Replication initiation is associated to divergent promoter regions in Schizosaccharomyces pombe genome

Koji Masuda[1], Claire Renard-Guillet[1], Takashi Sutani[1], Ryuichiro Nakato[1], Yuki Katou[1], Atsunori Yoshimura[1], Yutaka Kanoh[2], Hisao Masai[2], Katsuhiko Shirahige[1]

1 Institute of Molecular and Cellular Biosciences, The University of Tokyo,
1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan

2 Department of Genome Medicine, Tokyo Metropolitan Institute of Medical Science, Japan

DNA replication is a key process for cells, ensuring that the whole DNA is copied at each cell cycle. In eukaryotes, DNA replication initiates at multiple sites on a chromosome, called replication origins (ORI). During G1 phase, the pre-replication complexes (pre-RC) assemble at several points along the genome, and are then potentially licensed to fire as ORI. However, which element(s) determine the positioning of pre-RC is still an open question. A recent study support the model in which ORI may correspond to any nucleosome-free regions with high affinity for pre-RC proteins (Xu et al 2012). Complementarily, strong links were found, for a part of ORI, between ORI positions and firing times and active transcription initiation (review in Mechali 2010). It was even suggested that all ORI might be associated to TSS, active or not (Dellino et al. 2013). Yet, if the existence of a relationship between ORI and transcription is well established now, what determines this association remains unclear.

One hypothesis could be that ori are positioned so that the smooth progression of replication forks with regards to transcription process would be favored. Indeed, it was shown that in bacteria, whose DNA replication initiates at a single ORI, there was co-directionality between replication and transcription, presumably to avoid frontal collision between replication and transcription machineries (review in Lin and Pasero 2012). However, whether such co-directional bias exists also in eukaryotes is under question. If the deleterious effect of a head-on collision between transcription machinery and replication fork is well established (review in Lin and Pasero 2012), studies in budding yeast failed to find correlation on a genome-wide scale between ori position and gene orientation (*e.g.* Azvolinsky et al. 2009). In human, it was proposed to be the case, based on *in-silico* predicted ORI positions (Huvet et al 2007), but such genomic organization was not detected with an experimentally detected ORI dataset (Necsulea et al 2009).

In this work, we focused on ORI in the fission yeast (*Schizosaccharomyces pombe*) genome, a compact genome (12.5 Mb) but whose ORI have a structure similar to metazoan ones. We examine here, for the first time, the relationship between transcription and replication initiation with high accuracy.

We define a dataset of 395 ORI, as the overlap of the Mcm4 binding sites defined in (Hayano et al. 2012) and oris defined in (Hayashi et al. 2007) (both are ChIP-on-chip datasets). Timing of ORI firing is define using the BrdU dataset presented in (Hayano et al. 2012) ; ORI with top 50% highest BrdU ratio are considered as active, the others are defined as late or dormant. We also compare these replication data to an unpublished dataset of expression data (RNA-seq) in early S-phase, to have information not only on ORF position but also on their transcriptional activity.

We observe that a large part of ORI (77%) localize in intergenic regions. This distribution is highly biased since intergenic regions represent only 18% of the fission yeast genome.

We then focus on gene pairs immediately flanking intergenic ORI (ORI-flanking genes; figure 1A). Half of ORI are flanked by divergent gene pairs, which is significantly higher than the proportion of divergent pairs amongst all consecutive genes (figure 1B). This could support the hypothesis of co-orientation of transcription and replication; however this orientation is not conserved when considering the following genes. Since only 8% of ORI are in between convergent gene pairs, we can confirm the model in which (almost) all ORI are associated to transcription initiation sites, even with a preference for double transcription initiation. This finding suggests that the key point for proper genomic

replication might be an association between ORI and promoters more than a co-directionality of transcription and fork progression.

Active promoters are open-chromatin regions, and one can expect an association of ORI with promoters of highly transcribed genes. When comparing expression of divergent or uni-directional ORI-flanking genes, we do not observe expression levels significantly higher (figure 1C), However, all of these genes are transcribed (RPKM higher than 1). Interestingly, we do not observe difference in transcription level between genes flanking active or late ORI.

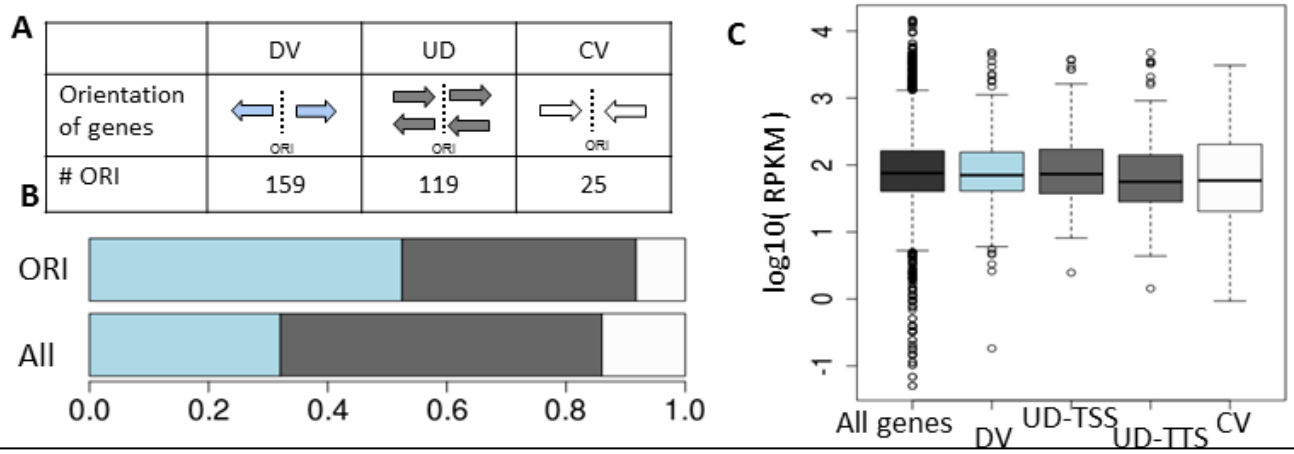

Figure 1: Proportion of the type of orientation of ORI-flanking genes, and their transcription level. A. Number of ORI-flanking gene pairs in divergent (DV), uni-directional (UD) or convergent (CV) orientation. B. Proportion of DV, UD and CV flanking gene pairs for ORI and all intergenic regions. C. Distribution of expression (RPKM) for ORI-flanking genes according to their orientation: DV, UD and CV. UD orientated genes are further separated in genes with TSS towards ORI (UD-TSS) or TTS (US-TTS). Distribution of expression for all genes is also shown. Width of box plot is related to the number of genes ; middle line indicates median and the limit of the 'box' are first and third quartile.

Our data then indicate that ORI are associated to active promoters, whatever the level of expression. However a large part of genes seem to be transcribed with no association to ORI; further analyses are then currently performed to determine whether particular features of intergenic regions (nucleosome profile, positioning of RNA polymeraseII…) may influence pre-RC positioning along the genome.

Bibliography

Azvolinsky, A.; Giresi, P. G.; Lieb, J. D. & Zakian, V. A., Highly transcribed RNA polymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae., *Mol Cell*, **2009**, *34*, 722-734

Dellino, G. I.; Cittaro, D.; Piccioni, R.; Luzi, L.; Banfi, S.; Segalla, S.; Cesaroni, M.; Mendoza-Maldonado, R.; Giacca, M. & Pelicci, P. G., Genome-wide mapping of human DNA-replication origins: levels of transcription at ORC1 sites regulate origin selection and replication timing., *Genome Res*, **2013**, *23*, 1-11

Hayano, M.; Kanoh, Y.; Matsumoto, S.; Renard-Guillet, C.; Shirahige, K. & Masai, H., Rif1 is a global regulator of timing of replication origin firing in fission yeast.,*Genes Dev*, **2012**, *26*, 137-150

Hayashi, M.; Katou, Y.; Itoh, T.; Tazumi, A.; Tazumi, M.; Yamada, Y.; Takahashi, T.; Nakagawa, T.; Shirahige, K. & Masukata, H.Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast.,*EMBO J*, **2007**, *26*, 1327-1339

Huvet, M.; Nicolay, S.; Touchon, M.; Audit, B.; d'Aubenton Carafa, Y.; Arneodo, A. & Thermes, C.,Human gene organization driven by the coordination of replication and transcription.,*Genome Res*, **2007**, *17*, 1278-1285

Lin, Y.-L. & Pasero, P., Interference between DNA replication and transcription as a cause of genomic instability., *Curr Genomics*, **2012**, *13*, 65-73

Mechali, M., Eukaryotic DNA replication origins: many choices for appropriate answers., *Nat Rev Mol Cell Biol*, **2010**, *11*, 728-738

Necsulea, A.; Guillet, C.; Cadoret, J.-C.; Prioleau, M.-N. & Duret, L., The relationship between DNA replication and human genome organization., *Mol Biol Evol*, **2009**, *26*, 729-741

Xu, J.; Yanagisawa, Y.; Tsankov, A. M.; Hart, C.; Aoki, K.; Kommajosyula, N.; Steinmann, K. E.; Bochicchio, J.; Russ, C.; Regev, A.; Rando, O. J.; Nusbaum, C.; Niki, H.; Milos, P.; Weng, Z. & Rhind, N., Genome-wide identification and characterization of replication origins by deep sequencing.,*Genome Biol*, **2012**, *13*, R27