD52 Facilitates Mitotic DNA Synthesis Following Replication Stress. Mol Cell. 2016;64:1117–26. https://doi.org/10.1016/j.molcel.2016.10.037.
19. Minocherhomji S, Ying S, Bjerregaard VA, Bursomanno S, Aleliunaite A, Wu W, et al. Replication stress activates DNA repair synthesis in mitosis. Nature. 2015;528:286–90. https://doi.org/10.1038/nature16139.
20. Le Tallec B, Millot GA, Blin ME, Brison O, Dutrillaux B, Debatisse M. Common Fragile Site Profiling in Epithelial and Erythroid Cells Reveals that Most Recurrent Cancer Deletions Lie in Fragile Sites Hosting Large Genes. Cell Rep. 2013;4:420–8. https://doi.org/10.1016/j.celrep.2013.07.003.
21. Letessier A, Millot GA, Koundrioukoff S, Lachagès A-M, Vogt N, Hansen RS, et al. Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. Nature. 2011;470:120–3. https://doi.org/10.1038/nature09745.
22. Glover TW, Berger C, Coyle J, Echo B. DNA polymerase ? inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum Genet. 1984;67:136–42. https://doi.org/10.1007/BF00272988.
23. Weise. Global screening and extended nomenclature for 230 aphidicolin-inducible fragile sites, including 61 yet unreported ones. Int J Oncol. 2010;36. https://doi.org/10.3892/ijo_00000572.
24. Balzano E, Di Tommaso E, Antoccia A, Pelliccia F, Giunta S. Characterization of Chromosomal Instability in Glioblastoma. Front Genet. 2022;12:810793. https://doi.org/10.3389/fgene.2021.810793.
25. Miron K, Golan-Lev T, Dvir R, Ben-David E, Kerem B. Oncogenes create a unique landscape of fragile sites. Nat Commun. 2015;6:7094. https://doi.org/10.1038/ncomms8094.
26. Vallabhaneni H, Hursh DA. Replication-associated DNA damage in induced pluripotent stem cells. In: Current Topics in iPSCs Technology. Elsevier; 2022. p. 177–96. https://doi.org/10.1016/B978-0-323-99892-5.00023-2.
27. Okamoto Y, Iwasaki WM, Kugou K, Takahashi KK, Oda A, Sato K, et al. Replication stress induces accumulation of FANCD2 at central region of large fragile genes. Nucleic Acids Res. 2018;46:2932–44. https://doi.org/10.1093/nar/gky058.

 doi:10.20944/preprints202607.1652.v1

28. Pentzold C, Shah SA, Hansen NR, Le Tallec B, Seguin-Orlando A, Debatisse M, et al. FANCD2 binding identifies conserved fragile sites at large transcribed genes in avian cells. Nucleic Acids Res. 2018;46:1280–94. https://doi.org/10.1093/nar/gkx1260.
29. Brison O, El-Hilali S, Azar D, Koundrioukoff S, Schmidt M, Nähse V, et al. Transcription-mediated organization of the replication initiation program across large genes sets common fragile sites genome-wide. Nat Commun. 2019;10:5693. https://doi.org/10.1038/s41467-019-13674-5.
30. Sarni D, Sasaki T, Irony Tur-Sinai M, Miron K, Rivera-Mulia JC, Magnuson B, et al. 3D genome organization contributes to genome instability at fragile sites. Nat Commun. 2020;11:3613. https://doi.org/10.1038/s41467-020-17448-2.
31. Ji F, Liao H, Pan S, Ouyang L, Jia F, Fu Z, et al. Genome-wide high-resolution mapping of mitotic DNA synthesis sites and common fragile sites by direct sequencing. Cell Res. 2020;30:1009–23. https://doi.org/10.1038/s41422-020-0357-y.
32. Dong Y, Qiu M, Tang H, Shi W, Lu Y, Ji F, et al. Early replication fragile sites are associated with cancer-related CNVs and SNVs in human embryonic stem cells. Stem Cell Rep. 2026;:102968. https://doi.org/10.1016/j.stemcr.2026.102968.
33. Kislova AV, Zheglo D, Pozhitnova VO, Sviridov PS, Gadzhieva EP, Voronina ES. Replication stress causes delayed mitotic entry and chromosome 12 fragility at the ANKS1B large neuronal gene in human induced pluripotent stem cells. Chromosome Res. 2023;31:23. https://doi.org/10.1007/s10577-023-09729-5.
34. Mehrjardi NZ, Molcanyi M, Hatay FF, Timmer M, Shahbazi E, Ackermann JP, et al. Acquisition of chromosome 1q duplication in parental and genome-edited human-induced pluripotent stem cell-derived neural stem cells results in their higher proliferation rate in vitro and in vivo. Cell Prolif. 2020;53:e12892. https://doi.org/10.1111/cpr.12892.
35. Varela C, Denis JA, Polentes J, Feyeux M, Aubert S, Champon B, et al. Recurrent genomic instability of chromosome 1q in neural derivatives of human embryonic stem cells. J Clin Invest. 2012;122:569–74. https://doi.org/10.1172/JCI46268.
36. Zheglo D, Pozhitnova VO, Kislova AV, Markova ZG, Kiselev D, Sviridov PS, et al. Chromosomal Aberrations in Induced Pluripotent Stem Cells: Identification of Breakpoints in the Large DCC Gene and HIST2 Histone Gene Cluster. Int J Mol Sci. 2025;26:7728. https://doi.org/10.3390/ijms26167728.
37. Avery S, Hirst AJ, Baker D, Lim CY, Alagaratnam S, Skotheim RI, et al. BCL-XL Mediates the Strong Selective Advantage of a 20q11.21 Amplification Commonly Found in Human Embryonic Stem Cell Cultures. Stem Cell Rep. 2013;1:379–86. https://doi.org/10.1016/j.stemcr.2013.10.005.
38. Halliwell JA, Baker D, Judge K, Quail MA, Oliver K, Betteridge E, et al. Nanopore Sequencing Indicates That Tandem Amplification of Chromosome 20q11.21 in Human Pluripotent Stem Cells Is Driven by Break-Induced Replication. Stem Cells Dev. 2021;30:578–86. https://doi.org/10.1089/scd.2021.0013.
39. Krivec N, Ghosh MS, Spits C. Gains of 20q11.21 in human pluripotent stem cells: Insights from cancer research. Stem Cell Rep. 2024;19:11–27. https://doi.org/10.1016/j.stemcr.2023.11.013.
40. Keller A, Lei Y, Krivec N, De Deckersberg EC, Dziedzicka D, Markouli C, et al. Gains of 12p13.31 delay WNT-mediated initiation of hPSC differentiation and promote residual pluripotency in a cell cycle dependent manner. 2021. https://doi.org/10.1101/2021.05.22.445238.
41. Lee C-T, Bendriem RM, Kindberg AA, Worden LT, Williams MP, Drgon T, et al. Functional Consequences of 17q21.31/WNT3-WNT9B Amplification in hPSCs with Respect to Neural Differentiation. Cell Rep. 2015;10:616–32. https://doi.org/10.1016/j.celrep.2014.12.050.