

Ligand-induced native G-quadruplex stabilization impairs transcription initiation

Conghui Li,^{1,7} Honghong Wang,^{1,7} Zhinang Yin,^{1,7} Pingping Fang,^{2,7} Ruijing Xiao,^{1,3} Ying Xiang,¹ Wen Wang,¹ Qiuzi Li,⁴ Beili Huang,¹ Jian Huang,⁴ and Kaiwei Liang^{1,5,6}

¹Department of Pathophysiology, School of Basic Medical Sciences, Wuhan University, Wuhan 430071, China; ²Department of Pharmacology, School of Basic Medical Sciences, Wuhan University, Wuhan 430071, China; ³Department of Immunology, School of Basic Medical Sciences, Wuhan University, Wuhan 430071, China; ⁴College of Life Sciences, Wuhan University, Wuhan 430071, China; ⁵Hubei Province Key Laboratory of Allergy and Immunology, School of Basic Medical Sciences, Wuhan University, Wuhan 430071, China; ⁶Research Center for Medicine and Structural Biology, School of Basic Medical Sciences, Wuhan University, Wuhan 430071, China

G-quadruplexes (G4s) are noncanonical DNA secondary structures formed through the self-association of guanines, and G4s are distributed widely across the genome. G4 participates in multiple biological processes including gene transcription, and G4-targeted ligands serve as potential therapeutic agents for DNA-targeted therapies. However, genome-wide studies of the exact roles of G4s in transcriptional regulation are still lacking. Here, we establish a sensitive G4-CUT&Tag method for genome-wide profiling of native G4s with high resolution and specificity. We find that native G4 signals are cell type-specific and are associated with transcriptional regulatory elements carrying active epigenetic modifications. Drug-induced promoter-proximal RNA polymerase II pausing promotes nearby G4 formation. In contrast, G4 stabilization by G4-targeted ligands globally reduces RNA polymerase II occupancy at gene promoters as well as nascent RNA synthesis. Moreover, ligand-induced G4 stabilization modulates chromatin states and impedes transcription initiation via inhibition of general transcription factors loading to promoters. Together, our study reveals a reciprocal genome-wide regulation between native G4 dynamics and gene transcription, which will deepen our understanding of G4 biology toward therapeutically targeting G4s in human diseases.

[Supplemental material is available for this article.]

G-quadruplexes (G4s) are four-stranded intramolecular structures that arise from the self-stacking of two or more guanine quartets (G-quartets), in which the four guanine molecules form a square planar arrangement in a cyclic hydrogen-bonding pattern (Bochman et al. 2012; Varshney et al. 2020). G4s are formed in guanine-rich nucleic acids and further stabilized in the presence of monovalent cations (Hänsel-Hertsch et al. 2017; Spiegel et al. 2020). They are evolutionarily conserved and stable under physiologic conditions (Chen et al. 2018; Marsico et al. 2019). In mammals, the genomic distribution of G4s is not random but rather peculiar to specific genomic regions, such as telomeres, gene promoters, transcription factor binding sites, and sites with DNA double-strand breaks (Biffi et al. 2013; Varshney et al. 2020; Zheng et al. 2020). Although the chemistry of G4s has been under investigation for decades, the important biological functions of G4 have just begun to emerge recently.

Based on their distribution in the genome, G4s have been implicated in several essential cellular processes, such as gene transcription, DNA replication, genomic instability, and telomere elongation and maintenance (Varshney et al. 2020). Gene transcription is a driving force of chromatin relaxation and single-stranded DNA (ssDNA) exposure, which is a prerequisite for G4 formation. Chromatin immunoprecipitation with an engineered G4 structure-specific antibody BG4 followed by high-throughput

sequencing (G4 ChIP-seq) has detected and mapped endogenous G4s in mammalian cells (Biffi et al. 2013; Hänsel-Hertsch et al. 2018). Using this method, ~10,000 G4 structures on chromatin have been identified in human cells, the majority of which mainly reside upstream of the transcription start sites (TSSs) of actively transcribed genes (Hänsel-Hertsch et al. 2016; Zheng et al. 2020), suggesting the potential interplay between G4s and transcriptional regulation.

Dysfunctions of G4s have been seen in neurodegenerative diseases and breast cancer (Hänsel-Hertsch et al. 2020; Wang et al. 2021a), and G4s were suggested to serve as potential therapeutic targets for DNA-targeted therapies, particularly in anticancer drug design (Neidle 2016; Zyner et al. 2019; Carvalho et al. 2020). The application of multiple G4-stabilizing compounds (G4 ligands), such as TMPyP4, pyridostatin (PDS), and PhenDC3, as potential anticancer drugs is currently being evaluated. These compounds were initially developed to interfere with telomere functions and alter transcription of oncogenes (De Cian et al. 2007; Rodriguez et al. 2012; Carvalho et al. 2020). Additionally, G4 formation was elevated in immortalized cells compared to their normal counterparts (Hänsel-Hertsch et al. 2016), and the differentially enriched G4 regions can function as genomic markers of regions that drive breast cancer and serve as predictors of drug

⁷These authors contributed equally to this work.

Corresponding author: KWliang@whu.edu.cn

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.275431.121>.

© 2021 Li et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <https://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

response to G4 ligands (Hänsel-Hertsch et al. 2020). Furthermore, genetic interaction studies of G4s uncovered many genetic vulnerabilities to G4 ligands, raising new therapeutic possibilities for G4 ligands in anticancer treatment (Zimmer et al. 2016; Zyner et al. 2019). After decades of development, some G4 ligands have reached advanced phase I and phase II trials as candidate therapeutic agents against several types of tumors (Drygin et al. 2009; Xu et al. 2017; Carvalho et al. 2020). However, elucidating how G4s are regulated by G4 ligands, especially at promoters, and the specific mechanisms underlying the biological roles of G4s and G4 ligands are still challenging.

Plenty of studies have linked G4 formation with transcriptional regulation, and different models have been proposed for G4 involved in transcription at promoters and gene bodies (Spiegel et al. 2020; Varshney et al. 2020). G4 has been reported to act as a direct or indirect roadblock for RNA polymerase II (Pol II) elongation (Varshney et al. 2020), promoting or inhibiting the recruitment of specific transcription factors and cofactors (Raiber et al. 2012; Gao et al. 2015; Li et al. 2017; Makowski et al. 2018). However, this evidence is largely based on computationally predicted G4 motifs, correlations between G4 and gene expression, or manipulation of G4 structures on individual genes in plasmid constructs (Spiegel et al. 2020; Varshney et al. 2020). The genome-scale interplay between native G4 and transcription remains unknown. More explicit evidence of native G4 involvement in transcription and scrutiny of the potential interference of indirect or network effects are imperative and would promote a better characterization of the direct roles of G4 in genome-scale transcriptional regulation. In this study, we established a sensitive G4-CUT&Tag (cleavage under targets and tagmentation) method (Kaya-Okur et al. 2019; Wang et al. 2019) to profile native G4s across the human genome with high resolution and specificity. Using G4 stabilizing ligands, we combined G4-CUT&Tag with multiple complementary methods to elucidate how native G4 participates in transcription on a genome-scale.

Results

Genome-wide profiling of native G4s by the cleavage under targets and tagmentation approach

To identify the G4s in the human genome, we purified the G4 structure-specific scFv antibody BG4 (Flag-BG4 antibody with a His-tag) (Supplemental Fig. S1A) and used the *in vitro* enrichment-based G4 ChIP-seq protocol (Supplemental Fig. S1B,C; Hänsel-Hertsch et al. 2018) to profile the G4s in HEK293T cells. G4 signals were successfully mapped and showed similar profiles as in a previously published G4 ChIP-seq data set (Supplemental Fig. S1D; Hänsel-Hertsch et al. 2018). In addition, motif analysis for the G4 peak regions in ChIP-seq revealed a significant enrichment for G-rich motifs, which is consistent with the property of G4 sequences and the reported G4 motifs (Supplemental Fig. S1E; Hänsel-Hertsch et al. 2018). However, we noticed that, in general, G4 ChIP-seq has a higher background, which is likely due to the *ex vivo* enrichment-based strategy of ChIP-seq. To improve the mapping of G4s and profile the native G4s in the genome, we combined the Tn5-based CUT&Tag approach (Kaya-Okur et al. 2019; Wang et al. 2019) with Flag-BG4 antibody and established an *in situ* G4 mapping method, named G4-CUT&Tag (Fig. 1A). In this method, G4 structures are recognized by Flag-BG4 antibody *in situ*, and G4-containing DNA fragments are cut and released by assembled Tn5 transposome-assisted tagmentation. We successfully

generated the G4-CUT&Tag libraries with pre-assembled Tn5 transposome efficiently (<1 d) using 1×10^5 HEK293T cells (Supplemental Fig. S1F).

Genome browser snapshots of G4s at individual genes (Fig. 1B; Supplemental Fig. S1G) and genome-wide analysis of G4 signals across all the protein-coding genes (Supplemental Fig. S1H) showed that G4-CUT&Tag produced much higher signals at the promoter regions with fewer background signals compared to G4 ChIP-seq. The metaplots and fingerprint plots of G4-CUT&Tag libraries and their corresponding controls showed a higher signal-to-noise ratio in HEK293T cells (Supplemental Fig. S1I,J). Peak calling with model-based analysis of ChIP-seq (MACS2) (Zhang et al. 2008) identified 17,888 G4 peaks along with an IgG-CUT&Tag control, whereas only 9202 peaks were characterized from G4 ChIP-seq with an IgG ChIP-seq control, suggesting the superior sensitivity of G4-CUT&Tag. Size distribution analysis of these peaks suggests that G4-CUT&Tag has a slightly higher resolution (Fig. 1C). The high correlation between biological replicates demonstrated the reproducibility of G4-CUT&Tag (Fig. 1D). We identified 5815 overlapped peaks between G4-CUT&Tag and G4 ChIP-seq, with 12,073 and 3387 unique peaks for G4-CUT&Tag and G4 ChIP-seq, respectively (Fig. 1E), suggesting that the *in situ* and *ex vivo* capturing strategies may result in different G4 profiles. Annotation of G4 peaks revealed that the majority of G4s localize at the gene promoters and a subset of G4 peaks were distributed at the gene body and intergenic regions (Fig. 1E). In agreement with a better signal-to-noise ratio, the 12,073 unique peaks in G4-CUT&Tag analysis were partly annotated to gene promoters, but a large proportion of them were enriched in distal intergenic regions, which often contain enhancers, indicating that G4-CUT&Tag is more sensitive than G4 ChIP-seq in detecting G4s (Fig. 1E).

To investigate the relationship between native G4 and epigenetic landscapes, we profiled the Pol II (POLR2A subunit) occupancy, histone modifications (including H3K27ac, H3K4me3, H3K4me2, and H3K4me1), as well as chromatin accessibility by ATAC-seq in HEK293T cells. We found that G4s colocalized with Pol II and active histone modifications and were accompanied with high chromatin accessibility at promoters and enhancers (Fig. 1F; Supplemental Fig. S2A,B). Genome-wide analysis showed that G4s positively correlated with active histone modifications, Pol II occupancy, and ATAC-seq signals (Fig. 1G). G4s were substantially prevalent at gene promoters (about 60%; 10,647 in 17,888) with high Pol II occupancy, chromatin accessibility, and high levels of active histone modifications, including H3K27ac, H3K4me2, and H3K4me3 (Fig. 1H). The metaplot analysis revealed that G4s colocalized with Pol II in the nucleosome-free regions between two nucleosomes marked by H3K27ac (Fig. 1I,J). We further compared the distribution of G4s with precision nuclear run-on sequencing (PRO-seq) signals (NCBI Gene Expression Omnibus [GEO: <https://www.ncbi.nlm.nih.gov/geo/>] accession numbers GSM4730174 and GSM4730175), which measures nascent RNA 3' ends and maps transcriptionally engaged Pol II. G4 signals overlapped with the TSSs and Pol II pausing sites at promoter regions (Fig. 1I,J), indicating that G4s may participate in the early stages of transcription.

Because G4-CUT&Tag could detect the G4 signals at enhancers (Fig. 1E), we further characterized the G4 structures at different enhancer types (defined by specific histone modifications). Active and poised enhancers were clustered by their relative intensities of H3K27ac and H3K4me1 (Supplemental Fig. S2C; Wang et al. 2015). Approximately 46% of G4 peaks (8168 in 17,888) reside at active enhancers, which are associated with a high H3K27ac/

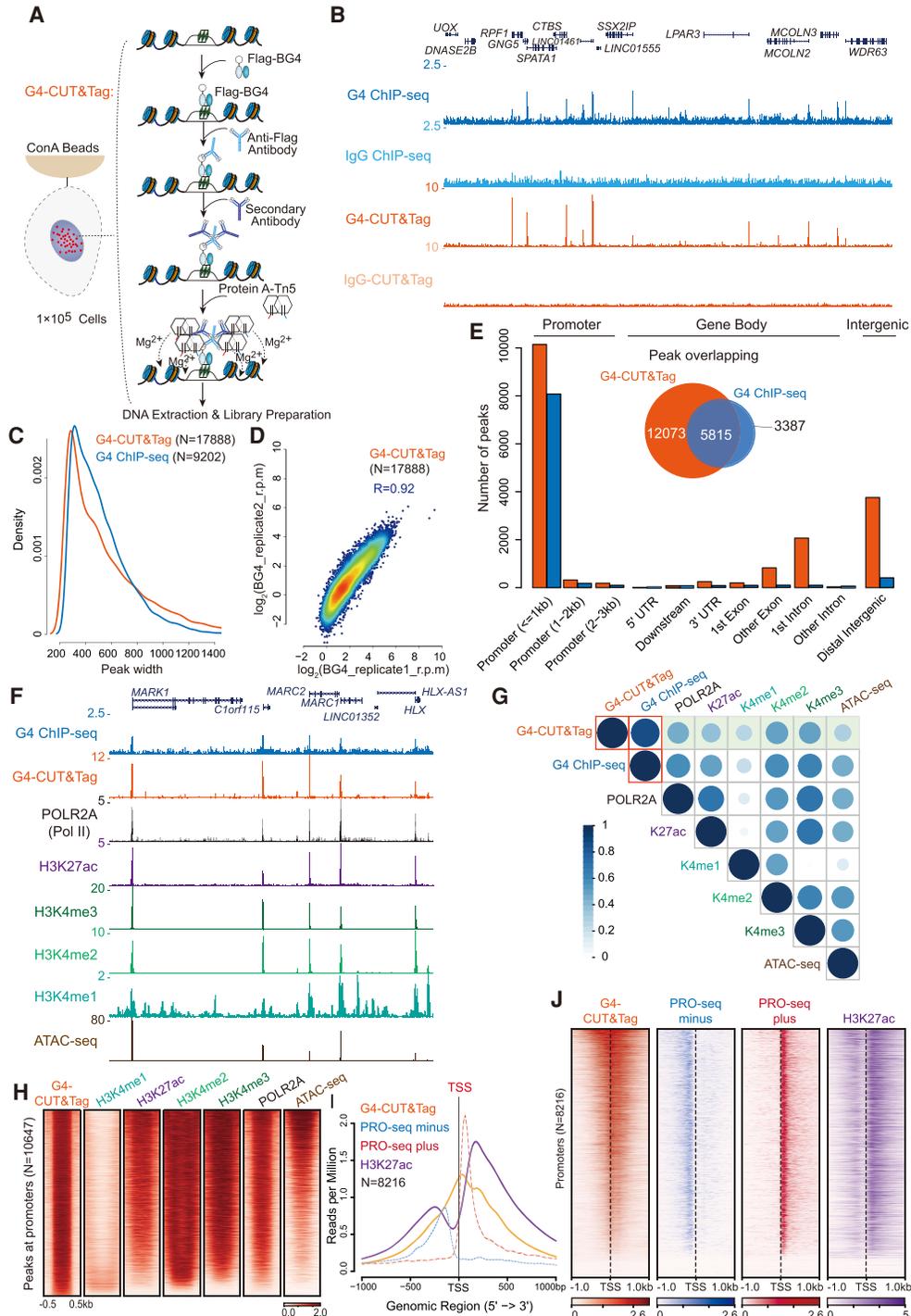


Figure 1. Comprehensive profiling of genome-wide native G4s by G4-CUT&Tag. (A) Schematic of G4-CUT&Tag with a G4-structure-specific single-chain antibody Flag-BG4. After binding of concanavalin A-coated magnetic beads to the cells and cell permeabilization, G4 structures were recognized by Flag-BG4 in situ. Anti-Flag antibody and secondary antibody were added sequentially to tether protein A-Tn5 transposase. Tn5 was activated by magnesium to cut the chromatin close to Flag-BG4 binding sites and simultaneously integrate adapters. After extraction, the Tn5-tagmented genomic DNA was amplified for library preparation and second-generation sequencing. For IgG control, mouse IgG was used to replace Flag-BG4 antibody. (B) UCSC Genome Browser tracks of G4 ChIP-seq and G4-CUT&Tag in HEK293T. The reads were aligned to human genome hg38 and normalized by reads per million. (C) Distributions of peak width of G4-CUT&Tag and G4 ChIP-seq. G4-CUT&Tag and G4 ChIP-seq peaks were called by MACS2 using corresponding IgG controls. A total of 17,888 G4-CUT&Tag peaks and 9202 G4 ChIP-seq peaks were identified. (D) Correlation plot between two G4-CUT&Tag biological replicates in HEK293T cells. (E) Genome-wide annotation of HEK293T G4-CUT&Tag and G4 ChIP-seq peaks. G4-CUT&Tag detected more G4s at the promoter, gene body, and intergenic regions than the G4 ChIP-seq method. (F) Tracks example of G4 ChIP-seq, G4-CUT&Tag, Pol II (subunit: POLR2A), H3K27ac, H3K4me3, H3K4me2, H3K4me1 ChIP-seq, and chromatin accessibility (ATAC-seq) in HEK293T, showing G4s mark transcriptional regulatory elements enriched with Pol II, multiple active chromatin marks, and open chromatin. (G) Correlation plot of G4s with POLR2A, histone modifications, and chromatin accessibility. Signals at the 17,888 G4-CUT&Tag peaks were counted and calculated by Pearson's correlation. (H) Heat map analysis of G4, histone modifications, POLR2A occupancy, and ATAC-seq signals at promoter regions in HEK293T. Heat maps were sorted by decreasing intensity of G4-CUT&Tag signals. (I, J) Metaplot and heat map of G4-CUT&Tag, PRO-seq, and H3K27ac signals at promoter regions. The TSSs are denoted by the vertical black-dashed lines.

H3K4me1 ratio, high Pol II occupancy, and chromatin accessibility, indicating that these enhancers are transcriptionally engaged (Supplemental Fig. S2C,D). Approximately 14% of G4 peaks (2480 in 17,888) were distributed at poised enhancers (Supplemental Fig. S2C), which display weaker G4 signals, H3K27ac, Pol II, and ATAC-seq intensities compared to those at promoters and active enhancers (Supplemental Fig. S2D–F). Additionally, motif discovery revealed enrichment of six G-rich recurring motifs, which are consensus motifs of the SP1, FOXO, and E2F family proteins (Supplemental Fig. S2G,H). Together, these results demonstrate that G4-CUT&Tag is a sensitive and reproducible method for comprehensive profiling of native G4s, which mark transcriptional regulatory elements across the genome.

Characterization of native G4 signals by G4-CUT&Tag and their cell type specificity

To test the specificity of G4-CUT&Tag signals, we analyzed the 17,888 G4-CUT&Tag peaks and their equivalent random sequences with the widely used G4Hunter prediction program. This program computationally estimates the G4 forming ability relying on the G-richness and G-skewness of DNA sequence and provides quadruplex propensity scores (Bedrat et al. 2016). Prediction of G4-CUT&Tag peak sequences yielded 1,501,354 G4 hits, which were nearly four times as many as those from equivalent random sequences. Furthermore, these hits were associated with higher quadruplex propensity scores (Supplemental Fig. S3A,B). Next, we compared the G4-CUT&Tag signals with G4-seq signals, which represent observed G4 sequences (OQs) in purified, single-stranded genomic DNA (Marsico et al. 2019). We found that G4-CUT&Tag signals largely overlapped with OQs at individual gene promoters, as shown in the genome browser tracks (Fig. 2A,B). Genome-wide analysis revealed that the majority of G4-CUT&Tag peaks contained OQs detected by G4-seq (Fig. 2C; Supplemental Fig. S3C) and their signals were generally correlated (Fig. 2C) with similar distribution patterns across different genomic regions (Supplemental Fig. S3D–F), confirming the specificity of G4-CUT&Tag signals.

R-loops are noncanonical three-stranded structures in which the nascent RNA invades into double-stranded DNA (dsDNA) and forms DNA:RNA hybrids during transcription (Niehrs and Luke 2020). G4s are suggested to be compatible and strongly correlated with R-loops, and G4s are preferentially distributed in the nontemplate and displaced DNA strand of an R-loop (Duquette et al. 2004; De Magis et al. 2019). Thus, we analyzed G4-CUT&Tag signals with R-loop signals detected by 2×HBD-CUT&Tag and S9.6-CUT&Tag (GSE156400) in HEK293T cells (Wang et al. 2021b). The *NPM1* and *GRK6* loci each displayed G4 and R-loop signals, and the G4s were located upstream of the mapped R-loops (Fig. 2D), which is consistent with previous studies (Chen et al. 2017). We found that 81.1% and 80.7% G4 peaks overlapped with R-loop signals detected by R-loop 2×HBD-CUT&Tag and S9.6-CUT&Tag (Fig. 2E), respectively. Also, G4-CUT&Tag intensities were positively correlated with R-loop 2×HBD-CUT&Tag signals ($r=0.664$) and S9.6-CUT&Tag signals ($r=0.599$) (Fig. 2F,G). Collectively, these results indicate a genome-wide tight association of G4s with R-loops and further confirm the specificity of G4 signals from G4-CUT&Tag.

In order to measure the cell type specificity of native G4s, we analyzed the G4 profiles in five cell lines of different origins, including embryonic kidney cells (HEK293T), chronic myelogenous leukemia cells (K562), cervical cancer cells (HeLa), small cell lung

cancer cells (SW1271), and breast cancer cells (MBD-231-LM2) (Fig. 2H). We performed peak calling with input DNA sequencing data from each cell line (NCBI Sequence Read Archive [SRA]; <https://www.ncbi.nlm.nih.gov/sra>) accession numbers SRX5466670, SRX3358201, SRX5449793, SRX6858029, and ERX4517391). The low degree of G4 peak overlap between these cell lines indicated that the G4s were highly cell type-specific, as shown in the Venn diagram (Fig. 2I). Pearson's correlation of G4-CUT&Tag signals also exhibited great heterogeneity of G4s among different cell lines (Fig. 2J). Together, these data demonstrate that the G4 distribution and signal intensities vary widely among different cells, suggesting the cell type specificity of G4s and likely reflecting diverse epigenetic landscapes and transcription programs among different cell lines.

The reciprocal interaction between RNA polymerase II and native G4

Given the findings that G4 marks transcriptional regulatory elements and colocalizes with Pol II at promoters, we further investigated the interplay between G4 and Pol II in two ways: first, by perturbing the dynamics of transcription; and second, by stabilizing native G4. We used the well-known CDK9 inhibitor flavopiridol (FP) to induce global promoter-proximal Pol II pausing at the promoter and enhancer regions. The induction of Pol II pausing is mediated through the inhibition of phosphorylation of three key polypeptides: the Pol II C-terminal domain; the negative transcription elongation factors (NELF); and the DRB sensitivity inducing factor (DSIF) (Liang et al. 2015). As measured by ChIP with reference exogenous genome (ChIP-Rx) (Orlando et al. 2014), 1-h treatment with FP rapidly induced retention of the Pol II (POLR2A) occupancy (Fig. 3A,B). Moreover, FP increased native G4 signals at promoters (Fig. 3A,B). We also found that FP increased the Pol II occupancy and G4-CUT&Tag signals at both active and poised enhancers (Fig. 3C,D). In summary, these data suggest that enhanced Pol II pausing by FP promotes native G4 formation at promoters and enhancers.

To investigate the effects of native G4s on Pol II-mediated transcription, we utilized two G4 stabilizing compounds, TMPyP4 tosylate (Anantha et al. 1998) and pyridostatin (Hänsel-Hertsch et al. 2017), to modulate the native G4s in the genome. We confirmed the specificity of BG4 and BG4-EGFP to G4s by dot-blot assays (Supplemental Fig. S3G) and found that TMPyP4 or PDS treatment did not block the interaction between annealed MYC-G4 oligonucleotides and BG4 or BG4-EGFP (Supplemental Fig. S3H). In addition, we checked the purity of the G4 ligands TMPyP4 and PDS by liquid chromatography-mass spectrometry (Supplemental Fig. S4A,B). To minimize potential secondary effects, we treated the HEK293T cells with these ligands for 1 h only. Immunostaining of G4s with BG4-EGFP antibody showed that TMPyP4 rapidly and robustly increased the nuclear G4s (Fig. 3E,F). Next, we profiled the native G4s in the genome by G4-CUT&Tag after 1 h of TMPyP4 treatment. TMPyP4 treatment increased the G4-CUT&Tag signals at the *NPM1* promoter (Fig. 3G). Genome-wide analysis confirmed that TMPyP4 induced a global increase of native G4s at promoters (Fig. 3H), especially around the TSSs (Supplemental Fig. S4C). We also measured the Pol II subunits POLR2A and POLR2B, along with Pol II-associated NELF complex member B (NELFB) by ChIP-Rx to investigate the effects of G4 stabilization on gene transcription. We found that TMPyP4 treatment rapidly reduced Pol II and NELF occupancy at TSSs (Fig. 3G). The heat map and metaplot analyses also revealed

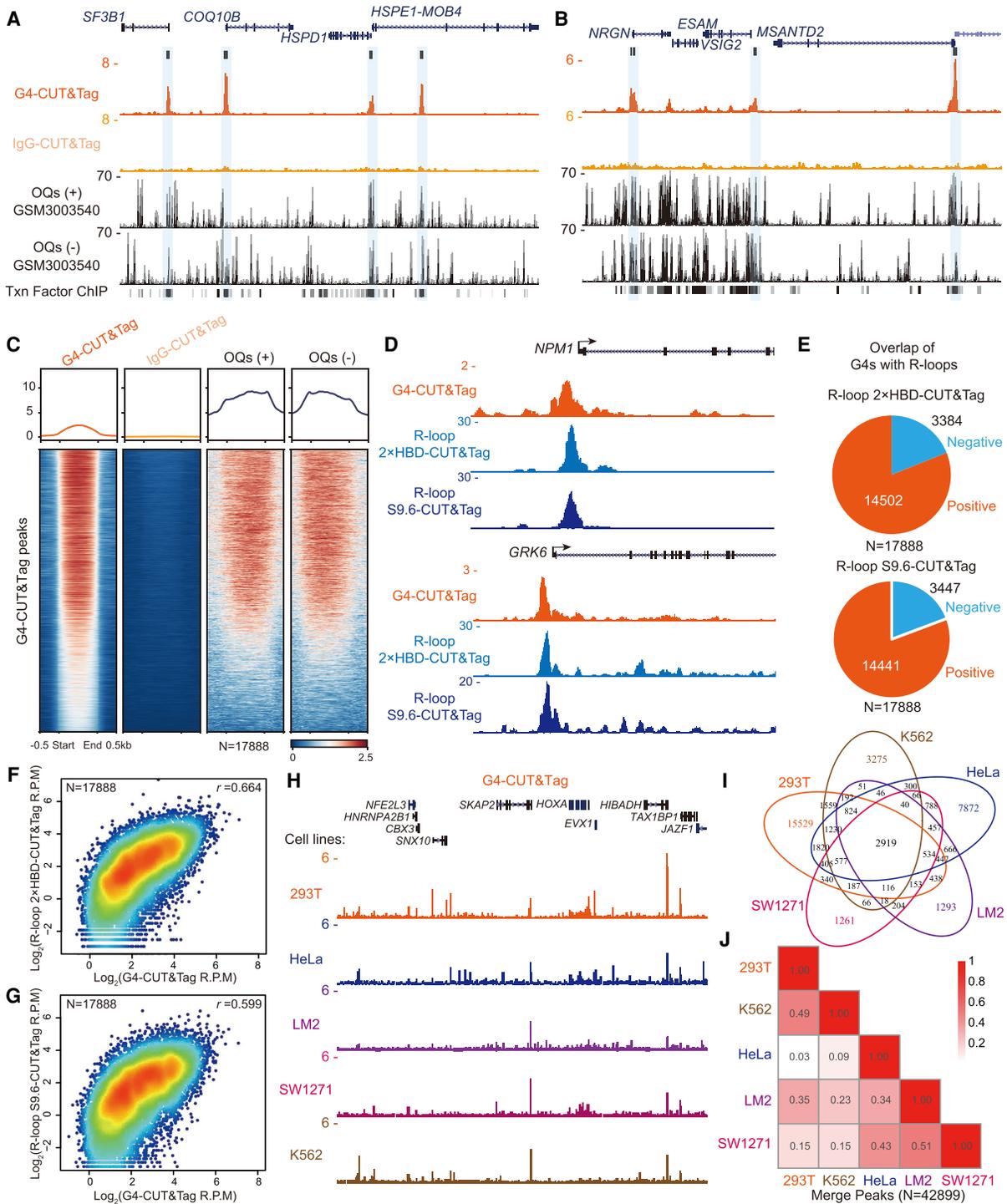


Figure 2. Characterization of native G4 signals in G4-CUT&Tag and their cell type specificity. (A,B) Genome browser tracks of G4-CUT&Tag and G4-seq in HEK293T cells. Observed G4 sequences (OQs) detected by G4-seq (GSM3003540) on the forward (+) and reverse (-) strands (Marsico et al. 2019) were compared with G4-CUT&Tag signals. Txn Factor ChIP track was derived from combining ChIP-seq peaks of 161 transcription factors in 91 cell types performed by the ENCODE Project. (C) Heat map and metaplot analyses of G4-CUT&Tag and G4-seq signals at G4 peaks. G4-CUT&Tag signals were positively correlated with OQs signals. Metaplots of these signals are shown at the top. (D) Track examples of G4-CUT&Tag, R-loop 2 × HBD-CUT&Tag, and R-loop S9.6-CUT&Tag signals at the *NPM1* and *GRK6* loci. (E) Pie graphs showing the overlap of G4-CUT&Tag peaks with R-loop signals identified by R-loop CUT&Tag with the DNA:RNA hybrid sensor 2 × HBD and S9.6 monoclonal antibody, respectively (Wang et al. 2021b). (F,G) Pearson’s correlation of G4-CUT&Tag signals with R-loop 2 × HBD-CUT&Tag signals (F) and R-loop S9.6-CUT&Tag signals (G) at the 17,888 G4 peaks. (H) A snapshot from the UCSC Genome Browser showing G4-CUT&Tag signals in HEK293T, K562, HeLa, SW1271, and MBD-231-LM2 (LM2) cells. (I) G4 peaks were called with the corresponding whole-genome sequencing data of each cell line (SRX5466670, SRX3358201, SRX5449793, SRX6858029, and ERX4517391). Overlapping of these G4 peaks showed heterogeneity of G4s among different cell lines. (J) Correlation analysis of log₂ fold change of G4 signals versus input in HEK293T, K562, HeLa, SW1271, and LM2 cells, showing the intercellular heterogeneity of G4s.

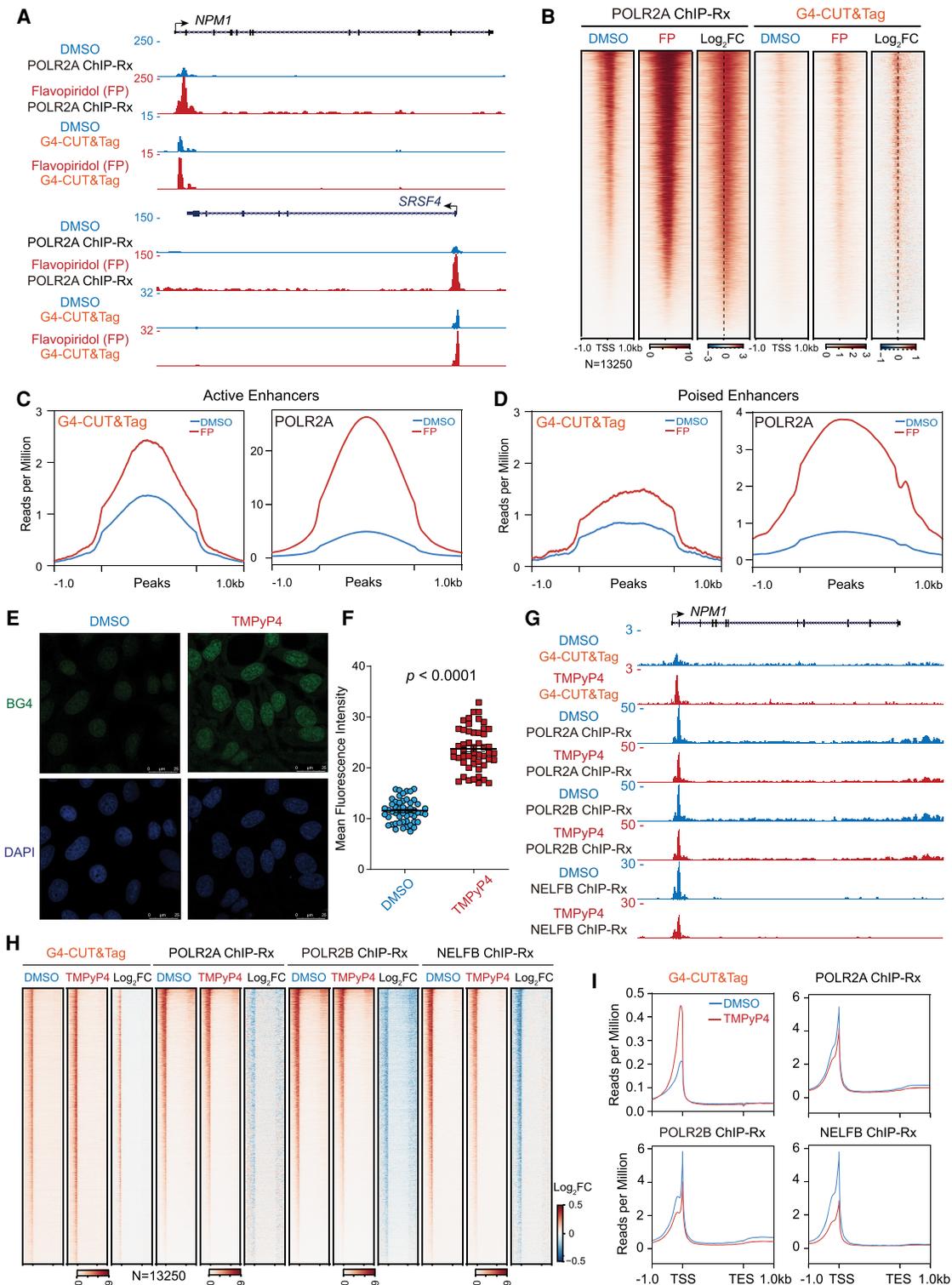


Figure 3. The reciprocal interaction between RNA polymerase II and native G4 in the genome. (A,B) Pol II subunit POLR2A ChIP-Rx and G4-CUT&Tag analyses showed that flavopiridol treatment resulted in increased Pol II occupancy and slightly increased G4 signals as seen at the promoter-proximal regions of *NPM1* and *SRSF4* (A) or around TSSs (B) by heat map analysis. (C,D) Metaplot analysis of G4 and POLR2A peak regions at active (C) and poised enhancers (D). (E,F) Immunostaining of G4 using BG4-EGFP (green) in HeLa cells treated with DMSO or TMPyP4 for 1 h. The nuclei were stained with DAPI (blue) (E). For each sample, images of 4–6 fields of vision were acquired with fixed parameters. Fifty-one nuclei in each sample were randomly selected, and the mean fluorescence intensities of G4s in each nuclear area were calculated by dividing the total intensities of BG4-EGFP with the nuclear area marked by DAPI (F). (G–I) Analysis of G4-CUT&Tag signals and ChIP-Rx signals detected by the Pol II subunits POLR2A and POLR2B antibodies as well as the NELF subunit NELFB antibody in HEK293T cells with or without TMPyP4 treatment. As shown at *NPM1* loci (G) or heat map (H) and metaplots (I) at gene coding regions, TMPyP4 treatment induced a global increase of G4 signals and decrease of Pol II and NELF occupancy around TSSs.

a global reduction of Pol II occupancy after G4 stabilization with TMPyP4 (Fig. 3H,I).

To verify these findings, we further tested the effect of another well-known G4 stabilizing ligand, PDS, on G4 and Pol II occupancy. PDS exhibited weaker effects than TMPyP4 on G4 as shown by the BG4-EGFP staining and quantification (Supplemental Fig. S4D,E). The G4-CUT&Tag analysis showed that PDS also increased genome-wide G4 signals at gene promoters, albeit to a lesser extent (Supplemental Fig. S4F). These findings are consistent with the BG4-EGFP immunostaining analysis and demonstrate that both TMPyP4 and PDS enhance the genome-wide formation of G4s. POLR2A and NELFB ChIP-Rx analyses showed that PDS also substantially decreased the Pol II and NELF occupancy at gene promoters (Supplemental Fig. S4F), confirming that ligand-induced G4 stabilization impaired Pol II and NELF occupancy at gene promoters throughout the genome.

Native G4 stabilization by G4 ligands rapidly modulates chromatin states

Next, we tried to investigate the underlying mechanisms by which G4 regulates transcription. First, we performed a POLR2A ChIP-Rx assay with HEK293T cells after 1 and 4 h of TMPyP4 treatment. We found that TMPyP4 decreased Pol II occupancy in a time-dependent manner as shown by a further decrease of POLR2A signal intensity at promoters after extended TMPyP4 treatment (Supplemental Fig. S5A,B). The reduction of Pol II occupancy started at the regions upstream of TSSs, indicating that G4 stabilization may impair gene transcription at the transcription initiation stage (Supplemental Fig. S5B). Surprised by such a rapid alteration of Pol II and its associated factors by G4 stabilizing ligands, we were curious about the changes in the chromatin states at promoters, including dsDNA melting and chromatin accessibility. Thus, we performed kethoxal-assisted single-stranded DNA sequencing (KAS-seq) (Wu et al. 2020b), which uses N_3 -kethoxal to rapidly and specifically label the guanines in single-stranded DNA, to track changes of chromatin states by G4 ligands. The G-quartet utilized the Hoogsteen bonds from four guanine bases to form a planar tetrad, which blocked the reaction between guanine and N_3 -kethoxal (Fig. 4A). TMPyP4 treatment for 1 h diminished the KAS-seq signals at the promoters, gene body, and transcription end sites (TESs) of *NPM1* and *SRSF4* genes (Fig. 4B). Heat map and metaplot analyses with all of the active promoters (Fig. 4C,D) showed that TMPyP4 decreased the KAS-seq signals globally. Moreover, from metagene analysis (Supplemental Fig. S5C,D), we found that TMPyP4 treatment also caused genome-wide reductions of the KAS-seq signals at the gene body and TES regions, which agrees with the reducing POLR2A ChIP-Rx signals at gene bodies after TMPyP4 treatment (Fig. 3I). Consistent with the competition between G-quartet formation and N_3 -kethoxal labeling, we found that the vast majority of promoters with enhanced G4s after TMPyP4 treatment showed attenuated KAS-seq signals (Supplemental Fig. S5E).

To further investigate the degrees of dsDNA melting and single-stranded DNA exposure, we utilized non-denaturing bisulfite conversion assays to convert the unmethylated cytosines (C) on single-stranded DNA into uracils (U) and used Sanger sequencing to measure the C-T conversion after various G4 ligand treatments. We observed that both *NPM1* and *SRSF4* promoters displayed more non-denaturing bisulfite conversion after TMPyP4 or PDS treatment (Fig. 4E; Supplemental Fig. S5F). Combined with KAS-seq results, the data reflect more dsDNA melting and ssDNA opening caused by G4 stabilization. To further measure the consequences

on chromatin accessibility by G4 stabilization, we performed transposase-accessible chromatin using sequencing (ATAC-seq) (Corces et al. 2017) to map the chromatin accessibility after TMPyP4 treatment for 1 h. We found that G4 stabilization by TMPyP4 increased the chromatin accessibility at the *NPM1*, *SRSF4*, and *HSPA8* promoters (Fig. 4F; Supplemental Fig. S5G). Genome-wide analysis of ATAC-seq signals at promoters uncovered substantially increased chromatin accessibility after G4 stabilization (Fig. 4G,H). Together, these results demonstrate that native G4 stabilization by G4 ligands rapidly modulates chromatin states by increasing dsDNA melting and chromatin accessibility in the genome.

G4 stabilization substantially impairs nascent RNA synthesis

To further characterize the alterations of gene transcription by G4 ligands, we measured the nascent RNA synthesis after 1-h DMSO or TMPyP4 treatment using transient transcriptome sequencing (TT-seq) (Schwalb et al. 2016; Liang et al. 2018) with exogenous S2 spike-in controls to minimize potential interference of indirect or network effects. Track examples (Fig. 5A,B) showed that TMPyP4 treatment decreased the nascent RNA synthesis of the *NPM1* and *SRSF4* genes. The differential gene expression analysis revealed reduced nascent RNA synthesis after TMPyP4 treatment for 1 h (Fig. 5C; Supplemental Fig. S6A). We identified 5870 differentially regulated genes (fold change > 1.5 and $P < 0.05$) by DESeq2 with S2 spike-in controls (Fig. 5C,D). The vast majority of these 5865 genes were down-regulated with 1-h TMPyP4 treatment (Fig. 5D), which is consistent with the POLR2A ChIP-Rx and KAS-seq results (Figs. 3, 4). Furthermore, the metagene plot of strand-specific TT-seq signals showed that TMPyP4 reduced the nascent RNA synthesis at the gene body and TES regions, which was in agreement with the reduction of Pol II occupancy and KAS-seq signals at these regions (Fig. 5E; Supplemental Fig. S6B). TMPyP4 treatment also decreased the enhancer RNAs (Supplemental Fig. S6C) and bidirectional antisense RNA synthesis at promoters (Fig. 5E). Together, these results demonstrate that G4 stabilization by TMPyP4 quickly inhibits genome-wide nascent RNA synthesis. To visualize the effects of TMPyP4 on individual gene transcription, we performed Stellaris RNA fluorescence in situ hybridization (FISH) (Liang et al. 2015) in HeLa cells with or without TMPyP4 treatment to measure the nascent RNA synthesis of *GAPDH* and *ACTG1* genes in the nucleus (Supplemental Fig. S6D). Without TMPyP4 treatment, we detected bright RNA FISH spots using *GAPDH* and *ACTG1* probes (Supplemental Fig. S6E-H). These punctate spots in the nucleus were brighter than the cytoplasmic RNA transcripts, suggesting that they likely represent sites of multiple nascent RNA transcripts. TMPyP4 treatment reduced the number of punctate spots (Supplemental Fig. S6E,G) and significantly decreased the intensity of these RNA FISH spots (Supplemental Fig. S6F,H).

Although native G4s were mainly present at the promoter regions, G4s were proposed to act as direct or indirect roadblocks for Pol II elongation in the gene body as evidenced from the artificial reporter gene assays (Varshney et al. 2020). To check the influences of native G4s on elongating Pol II in the gene body, we performed quick Precision Run-On and sequencing (qPRO-seq) (Judd et al. 2020) in the presence or absence of TMPyP4. If G4s could function as roadblocks for Pol II elongation, treatment of cells with TMPyP4 would most likely retain the elongating Pol II at the gene bodies or cause early termination of Pol II at TESs. However, we did not observe apparent retention of elongating Pol II from qPRO-seq at the

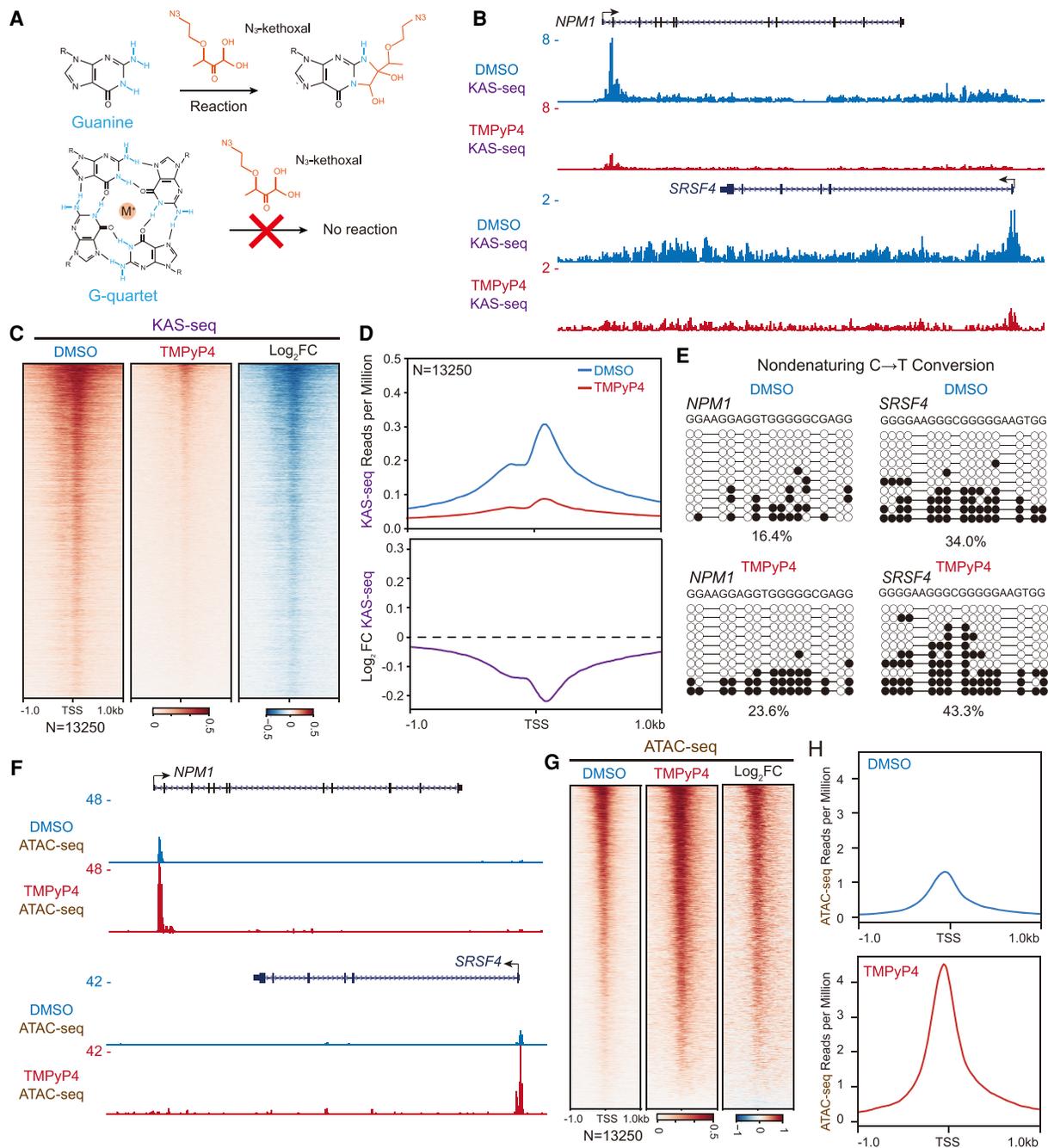


Figure 4. G4 stabilization by G4 ligands rapidly alters chromatin states with increased dsDNA melting and chromatin accessibility. (A) Schematic for N_3 -kethoxal labeling reaction in KAS-seq. In the G-quartet structure, the formation of Hoogsteen bonds blocks the chemical reaction between guanine and N_3 -kethoxal. (B–D) Analysis of KAS-seq signals in HEK293T cells treated with DMSO or TMPyP4. Tracks examples illustrate that TMPyP4 reduces KAS-seq signals at the promoters, gene bodies, and TSSs of *NPM1* and *SRSF4* genes (B). Heat maps (C) and metaplots (D) demonstrate that TMPyP4 induced a global reduction of KAS-seq signals at TSSs. (E) Nondenaturing C to T conversion at two G4 forming motifs in *NPM1* and *SRSF4* promoters after TMPyP4 treatment in HEK293T cells. Each row shows one representative Sanger sequencing result. The converted sites are depicted as black dots, whereas the unconverted sites are represented as white dots. The percentages of C-T conversion sites for each group were calculated from these 10 sequences. (F–H) Analysis of ATAC-seq signals in the absence or presence of TMPyP4 reveals that TMPyP4 enhanced the chromatin accessibility around TSSs. Snapshots of ATAC-seq tracks at *NPM1* and *SRSF4* loci (F), heat maps (G), and metaplots (H) are shown.

NPM1 and *SRSF4* genes (Fig. 5F). We further analyzed the genome-wide strand-specific qPRO-seq signals along with the G4-CUT&Tag signals with or without TMPyP4 (Fig. 5G,H). The results showed that TMPyP4 decreased the overall elongating Pol II at the gene bodies, which is consistent with the TT-seq and POLR2A

ChIP-Rx data showing decreased nascent RNA synthesis (Fig. 5E) and reduced Pol II occupancy at the gene body (Fig. 3I), respectively. However, TMPyP4 treatment neither caused apparent early termination of Pol II at the TSSs nor induced evident elongating Pol II retention in the gene bodies (Fig. 5G,H). We also noted that

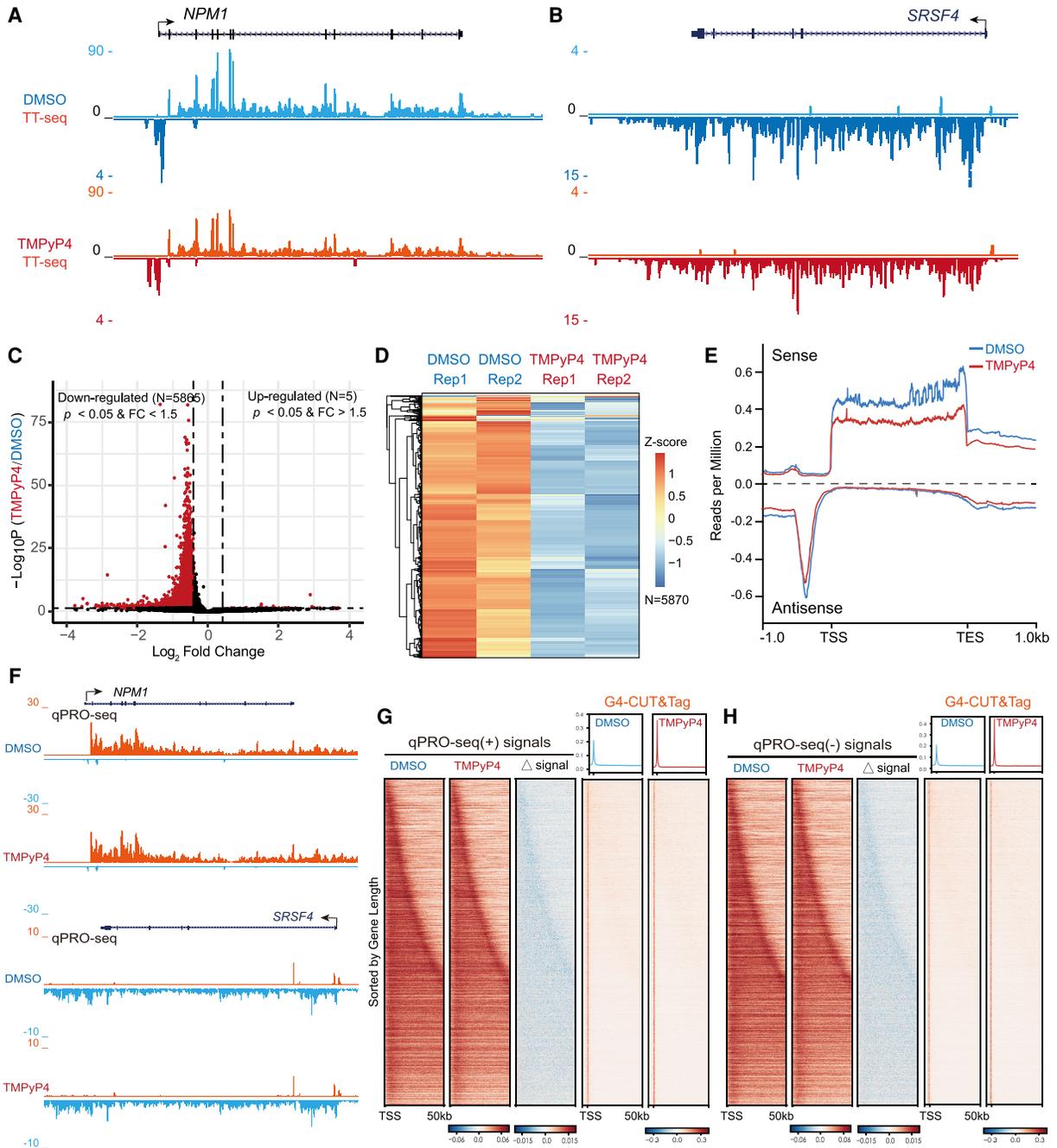


Figure 5. Disruption of G4 dynamics with G4 stabilizing ligands substantially restrains nascent RNA transcription. (A,B) UCSC Genome Browser snapshots of TT-seq tracks with DMSO or TMPyP4 treatment at gene coding regions of *NPM1* (A) and *SRSF4* (B) show that TMPyP4 down-regulated nascent RNA synthesis. (C,D) Volcano plot (C) and heat map (D) show substantially impaired nascent RNA synthesis in response to 1-h TMPyP4 treatment. A total of 5870 differentially genes (fold change > 1.5 and $P < 0.05$) were identified by DESeq2 with S2 spike-in controls. (E) Metagenome plot analysis of strand-specific TT-seq signals illustrates that TMPyP4 decreased nascent RNA synthesis at the gene body and bidirectional antisense RNA at promoter regions. (F) UCSC Genome Browser tracks of qPRO-seq at *NPM1* and *SRSF4* loci. (G,H) Heat map and metaplot showing strand-specific qPRO-seq signals and changes of qPRO-seq signals after TMPyP4 treatment. Rows represent genes and are sorted by gene length from shortest to longest. Corresponding G4-CUT&Tag signals are shown on the right sides.

TMPyP4 treatment did not induce native G4 formation in the gene bodies (Fig. 5G,H). These data are consistent with the genomic distribution of native G4s at promoters and suggest that blocking of Pol II elongation in the gene body is likely only a minor function of native G4 in the genome.

G4 stabilization inhibits transcription initiation by impairing the loading of general transcription factors to promoters

Next, we investigated the underlying mechanisms by which native G4 affects transcriptional regulation. We annealed the biotin-

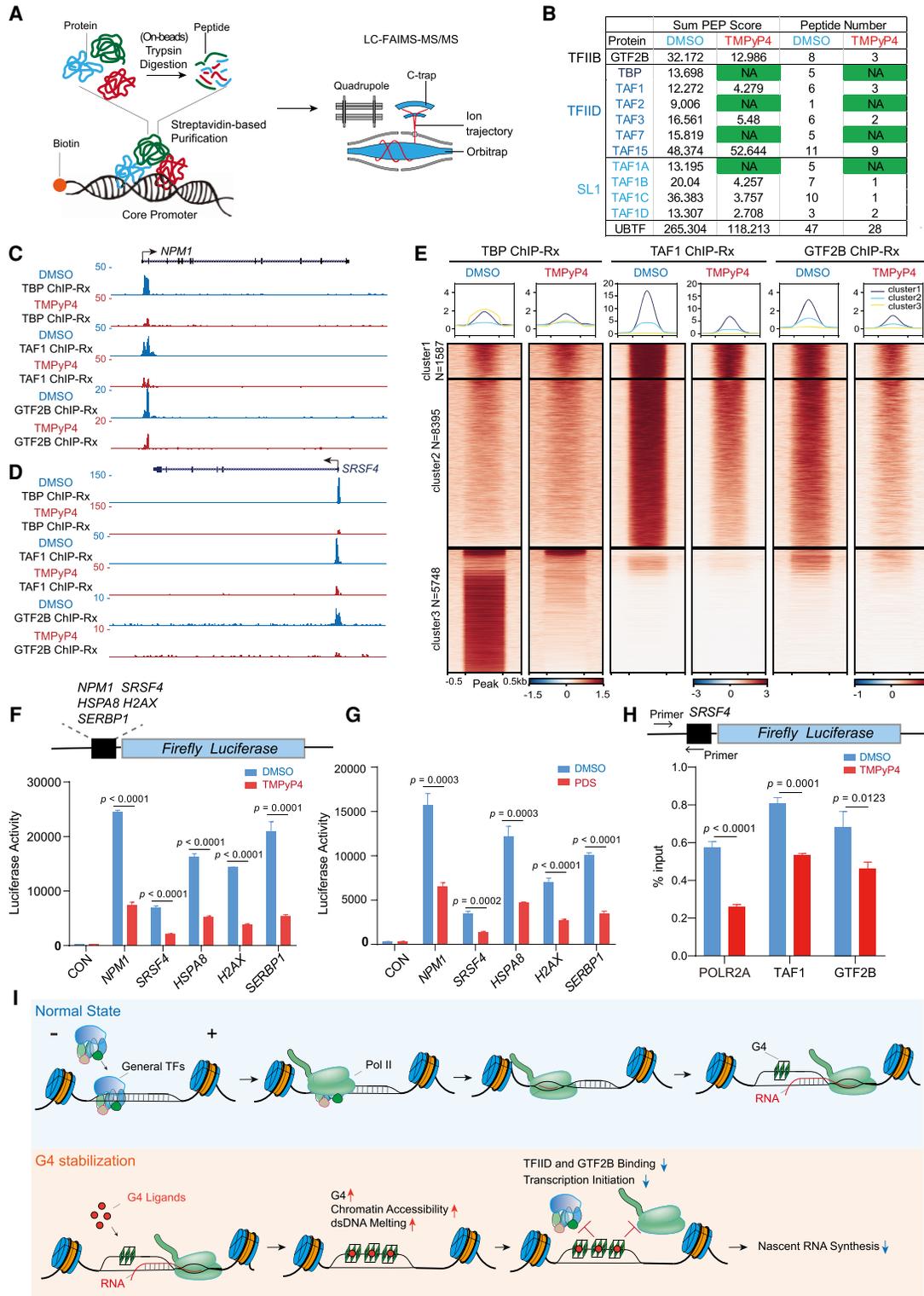


Figure 6. G4 stabilization inhibits transcription initiation by impairing general transcription factor loading. (A) Schematic of immunoprecipitation of G4-containing core promoter and nanoLC-FAIMS-MS/MS analysis. Streptavidin MagPoly beads were incubated with biotin-labeled G4-containing core promoter DNA and nuclear protein extract sequentially for immunoprecipitation. After on-beads trypsin digestion, the interacting proteins with core promoter DNA were analyzed on a Thermo Exporis 480 mass spectrometer. (B) List of core promoter interacting proteins with or without TMPyP4 treatment. (SL1) Selective Factor 1. Sum PEP Score is calculated as the sum of the negative logarithms of the posterior error probability (PEP) values of the connected peptide spectrum matches (PSMs). All the proteins in the list have high confidence, with q -values higher than the specified threshold 0.01. (C–E) Analysis of TBP, TAF1, and GTF2B ChIP-Rx signals in cells treated with DMSO or TMPyP4 for 2 h. Track examples at the *NPM1* (C) and *SRSF4* (D) loci are shown. Heat map and metaplot analyses (E) showed that TMPyP4 impaired TBP, TAF1, and GTF2B occupancy after TMPyP4 treatment. (F, G) Luciferase reporter assay of G4-containing promoters showed that G4 stabilization induced by TMPyP4 (F) or PDS (G) significantly inhibited the luciferase reporter gene expression. (H) ChIP-qPCR analysis in HEK293T cells transfected with recombinant luciferase reported plasmid. These cells were treated with or without TMPyP4. The occupancy of POLR2A, TAF1, and GTF2B in the insertion-vector junction of the recombinant plasmid was measured by ChIP-qPCR. The statistical analysis was performed using an unpaired Student's t -test. (I) Proposed model for the interplay between G4 and gene transcription. G4s are generated during transcription and mark transcription regulatory chromatin in the genome. They are correlated with active transcription and coincident with R-loops. However, disruption of G4 dynamics by G4 stabilizing ligands could alter the chromatin states by increasing the dsDNA melting and chromatin accessibility at the promoters, which impair the loading of general transcription factors, such as TBP, TFIID TAFs, and GTF2B. Thus, G4 stabilization leads to inhibition of transcription initiation and nascent RNA synthesis.

labeled core promoter DNA to form G4s, incubated the DNA with or without TMPyP4, and performed pulldown assays with nuclear extracts. We then performed mass spectrometry analysis to identify the proteins interacting with G4-forming core promoter DNA (Fig. 6A). Compared to the sample without TMPyP4 treatment, we identified fewer unique peptides for transcription initiation factors, such as GTF2B (TFIIB) and two TBP-associated factors (TFIID complex composed of TBP and TAF1-15 and Selective Factor 1 complex containing TBP and at least TAF1A, TAF1B, TAF1C, and TAF1D), in those treated with TMPyP4 (Fig. 6B). Distinct transcription initiation complexes are selectively required for different RNA polymerases (Zomerdijk et al. 1994; Hochheimer and Tjian 2003). The general transcription factors TFIID and GTF2B are crucial for mediating the selective and accurate promoter recognition of Pol II through protein–DNA and protein–protein interactions (Sainsbury et al. 2015; Roeder 2019). Therefore, loading of TFIID and GTF2B to core promoter dsDNA are key steps for the formation of the Pol II transcription preinitiation complex in a closed state (Cramer 2019). Thus, we performed ChIP-Rx with antibodies against TBP, TFIID subunit TAF1, and GTF2B after 1-h TMPyP4 treatment in HEK293T cells to check the effects of G4 stabilization on transcription preinitiation complex assembly. UCSC Genome Browser snapshots illustrated that G4 stabilization by TMPyP4 rapidly and substantially reduced the occupancy of TBP, TAF1, and GTF2B at *NPM1* and *SRSF4* loci (Fig. 6C,D), whereas promoter regions without G4-CUT&Tag signals were generally not occupied by general transcription factors (GSE178668). Genome-wide analysis with these ChIP-Rx peaks (Fig. 6E) revealed that TMPyP4 impaired the recruitment of TBP, TAF1, and GTF2B at their binding sites, indicating that TMPyP4-mediated G4 stabilization impairs the recruitment of general transcription factors to the chromatin.

Furthermore, we speculated that the general mechanism of impairment of transcription initiation could apply to the plasmid templates. To test this hypothesis, we cloned the promoters of *NPM1*, *SRSF4*, *HSPA8*, *H2AX*, and *SERBP1* genes, which showed native G4 signals in the G4-CUT&Tag analysis, into the firefly luciferase reporter plasmid pGL3, respectively. We tested the effects of G4 stabilizing drugs on these recombinant plasmids by measuring the luciferase activities (Fig. 6F). Compared to the vehicle-treated controls, TMPyP4 or PDS treatment inhibited the promoter activities of *NPM1*, *SRSF4*, *HSPA8*, *H2AX*, and *SERBP1* genes (Fig. 6F,G). ChIP-qPCR with the *SRSF4* recombinant reporter plasmid also showed that TMPyP4 treatment reduced the recruitment of POLR2A, TAF1, and GTF2B to the promoter (Fig. 6H), confirming that the general mechanism of transcription initiation inhibition applies to the plasmid templates. Together, these data demonstrate that disruption of G4 dynamics by G4 stabilizing ligands inhibits assembly of Pol II transcription preinitiation complex at promoter regions.

Discussion

In this study, we reported a native G4 profiling method, G4-CUT&Tag, by combining the G4-specific BG4 antibody with the sensitive CUT&Tag technique (Kaya-Okur et al. 2019; Wang et al. 2019). Compared with the well-known enrichment-based G4 ChIP-seq method, G4-CUT&Tag showed a higher resolution and profiles almost twice as many G4 peaks. It is likely attributed to the in situ capture strategy which leads to the lower signal loss and higher signal/noise ratio of the CUT&Tag technology. Moreover, the G4-CUT&Tag method is easier to perform and avoids many laborious and time-consuming steps. Specifically, this meth-

od does not require fixation, sonication, or immunoprecipitation and is convenient in sequencing library preparation due to the unique characteristics of Tn5 transposase. It only takes <24 h from collecting cells to the library preparation. This method allows starting with less than half a million cells, and this could be optimized further for even fewer cells, providing possibilities to profile the native G4s in the whole genome with limited starting materials.

In general, enhancers are associated with low transcriptional output, and gene body regions are highly dynamic due to transcription. Here, with this improved native G4 mapping method, we are able to identify native G4s at active enhancers, poised enhancers, and gene bodies, suggesting that G4-CUT&Tag is capable of capturing the dynamic and short-lived G4s at these sites. Novel G4s were able to be identified from enhancers by G4-CUT&Tag (Fig. 1E; Supplemental Fig. S2). G4s may regulate the activity of these identified key *cis*-regulatory elements which maintain gene expression programs. The sensitive G4-CUT&Tag could facilitate further studies of enhancer G4s in gene transcription, enhanceropathies, development, lineage specification, and pluripotency (Creyghton et al. 2010; Rada-Iglesias et al. 2011; Smith and Shilatifard 2014; Wang et al. 2015; Local et al. 2018).

Computational prediction of G4-CUT&Tag peaks demonstrates that a high association with G4 motifs and G4-CUT&Tag signals is generally correlated with observed G4 sequence signals by G4-seq (Marsico et al. 2019). Furthermore, G4-CUT&Tag peaks are well-correlated with R-loop CUT&Tag signals, and they share a matched distribution pattern, suggesting the G4 signals are bona fide. Additionally, native G4s in the genome show cell type specificity. In agreement with previous studies (Hänsel-Hertsch et al. 2016, 2020), G4s mark transcription regulatory elements and are associated with active transcription as evidenced from the further characterization of G4s with active epigenetic modifications and high Pol II occupancy. Moreover, G4s mark both active and poised enhancers with high H3K4me1/H3K27ac ratios, suggesting that G4s at the poised enhancers may participate in the maintenance of the potential to activate transcription and may be involved in development processes (Creyghton et al. 2010; Rada-Iglesias et al. 2011; Local et al. 2018). Distribution of G4 in transcription regulatory regions has been associated with chromatin relaxation, because the treatment of cells with the histone deacetylase inhibitor entinostat (Saito et al. 1999) stabilized histone H3K27 acetylation and led to an increase of G4 prevalence (Hänsel-Hertsch et al. 2016). Here, we found that CDK9 inhibitor flavopiridol treatment (Liang et al. 2015) for 1 h induced a rapid promoter-proximal Pol II pausing and enhanced native G4 formation at both promoter and enhancer regions (Fig. 3).

Dysfunction of G4s is involved in several human diseases, such as alpha-thalassemia X-linked intellectual disability (ATR-X) syndrome, C9orf72 amyotrophic lateral sclerosis and frontotemporal dementia (C9ALS/FTD), and breast cancer (Law et al. 2010; DeJesus-Hernandez et al. 2011; Renton et al. 2011; Shioda et al. 2018; Hänsel-Hertsch et al. 2020). Targeting G4s has emerged as the new type of DNA-targeted strategy, particularly for anticancer treatment (Neidle 2016; Zyner et al. 2019; Carvalho et al. 2020). The application of G4 ligands as potential anticancer drugs is currently being evaluated (Zimmer et al. 2016; Zyner et al. 2019), with some of them currently reaching advanced phase I and phase II trials as candidate therapeutic agents (Drygin et al. 2009; Xu et al. 2017; Carvalho et al. 2020). In this study, we found that stabilization of native G4 with G4 ligands rapidly changed the local chromatin states and that 1-h TMPyP4 treatment increased the

chromatin accessibility and altered the single-strand DNA conformation at promoter regions (Fig. 4; Supplemental Fig. S5). These findings suggest that G4s modulate biological processes such as transcription by altering the chromatin states. Indeed, TMPyP4 could inhibit Pol II occupancy at the promoter regions in a time-dependent manner and this effect could happen within 1 h, indicating that the mutual regulation of native G4, chromatin states, and transcription is a fast and dynamic process. G4 ligands are generally used for 1–2 d or even longer in cells to target G4s during anticancer strategies or when studying G4 function in telomere dysfunction and genome integrity (Hänsel-Hertsch et al. 2017; Carvalho et al. 2020). Thus, our studies indicate that it is critical to consider the rapid response of transcription changes and chromatin state alteration induced by native G4 stabilization when investigating functions of G4 ligands.

Although native G4s were coincident with active gene transcription, transient transcriptome analysis revealed that G4 stabilization rapidly decreased nascent synthesis of the gene body, antisense, and enhancer RNAs (Fig. 5; Supplemental Fig. S6). Mechanistically, we found that G4 stabilization inhibits transcription initiation by impairing the loading of general transcription factors to core promoters. This could be mediated through the impaired interaction of preinitiation factors with promoter DNA and the conformational changes of promoter DNA due to increased dsDNA melting (Fig. 6I). We observed impaired transcription preinitiation complex assembly at promoters and distribution of the majority of G4s at promoters and enhancers. The genome-wide analysis of qPRO-seq indicated that blocking Pol II elongation in the gene body is not likely to be a dominant function of native G4 in the genome (Fig. 5). Instead, mass spectrometry analysis with promoter DNA and ChIP-Rx analysis with TBP, TFIID subunit TAF1, and GTF2B antibodies suggested that blocking of transcription initiation plays a dominant role in gene transcription inhibition induced by G4 stabilizing agents.

Together, our studies have provided a more effective strategy for comprehensively profiling native G4s in the human genome and proposed a model for the interplay between G4 and gene transcription (Fig. 6I). G4s are generated after general transcription factor loading, Pol II recruitment, promoter melting, and Pol II elongation. G4s are correlated with active transcription and coincident with R-loops at the promoter regions. However, disruption of G4 dynamics by G4 stabilizing ligands modulates the chromatin states with enhanced chromatin accessibility and increases dsDNA melting as well as ssDNA opening. The alterations of chromatin states impair the loading of general transcription factors, such as TBP, TAF1, and GTF2B, to promoters, leading to inhibition of transcription initiation and nascent RNA synthesis (Fig. 6I). Our findings reveal the underlying mechanisms for the intertwinement of RNA polymerase II, chromatin states, and native G4s, provide a paradigm for functional studies of noncanonical DNA secondary structures, and advance our understanding of quadruplex-targeting therapies.

Methods

BG4 and BG4-EGFP antibodies

The scFv antibody, BG4, was purified using the expression vector pSANG10-3F-BG4 (Addgene plasmid no. 55756) according to the previous study (Hänsel-Hertsch et al. 2018). To get BG4-EGFP antibody, the EGFP coding sequence was amplified from the EGFP-Tpr (Addgene #35024) and subcloned into pSANG10-3F-BG4 to

generate the BG4-EGFP expression plasmid, which was chemically transformed into BL21 (DE3) competent cells. Expression of BG4-EGFP was induced by 0.2 mM IPTG for 14 h at 16°C, 200 rpm. BG4-EGFP protein was then purified with the Ni NTA beads 6FF (Smart-Lifesciences). See Supplemental Methods for further details.

G4 ChIP-seq

G4 ChIP-seq was performed according to a previously published protocol (Hänsel-Hertsch et al. 2018) with minor modifications. Briefly, a total of 12.5 μ L sheared chromatin was immunoprecipitated with 800 ng BG4 antibodies and preblocked 5 μ L Anti-DYKDDDDK affinity beads (Smart-Lifesciences) sequentially. After washes, the captured DNA was eluted and purified for library preparation with the NEBNext ultra II DNA library prep kit for Illumina. Experimental details and sequencing data analysis are described in Supplemental Methods.

G4-CUT&Tag

Recombinant pA-Tn5 and pA-Tn5 transposome assembly have been described in detail (Wang et al. 2021b). Cleavage under targets and tagmentation was performed as described before (Kaya-Okur et al. 2019; Wang et al. 2019) with some modifications. Briefly, cells were incubated with activated concanavalin A-coated magnetic beads (Smart-Lifesciences). The bead-bound cells were permeabilized and incubated first with BG4 primary antibody, followed by anti-Flag M2 antibody (Sigma-Aldrich), and lastly with rabbit antimouse IgG antibody. Diluted pA-Tn5 adapter complex was then added followed by the tagmentation reaction. Extracted DNA fragments were used for library preparation. More details of this experiment and data analysis are described in Supplemental Methods.

ChIP-seq and ChIP-Rx

Chromatin immunoprecipitation sequencing was performed with 1×10^7 cells as previously described (Liang et al. 2015). The sheared chromatin was immunoprecipitated with 10 μ g individual antibodies and 15 μ L preblocked Protein A/G beads (Smart-Lifesciences). Library was prepared using the NEBNext ultra II DNA library prep kit for Illumina. For ChIP-Rx experiments, sonicated human chromatin was spiked-in with 10%–30% mouse chromatin to quantitatively normalize across experiments. Experimental details and sequencing data analysis are included in Supplemental Methods.

KAS-seq, ATAC-seq, TT-seq, and qPRO-seq

KAS-seq was performed as reported (Wu et al. 2020a) with minor modifications. ATAC-seq was performed as previously described (Corces et al. 2017). TT-seq was performed as described previously (Schwalb et al. 2016; Liang et al. 2018) with minor modifications. qPRO-seq was performed as previously reported (Judd et al. 2020) with some modifications. For experimental details and data analysis, see Supplemental Methods.

Native bisulfite conversion assay

A native bisulfite conversion assay was performed as previously described (Zheng et al. 2017) using an EpiMark bisulfite conversion kit (NEB) with some modifications, as elaborated in Supplemental Methods.

Immunofluorescence and Stellaris RNA FISH

Immunofluorescence and Stellaris RNA FISH were performed as previously described (Raj et al. 2008; Liang et al. 2015). To detect G4 formation, HeLa cells were incubated with BG4-EGFP antibody and stained with DAPI (4', 6-Diamidino-2-Phenylindole, Dihydrochloride) (Thermo Fisher Scientific D1306). For Stellaris RNA FISH, HeLa cells were immunostained with specific Stellaris FISH probes followed by DAPI staining. Images were acquired with a 63 × 1.4 oil objective on a Leica SP8 STED confocal microscope. See the Supplemental Methods for details.

Reporter gene assay and ChIP-qPCR

Putative G4 forming regions in the promoters of *NPM1*, *SRSF4*, *HSPA8*, *H2AX*, and *SERBP1* that showed signal enrichments in G4-CUT&Tag were amplified and cloned into the pGL3 basic luciferase reporter vector (Promega) (primers are shown in Supplemental Table S1). For the luciferase reporter assay, HEK293T cells were transfected with these plasmids and treated with TMPyP4 or PDS followed by luciferase activity detection with the firefly luciferase assay kit (US Everbright Inc.).

For ChIP-qPCR assays, 1×10^7 HEK293T cells were used for immunoprecipitation with 5 µg of specific antibody and 15 µL of preblocked Dynabeads Protein G (Invitrogen). After extensive washes, the captured DNA was eluted, purified, and subjected to real-time PCR (qPCR) analysis. To avoid the interference of genomic DNA, the primers were designed to span across the insertion-vector junctions (Supplemental Table S1). For the input control, 10% of the sheared DNA was directly purified without immunoprecipitation and was used for qPCR with the same primers. More details are described in Supplemental Methods.

Liquid chromatography with electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS)

LC-ESI-MS/MS experiments were performed on a Thermo TSQ Quantis triple-stage quadrupole mass spectrometer under the positive ion mode. See the Supplemental Methods for details.

Coimmunoprecipitation of core promoter and nanoLC-FAMIS-MS/MS

A biotin-labeled DNA oligonucleotide (5'-GCGTGTAGTGCAC TTGGGCGCCTATAATAATGGGTGGGTGGGACATCTCATTCTT TGGGACGGGAGGGTTCATACATCTAGACGCG-biotin-3'), which was G-rich and contained core promoter elements, and its reverse complement DNA oligo were synthesized by Sangon Biotech. The G-rich oligo was annealed and mixed with its reverse complement DNA oligo to form G4-containing core promoter DNA. Preblocked streptavidin MagPoly beads (Smart-Lifesciences) were incubated with annealed double-stranded DNA with or without TMPyP4. These streptavidin MagPoly beads with core promoter DNA were then incubated with nuclear protein extract. After extensive washes, the captured proteins were eluted and digested via two-step digestion in urea buffer as previously reported (Keilhauer et al. 2015; Zhang et al. 2017). The acidified peptides were desalted with the C18 Stage tips (Thermo Fisher Scientific) and analyzed by an Orbitrap Exploris 480 mass spectrometer equipped with the FAIMS Pro interface. Samples were analyzed on an EASY-nLC system using a Hypersil GOLD C18 Selectivity HPLC column and 3-h preprogrammed gradients. Full MS resolutions were set to 60,000 at m/z 200 and Mass range was set to 350–1500. Raw files were processed with Proteome Discoverer 2.4 (Thermo Fisher Scientific) using a four-stage searching program. More details of

experiment and data analysis are described in Supplemental Methods.

Quantification and statistical analyses

Data are presented as mean ± SD. The sample sizes (*n*) in the figure legends indicate the number of replicates in each experiment and are provided in the corresponding figure legends. The peak or gene size (*N*) in the heat maps indicates the number of peaks or genes included. A two-tailed Mann–Whitney *U* test was used to determine the significant differences between the two groups (Fig. 3F; Supplemental Figs. S4E, S6F, S6H). Statistical analysis in Figure 6, F through H, was performed by unpaired Student's *t*-tests, and the *P* values were denoted in each figure.

Data access

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE178668. The mass spectrometry proteomics data generated in this study have been submitted to the ProteomeXchange Consortium (<http://www.proteomexchange.org>) via the PRIDE (Perez-Riverol et al. 2019) partner repository under the data set identifier PXD026830.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank Dr. Shankar Balasubramanian at the Department of Chemistry at the University of Cambridge for the pSANG10-3F-BG4 plasmid. We thank Dr. Steven Henikoff at the Division of Basic Sciences of Fred Hutchinson Cancer Research Center for providing the 3 × Flag-pA-Trn5-Fl plasmid. We also thank the staffs at the Research Center for Medicine and Structural Biology of Wuhan University for technical support and Dr. Xiaocheng Weng at the College of Chemistry and Molecular Science of Wuhan University for providing the N₃-kethoxal. This study was supported by the National Key Research and Development Program of China Stem Cell and Translational Research (No. 2019YFA0111100), the Thousand Youth Talents Plan award to K.L., and the Youth Program of National Natural Science Foundation of China (No. 32001049 and No. 82000114) awards to P.F. and Y.X.

Author contributions: K.L., C.L., J.H., and P.F. conceived and designed the experiments. C.L. and H.W. conducted most of the cell and biochemical experiments. K.L., Z.Y., and W.W. performed bioinformatics analyses. P.F. and R.X. performed the mass spectrometry and data analysis. Y.X. performed the immunostaining and confocal analysis. B.H. performed part of the immunoprecipitation and luciferase reporter gene assays. Q.L. established the KAS-seq method. P.F., K.L., and C.L. analyzed and interpreted results and wrote the manuscript with input from all authors. All authors contributed to editing the manuscript.

References

- Anantha NV, Azam M, Sheardy RD. 1998. Porphyrin binding to quadruplexed T₄G₄. *Biochemistry* **37**: 2709–2714. doi:10.1021/bi973009v
- Bedrat A, Lacroix L, Mergny JL. 2016. Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res* **44**: 1746–1759. doi:10.1093/nar/gkw006

- Biffi G, Tannahill D, McCafferty J, Balasubramanian S. 2013. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem* **5**: 182–186. doi:10.1038/nchem.1548
- Bochman ML, Paeschke K, Zakian VA. 2012. DNA secondary structures: stability and function of G-quadruplex structures. *Nat Rev Genet* **13**: 770–780. doi:10.1038/nrg3296
- Carvalho J, Mergny JL, Salgado GF, Queiroz JA, Cruz C. 2020. G-quadruplex, friend or foe: the role of the G-quartet in anticancer strategies. *Trends Mol Med* **26**: 848–861. doi:10.1016/j.molmed.2020.05.002
- Chen L, Chen JY, Zhang X, Gu Y, Xiao R, Shao C, Tang P, Qian H, Luo D, Li H, et al. 2017. R-ChIP using inactive RNase H reveals dynamic coupling of R-loops with transcriptional pausing at gene promoters. *Mol Cell* **68**: 745–757.e5. doi:10.1016/j.molcel.2017.10.008
- Chen MC, Tippiana R, Demeshkina NA, Murat P, Balasubramanian S, Myong S, Ferré-D'Amaré AR. 2018. Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature* **558**: 465–469. doi:10.1038/s41586-018-0209-9
- Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, et al. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* **14**: 959–962. doi:10.1038/nmeth.4396
- Cramer P. 2019. Organization and regulation of gene transcription. *Nature* **573**: 45–54. doi:10.1038/s41586-019-1517-4
- Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, et al. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci* **107**: 21931–21936. doi:10.1073/pnas.1016071107
- De Cian A, Cristofari G, Reichenbach P, De Lemos E, Monchaud D, Teulade-Fichou MP, Shin-Ya K, Lacroix L, Lingner J, Mergny JL. 2007. Reevaluation of telomerase inhibition by quadruplex ligands and their mechanisms of action. *Proc Natl Acad Sci* **104**: 17347–17352. doi:10.1073/pnas.0707365104
- DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, et al. 2011. Expanded GGGGCC hexanucleotide repeat in noncoding region of *C9ORF72* causes chromosome 9p-linked FTD and ALS. *Neuron* **72**: 245–256. doi:10.1016/j.neuron.2011.09.011
- De Magis A, Manzo SG, Russo M, Marinello J, Morigi R, Sordet O, Capranico G. 2019. DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc Natl Acad Sci* **116**: 816–825. doi:10.1073/pnas.1810409116
- Drygin D, Siddiqui-Jain A, O'Brien S, Schwabe M, Lin A, Bliesath J, Ho CB, Proffitt C, Trent K, Whitten JP, et al. 2009. Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis. *Cancer Res* **69**: 7653–7661. doi:10.1158/0008-5472.CAN-09-1304
- Duquette ML, Handa P, Vincent JA, Taylor AF, Maizels N. 2004. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev* **18**: 1618–1629. doi:10.1101/gad.1200804
- Gao J, Zybailov BL, Byrd AK, Griffin WC, Chib S, Mackintosh SG, Tackett AJ, Raney KD. 2015. Yeast transcription co-activator Sub1 and its human homolog PC4 preferentially bind to G-quadruplex DNA. *Chem Commun (Camb)* **51**: 7242–7244. doi:10.1039/C5CC00742A
- Hänsel-Hertsch R, Beraldi D, Lensing SV, Marsico G, Zyner K, Parry A, Di Antonio M, Pike J, Kimura H, Narita M, et al. 2016. G-quadruplex structures mark human regulatory chromatin. *Nat Genet* **48**: 1267–1272. doi:10.1038/ng.3662
- Hänsel-Hertsch R, Di Antonio M, Balasubramanian S. 2017. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat Rev Mol Cell Biol* **18**: 279–284. doi:10.1038/nrm.2017.3
- Hänsel-Hertsch R, Spiegel J, Marsico G, Tannahill D, Balasubramanian S. 2018. Genome-wide mapping of endogenous G-quadruplex DNA structures by chromatin immunoprecipitation and high-throughput sequencing. *Nat Protoc* **13**: 551–564. doi:10.1038/nprot.2017.150
- Hänsel-Hertsch R, Simeone A, Shea A, Hui WWI, Zyner KG, Marsico G, Rueda OM, Bruna A, Martin A, Zhang X, et al. 2020. Landscape of G-quadruplex DNA structural regions in breast cancer. *Nat Genet* **52**: 878–883. doi:10.1038/s41588-020-0672-8
- Hochheimer A, Tjian R. 2003. Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. *Genes Dev* **17**: 1309–1320. doi:10.1101/gad.1099903
- Judd J, Wojenski LA, Wainman LM, Tippens ND, Rice EJ, Dziubek A, Villafano GJ, Wissink EM, Vershuis P, Bagepalli L. 2020. A rapid, sensitive, scalable method for Precision Run-On sequencing (PRO-seq). bioRxiv doi:10.1101/2020.05.18.102277
- Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. 2019. CUT&tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* **10**: 1930. doi:10.1038/s41467-019-09982-5
- Keilhauer EC, Hein MY, Mann M. 2015. Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol Cell Proteomics* **14**: 120–135. doi:10.1074/mcp.M114.041012
- Law MJ, Lower KM, Voon HP, Hughes JR, Garrick D, Viprasakit V, Mitsun M, De Gobbi M, Marra M, Morris A, et al. 2010. ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell* **143**: 367–378. doi:10.1016/j.cell.2010.09.023
- Li PT, Wang ZF, Chu IT, Kuan YM, Li MH, Huang MC, Chiang PC, Chang TC, Chen CT. 2017. Expression of the human telomerase reverse transcriptase gene is modulated by quadruplex formation in its first exon due to DNA methylation. *J Biol Chem* **292**: 20859–20870. doi:10.1074/jbc.M117.808022
- Liang K, Woodfin AR, Slaughter BD, Unruh JR, Box AC, Rickels RA, Gao X, Haug JS, Jaspersen SL, Shilatifard A. 2015. Mitotic transcriptional activation: clearance of actively engaged Pol II via transcriptional elongation control in mitosis. *Mol Cell* **60**: 435–445. doi:10.1016/j.molcel.2015.09.021
- Liang K, Smith ER, Aoi Y, Stoltz KL, Katagi H, Woodfin AR, Rendleman EJ, Marshall SA, Murray DC, Wang L, et al. 2018. Targeting processive transcription elongation via SEC disruption for MYC-induced cancer therapy. *Cell* **175**: 766–779.e17. doi:10.1016/j.cell.2018.09.027
- Local A, Huang H, Albuquerque CP, Singh N, Lee AY, Wang W, Wang C, Hsia JE, Shiau AK, Ge K, et al. 2018. Identification of H3K4me1-associated proteins at mammalian enhancers. *Nat Genet* **50**: 73–82. doi:10.1038/s41588-017-0015-6
- Makowski MM, Gräwe C, Foster BM, Nguyen NV, Bartke T, Vermeulen M. 2018. Global profiling of protein–DNA and protein–nucleosome binding affinities using quantitative mass spectrometry. *Nat Commun* **9**: 1653. doi:10.1038/s41467-018-04084-0
- Marsico G, Chambers VS, Sahakyan AB, McCauley P, Boutell JM, Antonio MD, Balasubramanian S. 2019. Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res* **47**: 3862–3874. doi:10.1093/nar/gkz179
- Neidle S. 2016. Quadruplex nucleic acids as novel therapeutic targets. *J Med Chem* **59**: 5987–6011. doi:10.1021/acs.jmedchem.5b01835
- Niehrs C, Luke B. 2020. Regulatory R-loops as facilitators of gene expression and genome stability. *Nat Rev Mol Cell Biol* **21**: 167–178. doi:10.1038/s41580-019-0206-3
- Orlando DA, Chen MW, Brown VE, Solanki S, Choi YJ, Olson ER, Fritz CC, Bradner JE, Guenther MG. 2014. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep* **9**: 1163–1170. doi:10.1016/j.celrep.2014.10.018
- Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Ingugiari A, Griss J, Mayer G, Eisenacher M, et al. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**: D442–d450. doi:10.1093/nar/gky1106
- Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. 2011. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**: 279–283. doi:10.1038/nature09692
- Raiber EA, Kranaster R, Lam E, Nikan M, Balasubramanian S. 2012. A non-canonical DNA structure is a binding motif for the transcription factor SP1 *in vitro*. *Nucleic Acids Res* **40**: 1499–1508. doi:10.1093/nar/gkr882
- Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. 2008. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5**: 877–879. doi:10.1038/nmeth.1253
- Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, et al. 2011. A hexanucleotide repeat expansion in *C9ORF72* is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**: 257–268. doi:10.1016/j.neuron.2011.09.010
- Rodriguez R, Miller KM, Forment JV, Bradshaw CR, Nikan M, Britton S, Oelschlaegel T, Xhemalce B, Balasubramanian S, Jackson SP. 2012. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat Chem Biol* **8**: 301–310. doi:10.1038/nchembio.780
- Roeder RG. 2019. 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nat Struct Mol Biol* **26**: 783–791. doi:10.1038/s41594-019-0287-x
- Sainsbury S, Bernecky C, Cramer P. 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**: 129–143. doi:10.1038/nrm3952
- Saito A, Yamashita T, Mariko Y, Nosaka Y, Tsuchiya K, Ando T, Suzuki T, Tsuruo T, Nakanishi O. 1999. A synthetic inhibitor of histone deacetylase, MS-27-275, with marked *in vivo* antitumor activity against human tumors. *Proc Natl Acad Sci* **96**: 4592–4597. doi:10.1073/pnas.96.8.4592
- Schwab B, Michel M, Zacher B, Frühauf K, Demel C, Tresch A, Gagneur J, Cramer P. 2016. TT-seq maps the human transient transcriptome. *Science* **352**: 1225–1228. doi:10.1126/science.aad9841

- Shioda N, Yabuki Y, Yamaguchi K, Onozato M, Li Y, Kurosawa K, Tanabe H, Okamoto N, Era T, Sugiyama H, et al. 2018. Targeting G-quadruplex DNA as cognitive function therapy for ATR-X syndrome. *Nat Med* **24**: 802–813. doi:10.1038/s41591-018-0018-6
- Smith E, Shilatifard A. 2014. Enhancer biology and enhanceropathies. *Nat Struct Mol Biol* **21**: 210–219. doi:10.1038/nsmb.2784
- Spiegel J, Adhikari S, Balasubramanian S. 2020. The structure and function of DNA G-quadruplexes. *Trends Chem* **2**: 123–136. doi:10.1016/j.trechm.2019.07.002
- Varshney D, Spiegel J, Zyner K, Tannahill D, Balasubramanian S. 2020. The regulation and functions of DNA and RNA G-quadruplexes. *Nat Rev Mol Cell Biol* **21**: 459–474. doi:10.1038/s41580-020-0236-x
- Wang A, Yue F, Li Y, Xie R, Harper T, Patel NA, Muth K, Palmer J, Qiu Y, Wang J, et al. 2015. Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell* **16**: 386–399. doi:10.1016/j.stem.2015.02.013
- Wang Q, Xiong H, Ai S, Yu X, Liu Y, Zhang J, He A. 2019. CoBATCH for high-throughput single-cell epigenomic profiling. *Mol Cell* **76**: 206–216.e7. doi:10.1016/j.molcel.2019.07.015
- Wang E, Thombre R, Shah Y, Latanich R, Wang J. 2021a. G-Quadruplexes as pathogenic drivers in neurodegenerative disorders. *Nucleic Acids Res* **49**: 4816–4830. doi:10.1093/nar/gkab164
- Wang K, Wang H, Li C, Yin Z, Xiao R, Li Q, Xiang Y, Wang W, Huang J, Chen L, et al. 2021b. Genomic profiling of native R-loops with a DNA-RNA hybrid recognition sensor. *Sci Adv* **7**: eabe3516. doi:10.1126/sciadv.abe3516
- Wu T, Lyu R, You Q, He C. 2020a. Author correction: Kethoxal-assisted single-stranded DNA sequencing captures global transcription dynamics and enhancer activity in situ. *Nat Methods* **17**: 749. doi:10.1038/s41592-020-0881-1
- Wu T, Lyu R, You Q, He C. 2020b. Kethoxal-assisted single-stranded DNA sequencing captures global transcription dynamics and enhancer activity in situ. *Nat Methods* **17**: 515–523. doi:10.1038/s41592-020-0797-9
- Xu H, Di Antonio M, McKinney S, Mathew V, Ho B, O'Neil NJ, Santos ND, Silvester J, Wei V, Garcia J, et al. 2017. CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nat Commun* **8**: 14432. doi:10.1038/ncomms14432
- Zhang Y, Liu T, Meyer CA, Eickhout J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137. doi:10.1186/gb-2008-9-9-r137
- Zhang Y, Sun H, Zhang J, Brasier AR, Zhao Y. 2017. Quantitative assessment of the effects of trypsin digestion methods on affinity purification–mass spectrometry-based protein–protein interaction analysis. *J Proteome Res* **16**: 3068–3082. doi:10.1021/acs.jproteome.7b00432
- Zheng KW, He YD, Liu HH, Li XM, Hao YH, Tan Z. 2017. Superhelicity constrains a localized and R-loop-dependent formation of G-quadruplexes at the upstream region of transcription. *ACS Chem Biol* **12**: 2609–2618. doi:10.1021/acscchembio.7b00435
- Zheng KW, Zhang JY, He YD, Gong JY, Wen CJ, Chen JN, Hao YH, Zhao Y, Tan Z. 2020. Detection of genomic G-quadruplexes in living cells using a small artificial protein. *Nucleic Acids Res* **48**: 11706–11720. doi:10.1093/nar/gkaa841
- Zimmer J, Tacconi EMC, Folio C, Badie S, Porru M, Klare K, Tumiaty M, Markkanen E, Halder S, Ryan A, et al. 2016. Targeting BRCA1 and BRCA2 deficiencies with G-quadruplex-interacting compounds. *Mol Cell* **61**: 449–460. doi:10.1016/j.molcel.2015.12.004
- Zomerdijk JC, Beckmann H, Comai L, Tjian R. 1994. Assembly of transcriptionally active RNA polymerase I initiation factor SL1 from recombinant subunits. *Science* **266**: 2015–2018. doi:10.1126/science.7801130
- Zyner KG, Mulhearn DS, Adhikari S, Martinez Cuesta S, Di Antonio M, Erard N, Hannon GJ, Tannahill D, Balasubramanian S. 2019. Genetic interactions of G-quadruplexes in humans. *eLife* **8**: e46793. doi:10.7554/eLife.46793

Received February 28, 2021; accepted in revised form July 20, 2021.