English title

Genome-wide study of DNA replication and genome instability of human cells under both normal and stressed conditions

Keywords: DNA Replication; Genome instability; Replication stress; Single-molecule/cell approaches

Summary of thesis project

It is well known that DNA replication is a vital process in all living organisms. At each cell division, the activation of over 30,000 replication origins in a coordinated manner is essential to ensure the duplication of >6 billion base pairs of the human genome. During differentiation and development, this program must adapt to changes in chromatin organization and gene transcription. Its deregulation can challenge genome stability, which is a leading cause of many diseases including cancer and neurological disorders. However, despite intensive studies, the mechanisms that coordinate where and when replication initiates in the human genome remain poorly known.

Our team at Institut Curie (Paris, France) focuses on developing cutting-edge high-throughput genomic approaches and using genome-wide data analyses to study the spatio-temporal replication program of the human genome, and how its deregulation causes genome instability and leads to human diseases. Recently, we have published several pioneer papers showing how gene transcription as well as chromatin organization shape the replication landscape of human cells (*1–4*), and how transcription-replication conflicts lead to genome instability (*5–8*), which is amongst the key factors playing an important role in cancer development. Importantly, origin usage in human cells is remarkably dynamic and flexible. This is likely an important mechanism to prevent genome instability. However, studying the degree of such stochastic variation is challenging. Therefore, in addition to population-based methods, we have also developed new genome-wide approaches to study DNA replication at the single−molecule or single-cell level (*2*, *9*, *10*). Accessing to this level is critical given the inefficient and heterogeneous nature of mammalian origin firing.

Importantly, RIF1 (Rap1-interacting factor 1) has been recently identified as a key factor in controlling genome-wide replication timing (RT) from yeasts, fly, mouse to human. Remarkably, Rif1 Knockout (KO) human or mouse embryonic stem cells show the most severe phenotype with a nearly random RT program (i.e. flat population RT profile). Analysis of single-cell replication timing (scRT) supports that Rif1 KO disrupts RT by dramatically increasing cell-to-cell RT heterogeneity and genome instability within the cell population, while how Rif1 depletion affects genome-wide replication initiation program has not been studied so far. To address this important question, both team of Prof. Masato T. Kanemaki (NIG, University of Tokyo, Japan) and Chen's team at Institut Curie have recently applied complementary approaches to study the detailed replication initiation program in RIF1 depleted cells (unpublished results). By applying LD-OK-seq on Rif1 KO HCT116 cells, Kanemaki's team has mapped the replication fork direction and replication initiation at the population level, and by applying Optical Replication Mapping (ORM) on AID2-RIF-knockdown (KD) cells, Chen's team has mapped individual replication initiation events at the single-molecule level. However, the detailed mechanism on how RIF1 controls replication program and how its deregulation leads to genome instability are still unclear and need to be further investigated.

In this project in collaboration with Prof. Kanemaki, the Ph.D. candidate will combine molecular and cell biological, structural, and genomics approaches to elucidate the mechanisms regulating DNA replication program under both normal and stressed conditions in human cells. This innovative interdisciplinary research project will generate important novel insights into the molecular mechanisms of DNA replication initiation control in human cells, providing a functional framework for further understanding diseases associated with mutations of genes involved in DNA replication regulation, which are often manifested as developmental defects, such as dwarfism as well as tumorigenesis.

Objectifs: Using cutting-edge high-throughput genomic approaches recently developed by both teams (both experimental and computational methods) to study DNA replication program and genome instability in human cells under normal growth as well as under various replication stresses and mutants/depletions of key factors in replication initiation control (e.g. RIF1 and its partners), in order to gain new insights into the molecular mechanisms of DNA replication initiation control in human cells and how its deregulation leads to genome instability and human diseases.

Host team (Explain feasibility, expertise of the host team, availability of equipment, etc.)

The team leader, Chunlong Chen (ATIP-Avenir laureate in 2017, and Impulscience laureate in 2022) is is one of the world-wide leaders in the field. Combining both wet and dry lab, the team has strong expertise in developing new genome-wide approaches and bioinformatics analyses to study the replication program of the human genome and how its deregulation impacts on genome instability. We have published one of the first high-resolution replication timing profiles of the human genome by Repli-Seq (*Genome Res.* 2010) and a series of original studies on the impact of replication program in organization and mutational landscape of the human genome (*MBE* 2011, *PLoS Comput. Biol.* 2011, *Nat. Protoc.* 2013, *Euro. Urology* 2019, *Genome Medicine* 2020, etc.). More recently, the team has published a new genomic approach obtaining completely novel data on the human replication landscape based on the purification, sequencing and bioinformatics analysis of the Okazaki fragments (OK-Seq, *Nat. Commun.* 2016, *NAR* 2023, *Nat. Protocol* 2023). By using Repli-Seq and OK-Seq data, we have further revealed that gene transcription plays an important role in shaping DNA replication landscape, and transcription-mediated replication stresses induce genome instability at both early *(Nat. Commun.* 2020) and late replicating regions (*Nat. Commun.* 2019, *NSMB* 2023). With the support of the Impulscience project of the Fondation Bettencourt Schueller, we have developed original interdisciplinary methodologies by combining cutting-edge high-throughput single molecule/cell imaging/sequencing (*Mol Cell* 2021, *Nat Commun* 2022, 2025), mathematical/computational modelling and bioinformatics analyses to study the replication program in mammalian cell lines and developmental model organisms under normal and pathological conditions.

Institut Curie is a prestigious institution with its reputation as a foundation for cancer research. In terms of infrastructure, we have all the necessary equipment and computational resources for the project, including a high-performance computational cluster, the NGS, Genomics, single-cell, animal and bioinformatic facilities. The genome-wide data will be managed by the team at I. Curie that has a long team experience and all the informatic equipment necessary to store and analyze large amounts of data. We will set up an adapted data management plan with the help of the Direction of Data of I. Curie based on the Guidelines on FAIR Data Management in Horizon 2020.

Collaborations: The team has established secussfull collaborations with world-wide experts in the field of DNA replicaiton and genome instbility (such as P. Pasero, D.M. Gilbert, Y.L. Zhai, etc.). This internationial Ph.D. project will be developped under a co-supervision with Prof. M.T. Kanemaki (University of Tokyo), whom we have established successful collaboration with join publication (Emerson *Nature* 2022). The team of Pr. M.T. Kanemaki at National Institute of Genetics is a pioneer in the field in developing the innovative auxin-inducible degron (AID) technology to understand the mechanisms of DNA replication, from yeast to human cells and mice. Based on the complementary expertise of Pr. Kanemaki and Dr. Chen, together with the large numbers of preliminary data obtained by both teams, we are confident to carry out this new innovated interdisciplinary research project.

References

1. N. Petryk, et al., Replication landscape of the human genome. *Nature Communications* **7**, 10208 (2016).
2. W. Wang, et al., Genome-wide mapping of human DNA replication by optical replication mapping supports a stochastic model of eukaryotic replication. *Mol Cell* **81**, 2975-2988.e6 (2021).
3. Y. Liu, et al., C.-L. Chen, OKseqHMM: a genome-wide replication fork directionality analysis toolkit. *Nucleic Acids Research* **51**, e22–e22 (2023).
4. D. J. Emerson, et al., Cohesin-mediated loop anchors confine the locations of human replication origins. *Nature* **606**, 812–819 (2022).
5. O. Brison, et al., C. L. Chen, Transcription-mediated organization of the replication initiation program across large genes sets common fragile sites genome-wide. *Nature Communications* **10**, 5693 (2019).
6. A. Promonet, et al., Topoisomerase 1 prevents replication stress at R-loop-enriched transcription termination sites. *Nature Communications* **11**, 3940 (2020).
7. S. Gnan, Y. Liu, M. Spagnuolo, C.-L. Chen, The impact of transcription-mediated replication stress on genome instability and human disease. *Genome Instability & Disease* **1**, 207–234 (2020).
8. O. Brison, et al., C.-L. Chen, M. Debatisse, Mistimed origin licensing and activation stabilize common fragile sites under tight DNA-replication checkpoint activation. *Nature Structural & Molecular Biology*, 1–12 (2023).
9. S. Gnan, et al., C. L. Chen, Kronos scRT: a uniform framework for single-cell replication timing analysis. *Nat Commun* **13**, 2329 (2022).
10. J. M. Josephides, C.-L. Chen, Unravelling single-cell DNA replication timing dynamics using machine learning reveals heterogeneity in cancer progression. *Nat Commun* **16**, 1472 (2025).