



## RESEARCH HIGHLIGHT

# Twin peaks: finding fragile sites with MiDAS-seq

Thomas W. Glover<sup>1,2</sup> and Thomas E. Wilson<sup>1,2</sup>

Cell Research (2020) 30:944–945; <https://doi.org/10.1038/s41422-020-0376-8>

**Two papers from Macheret et al. and Ji et al. describe novel high-resolution direct sequencing approaches to map fragile sites and hundreds of additional genomic regions that can remain under-replicated prior to mitotic entry and complete replication in mitosis. They further establish many defining properties of these loci that greatly add to our mechanistic understanding of fragile sites and genome instability following replication stress.**

Partial inhibition of DNA replication creates replication stress, which in turn promotes genome instability. Common fragile sites (CFSs) are genomic loci that are especially prone to this instability.<sup>1</sup> CFSs form visible gaps and breaks on metaphase chromosomes under conditions that perturb DNA synthesis, such as treatment with low concentrations of the DNA polymerase inhibitor aphidicolin. Given their sensitivity to impaired DNA synthesis, CFSs in cultured cells have been widely utilized as signatures of replication stress. Current models for CFS instability posit that replication forks progressing inward into origin-poor CFSs, frequently dictated by active transcription of very large genes, stall and fail to replicate the DNA between them in S phase.<sup>2</sup> Replication is completed in mitosis (M phase) by mitotic DNA synthesis (MiDAS), a POLD3- and RAD52-dependent process that shares features with break-induced DNA replication (BIR).<sup>3</sup> Importantly, chromosome breaks and gaps are but one manifestation of CFS instability. Misrepair of CFS lesions can lead to chromosome rearrangements, most notably copy number variations that are structurally equivalent to the copy number alterations seen in cancers.<sup>4</sup> In addition, CFS genes are top hits in experiments that detect double-strand break-mediated translocations in cultured neural progenitors.<sup>5</sup> Unreplicated DNA at CFSs that persists into late mitosis can also lead to ultrafine anaphase bridges and chromosome mis-segregation.<sup>6</sup>

CFSs were first identified over 35 years ago using classic cytogenetics and now over 70 have been described.<sup>2</sup> The majority of fragile site gaps and breaks occur at ~10–15 of the most sensitive and unstable loci in any cell type. Others are less frequent with no clear consensus about what differentiates a CFS from a random gap or break. Only a small number of CFSs have been mapped at high resolution by fluorescence in situ hybridization; most are defined at the chromosome band level at times leading to uncertain correlations with other molecular and cellular events.

Macheret et al.<sup>7</sup> and Ji et al.<sup>8</sup> address these issues using novel high-resolution sequencing approaches to locate all regions that undergo replication in mitosis under replication stress. These methods, termed EdU-Seq or, more specifically, MiDAS-Seq, involve EdU incorporation into DNA selectively in M-phase cells

followed by sequencing of the labeled DNA. Both groups studied U2OS osteosarcoma cells, HeLa cells and normal human fibroblasts or epithelial cells. The MiDAS regions they identified encompassed all of the 73 known CFSs and hundreds of other CFS-like loci. There were over 250 MiDAS peaks in U2OS cells in both studies, 85–206 peaks in HeLa cells, and considerably fewer in normal cells with fibroblasts showing only 36 MiDAS regions. They ranged in size from 0.5 to 1.2 Mb in the Macheret study and 0.1 to 2.3 Mb by Ji et al. The regions overlapped substantially between the cell lines but with cell type differences that likely reflect different transcription profiles and the propensity of different cell lines to progress to M phase with unreplicated DNA.

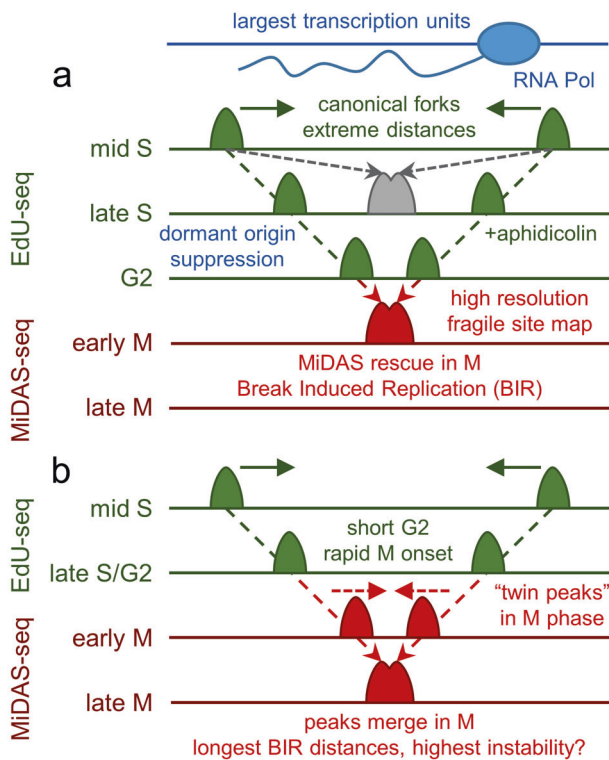
Known CFSs display a number of characteristics, including late replication, enrichment in transcribed large genes, cell type specificity, a paucity of replication origins and completion of replication in mitosis.<sup>3,4,9,10</sup> These properties were shared by the great majority of the loci identified by MiDAS-seq. Macheret et al. found that two-thirds of MiDAS regions mapped to single large genes > 630 kb, with some mapping to two adjacent genes. One-third mapped to intergenic regions that were nevertheless actively transcribed, with strong concordance between the size of the MiDAS regions and the length of the transcribed domains. These results reinforce a central mechanistic importance of transcription over long genomic distances in leading to incomplete replication in S phase, as previously shown for many known CFSs.<sup>4</sup>

One presumptive mechanism for the transcription effect is suggested by both studies wherein MiDAS regions consistently mapped to replication domains with a paucity of active and dormant origins, including a new demonstration of this phenomenon in RIF1-depleted cells by Macheret et al. A cause and effect relationship remains to be established, but the notion that origin paucity is secondary to origin suppression by transcription was supported by a multiple logistic regression by Ji et al. where only large transcription units and late replication were independent predictors of MiDAS.

Interestingly, a “twin-peak” EdU signal pattern was frequently observed at MiDAS sites that merged into a single peak as M phase progressed, especially at known CFSs (Fig. 1). Macheret et al. found that the median sizes of the twin-peak regions in U2OS cells ranged between 1.1 and 1.2 Mb, whereas single-peak regions in all cells were considerably smaller, similar to the results reported by Ji et al. Importantly, only the cancer cell lines showed twin peaks, whereas normal cells did not. Moreover, MiDAS regions were specific to samples subjected to replication stress. These patterns provide strong support for models of dual failure of two replication forks converging towards the centers of large

<sup>1</sup>Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109, USA and <sup>2</sup>Department of Pathology, University of Michigan Medical School, Ann Arbor, MI 48109, USA

Correspondence: Thomas W. Glover (glover@med.umich.edu)



**Fig. 1 EdU-seq in S, G2, and M phases reveals the dynamics of fragile site replication.** Panels show idealized signal peaks seen at the largest human genes associated with CFSs when EdU-seq is restricted to different cell cycle stages. **a** Transcription enforces extremely large replicons by suppressing internal origin firing, such that MiDAS (red color) is required to rescue replication in M phase when fork progression is impeded (even large genes usually complete replication in S phase without replication stress; gray color). **b** In cell types with less effective DNA damage responses, more rapid progression to M phase leads to the appearance of “twin peaks” of MiDAS signal that highlight the long distances traveled by the non-canonical low-fidelity forks associated with BIR.

genes,<sup>4,10</sup> wherein the extent of M-phase replication is determined by fork speed, their likelihood of failing, the distance they travel, and the time afforded to resolve replicative lesions in S/G2 by DNA damage responses.

MiDAS has been hypothesized to occur as a form of BIR, a conservative DNA replication process initiated at single-ended DNA breaks.<sup>3</sup> This atypical nature of replication by MiDAS was reinforced by Macheret et al. They found strand asymmetry in aligned sequence reads in HeLa cells suggesting uncoupling of leading and lagging strand synthesis in the two-stage BIR mechanism. Thus, under certain circumstances, potentially large spans of genomic DNA must be replicated at CFSs by an inefficient low-fidelity mechanism (Fig. 1).

In summary, MiDAS-seq has provided new high-density data that support previous hypotheses to explain CFS fragility, including the contributions of late replication timing, large transcription units and BIR acting in mitosis. Hundreds of additional CFS-like regions were identified that share these characteristics, yielding a high-resolution map of most known and potential human CFSs. A number of new questions now arise. Are all identified MiDAS regions at high risk for instability similar to known CFSs? Is MiDAS an error-free process required to suppress genomic rearrangements at CFSs or an error-prone rescue process that creates such mutations? Which cell types in normal tissues and during oncogenesis are most dependent on MiDAS vs replication resolution in S/G2? Answers to these questions will have important implications for genome instability after replication stress in normal development and disease.

## REFERENCES

- Glover, T. W., Berger, C., Coyle, J. & Echo, B. *Hum. Genet.* **67**, 136–142 (1984).
- Glover, T. W., Wilson, T. E. & Arlt, M. F. *Nat. Rev. Cancer* **17**, 489–501 (2017).
- Minocherhomji, S. et al. *Nature* **528**, 286–290 (2015).
- Wilson, T. E. et al. *Genome Res.* **25**, 189–200 (2015).
- Wei, P. C. et al. *Cell* **164**, 644–655 (2016).
- Chan, K. L., Palmari-Pallag, T., Ying, S. & Hickson, I. D. *Nat. Cell Biol.* **11**, 753–760 (2009).
- Macheret, M. et al. *Cell Res.* <https://doi.org/10.1038/s41422-020-0358-x> (2020).
- Ji, F. et al. *Cell Res.* <https://doi.org/10.1038/s41422-020-0357-y> (2020).
- Le Tallec, B. et al. *Nat. Struct. Mol. Biol.* **18**, 1421–1423 (2011).
- Letessier, A. et al. *Nature* **470**, 120–123 (2011).