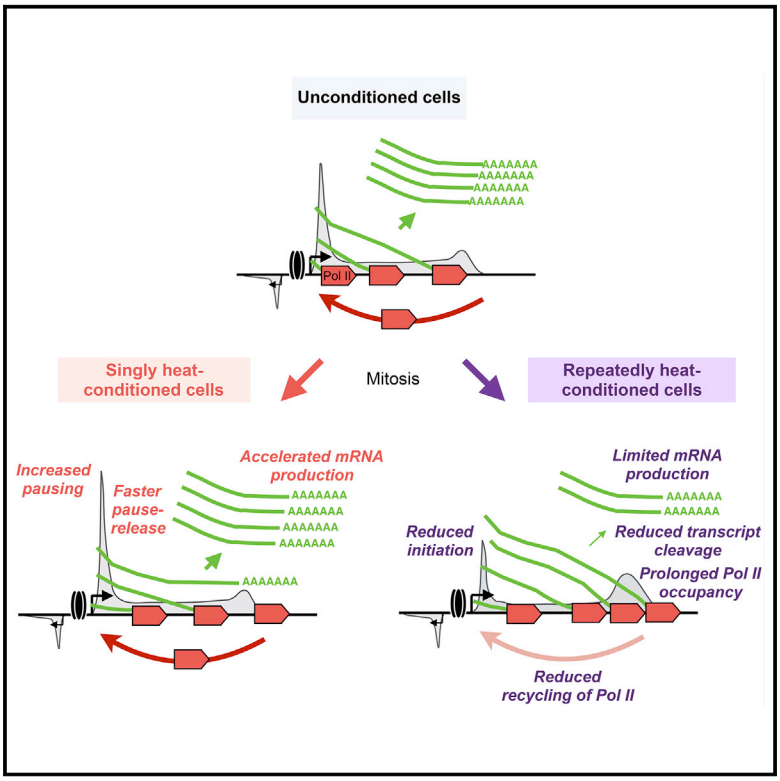
Molecular Cell

Article

# Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions

## Graphical abstract




## Highlights

- Cell-type-specific transcription precisely recovers after heat shock
- Daughters of stressed cells accelerate heat induction of quality control genes
- Daughters of repeatedly stressed cells refine basal and inducible transcription
- Transcriptional memory acts at Pol II pause release and transcription termination

## Authors

Anniina Vihervaara, Dig Bijay Mahat, Samu V. Himanen, Malin A.H. Blom, John T. Lis, Lea Sistonen

## Correspondence

viher@kth.se (A.V.),
jtl10@cornell.edu (J.T.L.),
lea.sistonen@abo.fi (L.S.)

## In brief

Heat shock reprograms transcription at genes and enhancers. Vihervaara et al. reveal that cells precisely restore transcription program after heat exposure and that cells establish a transcriptional memory of stress. In the daughters of stressed cells, increased Pol II pause release accelerates gene activation and reduced transcription termination decelerates mRNA processing.

Vihervaara et al., 2021, Molecular Cell *81*, 1–17
April 15, 2021 
https://doi.org/10.1016/j.molcel.2021.03.007

CellPress

Molecular Cell

CellPress

Article

# Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions

Anniina Vihervaara,[1,2,3,4,*] Dig Bijay Mahat,[3,5,6] Samu V. Himanen,[1,2,6] Malin A.H. Blom,[1,2] John T. Lis,[3,*] and Lea Sistonen[1,2,7,*]

[1]Faculty of Science and Engineering, Cell Biology, Åbo Akademi University, 20520 Turku, Finland
[2]Turku Bioscience Centre, University of Turku and Åbo Akademi University, 20520 Turku, Finland
[3]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA
[4]Department of Gene Technology, Science for Life Laboratory, KTH Royal Institute of Technology, 17165 Stockholm, Sweden
[5]Present address: Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[6]These authors contributed equally
[7]Lead contact
*Correspondence: viher@kth.se (A.V.), jtl10@cornell.edu (J.T.L.), lea.sistonen@abo.fi (L.S.)
https://doi.org/10.1016/j.molcel.2021.03.007

## SUMMARY

Heat shock instantly reprograms transcription. Whether gene and enhancer transcription fully recover from stress and whether stress establishes a memory by provoking transcription regulation that persists through mitosis remained unknown. Here, we measured nascent transcription and chromatin accessibility in unconditioned cells and in the daughters of stress-exposed cells. Tracking transcription genome-wide at nucleotide-resolution revealed that cells precisely restored RNA polymerase II (Pol II) distribution at gene bodies and enhancers upon recovery from stress. However, a single heat exposure in embryonic fibroblasts primed a faster gene induction in their daughter cells by increasing promoter-proximal Pol II pausing and by accelerating the pause release. In K562 erythroleukemia cells, repeated stress refined basal and heat-induced transcription over mitotic division and decelerated termination-coupled pre-mRNA processing. The slower termination retained transcripts on the chromatin and reduced recycling of Pol II. These results demonstrate that heat-induced transcriptional memory acts through promoter-proximal pause release and pre-mRNA processing at transcription termination.

## INTRODUCTION

Heat shock triggers transcription reprogramming, provoking an instant genome-wide change in RNA synthesis from genes and enhancers (reviewed in Vihervaara et al., 2018). Upon heat shock, hundreds of genes are rapidly induced by a potent *trans*-activator heat shock factor 1 (HSF1). Activated HSF1 binds to heat shock elements (HSEs) at architecturally primed promoters and enhancers (Rougvie and Lis, 1988; Rasmussen and Lis, 1993; Guertin and Lis, 2010; Vihervaara et al., 2013, 2017; Ray et al., 2019), and it can trigger the release of promoter-proximally paused RNA polymerase II (Pol II) into productive elongation (Duarte et al., 2016; Mahat et al., 2016). Concomitantly with the heat-induced escape of Pol II from the promoters of activated genes, thousands of genes are repressed via inhibition of the Pol II pause release. This restricted entry of Pol II into productive elongation causes the transcription machinery to accumulate at promoter-proximal regions of heat-repressed genes (Mahat et al., 2016; Vihervaara et al., 2017). As a consequence of the genome-wide re-coordination of Pol II pause release, heat-stressed cells promptly switch their transcription program to produce chaperones, reduce genome-wide transcription, and protect cellular integrity.

Stress responses are robustly activated and evolutionarily conserved to safeguard cells and organisms. Severe stress can have long-lasting consequences for an individual (Guan et al., 2002; Sailaja et al., 2012) and cause physiological changes over generations (Kaati et al., 2002; Wei et al., 2015; reviewed in Heard and Martienssen, 2014). The inheritance of physiological changes to many types of stresses has been described, but the cellular mechanisms that establish, maintain, and execute transcriptional memory remain poorly understood (reviewed in Perez and Lehner, 2019). Various stresses have been associated with long-term changes in the chromatin state (Guan et al., 2002; Tetievsky and Horowitz, 2010; Sailaja et al., 2012; D'Urso et al., 2016; Lämke et al., 2016; reviewed in D'Urso and Brickner, 2017) and shown to protect against protein misfolding by increasing chaperone expression (Gerner and Schneider, 1975; Maytin et al., 1990; Yost and Lindquist, 1991). However, stress-induced long-term changes in gene expression have been investigated

Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007



with steady-state RNA and protein analyses, which neither capture the processes of nascent transcription nor reveal the mechanistic control of Pol II. Thus, we do not yet know whether cells restore or adjust their program of nascent RNA synthesis when recovering from stress and whether regulation of Pol II at genes and enhancers encodes a memory of encountered stress.

Here, we provoked a genome-wide change in gene and enhancer transcription using heat shock and asked whether proteotoxic stress reprograms transcription and transcriptional responsiveness over mitotic divisions. We monitored nascent RNA synthesis at nucleotide resolution using precision run-on sequencing (PRO-seq) that provides genome-wide maps of transcription-engaged Pol II complexes at genes and enhancers (Kwak et al., 2013; Core et al., 2014; Vihervaara et al., 2017). By tracking engaged Pol II complexes through the rate-limiting steps of transcription, PRO-seq allows identification of regulatory decisions at high fidelity and spatiotemporal resolution (reviewed in Cardiello et al., 2020; Wissink et al., 2019). Simultaneously, changes in the chromatin accessibility were measured with an assay for transposase-accessible chromatin using sequencing (ATAC-seq) (Buenrostro et al., 2013). We used mouse embryonic fibroblasts (MEFs) and human K562 erythroleukemia cells that coordinate transcription upon heat shock with similar mechanisms (Mahat et al., 2016; Vihervaara et al., 2017) yet display different cellular identities, pathophysiological states, and stress sensitivities (Lozzio and Lozzio, 1975; Mivechi 1989; Ahn et al., 2001; Luft et al., 2001; Vihervaara et al., 2013; Elsing et al., 2014).

We found that transcriptional reprogramming by heat shock is followed by a precise restoration of basal-cell-type-specific transcription program within hours of recovery. In accordance, chromatin accessibility spread with transcription to heat-induced genes and enhancers and returned to pre-stress levels during the recovery. This transient transcriptional response to stress enabled us to investigate whether stress exposure establishes a transcriptional memory. In non-transformed MEFs, a single heat shock primed a subset of genes for an instant induction in the daughter cells. The faster responsiveness was established by increased promoter-proximal Pol II pausing and accelerated pause release upon an additional heat shock. In human K562 erythroleukemia cells, repeated stress exposures decreased transcription of genes for protein synthesis and increased transcription of pro-survival genes over mitotic division. The daughters of repeatedly heat-stressed cells also prolonged the residency of Pol II at the termination window of active genes, concurrently reducing transcript cleavage and recycling of Pol II to a new heat-induced initiation. These results uncovered promoter-proximal Pol II pausing, pause release, and transcription termination as the rate-limiting steps of transcription involved in establishing a memory over cell divisions.

## RESULTS

### Normalization of PRO-seq data to measure rapid transcription kinetics and prolonged transcription changes

We tracked the process of nascent transcription in acutely stressed cells, in cells recovering from stress, and in the daughters of stress-exposed cells using PRO-seq. PRO-seq is a highly sensitive method that maps engaged transcription complexes at nucleotide resolution across the genome (Kwak et al., 2013) and provides instant measures of rate-limiting regulatory steps at genes and enhancers upon transcriptional reprogramming (reviewed in Cardiello et al., 2020; Wissink et al., 2019). Because heat shock causes a global change in nascent transcription (reviewed in Vihervaara et al., 2018), robust normalization strategies are required to precisely quantify transcription between distinct stress conditions. We normalized the PRO-seq datasets of short (<1-h) heat shock (HS) kinetics using ends of over 150-kb-long genes, which provide sample-intrinsic normalization windows beyond the reach of acute heat-induced changes in transcription (Mahat et al., 2016; Vihervaara et al., 2017). For samples cultured more than 1 h under distinct conditions, we adopted a whole-genome spike-in strategy (Booth et al., 2018) and used *Drosophila* S2 cells as an extrinsic source of PRO-seq normalization counts (see STAR Methods). Accurate normalization was evident from the highly similar Pol II densities at gene bodies between biological replicate pairs (Figures S1 and S2) and close to identical transcription profiles of heat-unresponsive genes, as demonstrated in cells cultured several days under distinct conditions (Figures S1C and S2B).

### *De novo* identification of transcribed enhancers from nascent transcription profile

Active enhancers generally produce unstable and short enhancer RNAs (eRNAs) from divergent initiation regions (Core et al., 2014; Henriques et al., 2018; Mikhaylichenko et al., 2018; Tippens et al., 2018, 2020; Tome et al., 2018). The specific pattern of eRNA transcription is used for identification of transcribed enhancers *de novo* at high spatiotemporal resolution (Melgar et al., 2011; Azofeifa and Dowell, 2017; Vihervaara et al., 2017; Chu et al., 2018; Wang et al., 2019). There is no method for *in vivo* functional validation of all the computationally identified enhancers, but we confirmed that the putative enhancers that we identified from PRO-seq (http://www.dankolab.org/; Wang et al., 2019) precisely captured functionally verified enhancers of MYC (Fulco et al., 2016) and beta globin locus control element (Li et al., 2002; Song et al., 2007) in K562 cells (Figures S3A and S3B). The putative enhancers also contained the expected chromatin modifications (Figure S3C), and 76% of them localized to transcription-associated chromatin loops (Figure S3D). Our analyses strengthen and extend previous studies (Vihervaara et al., 2017; Henriques et al., 2018; Mikhaylichenko et al., 2018; Chu et al., 2018; Wang et al., 2019; Tippens et al., 2020), showing that promoter-distal transcription regulatory elements with divergently oriented Pol II include functional enhancers. For simplicity, we refer to the enhancer candidates identified from PRO-seq data as enhancers.

### Gene and enhancer transcription is precisely restored after an acute heat shock

A single heat shock induced hundreds and repressed thousands of genes and caused Pol II to accumulate at transcribed enhancers (Mahat et al., 2016; Vihervaara et al., 2017; Figure 1A). To address whether this heat-induced reprogramming of RNA

Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007



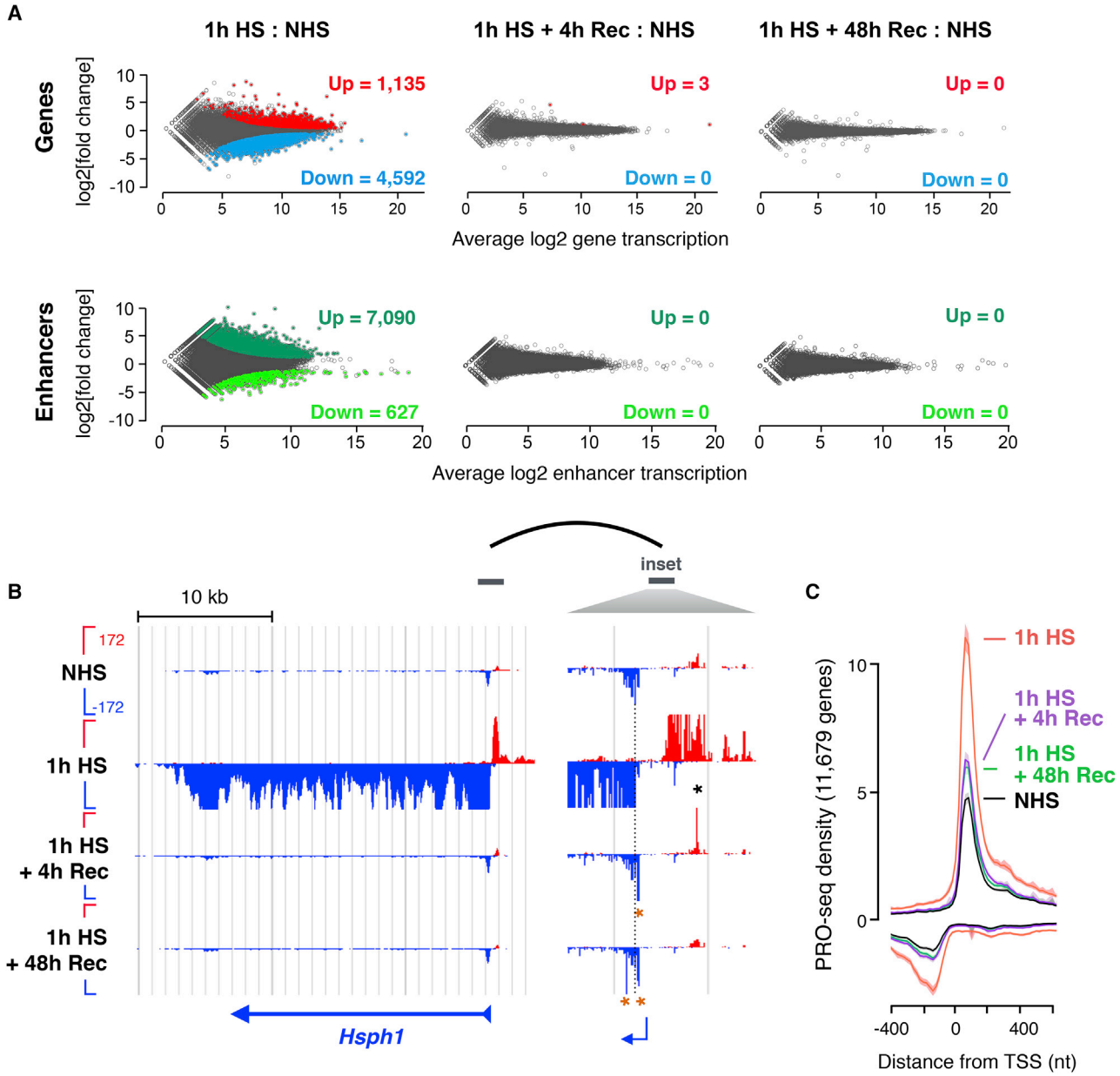



**Figure 1. Transcription of genes and enhancers precisely recovers after heat-induced reprogramming**

(A) Differential gene and enhancer transcription upon heat shock and recovery in MEFs. Up and down denote a statistically significant increase or decrease in Pol II density at gene bodies (upper panels) and enhancers (lower panels).

(B) Transcriptional profile of a heat-induced *Hsph1* gene in the non-heat-shock condition (NHS), upon a 1-h heat shock (HS), and upon recovery from a 1-h heat shock (Rec). Inset depicts promoter-proximal region. The dashed line indicates the highest Pol II pausing density in non-heat-shocked cells, and asterisks denote prominent Pol II pausing on sense (orange) and anti-sense (black) strand after recovery.

(C) Average promoter-proximal pausing measured at all transcribed genes. Shaded area: 12.5%–87.5% confidence interval.

The y axis in (B) is in linear scale from 172 to −172 for each track. See also Figures S1, S3, and S4.

synthesis is followed by restoration or readjustment of transcription, we measured nascent RNA synthesis in MEFs upon a 4- or 48-h recovery from a single 1-h heat shock (Figures 1 and S1A). We verified that the transiently heat-shocked MEFs continued to proliferate and did not undergo cell cycle arrest or apoptosis (Figure S4A). Moreover, Pol II levels remained constant throughout the experimentation (Figure S4B). Surprisingly, during only a 4-h recovery, the genome-wide profile of gene body and enhancer transcription was precisely restored to the level observed prior to the heat shock (Figures 1A and 1B). Despite the full recovery of transcription at enhancers and gene bodies, certain promoter-proximal regions gained new pause sites

Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007



(Figure 1B) or increased Pol II pausing at a single site (Figures S4C and S4D) during the recovery. Consequently, the genome-wide average of paused Pol II remained elevated, even when measured 48 h after the heat exposure (Figure 1C).

## Heat shock primes accelerated gene induction over mitotic divisions

Individual genes and whole transcription programs can be coordinated at the step of promoter-proximal pause release (Rougvie and Lis, 1988; Boettiger and Levine, 2009; Mahat et al., 2016; Vihervaara et al., 2017). To address whether the changed Pol II pausing in daughter cells alters genes' heat responsiveness, we preconditioned MEFs with a single 1-h heat shock, allowed a 48-h recovery, and measured transcription kinetics provoked by an additional heat shock. Instant and sustained changes in heat-induced transcription were assayed with PRO-seq upon 0, 12.5, 25, and 40 min of heat shock and by comparing the transcriptional stress response between unconditioned and preconditioned cells (Figures 2A and S1B). Analyses of productive elongation with DESeq2 (Love et al., 2014) showed clear differences in transcription upon 12.5 min of heat shock in preconditioned versus unconditioned cells (Figure S5A). Several genes, e.g., polyubiquitin-coding *Ubc* (Figure 2B) and metallothionein *Mt1* (Figure S5B), had gained a faster heat induction by preconditioning, whereas others, e.g., serum response factor (*Srf*), displayed a slower heat induction (Figure S5C). At *Ubc*, the promoter-proximal Pol II pausing was elevated after preconditioning (0-min inset in Figure 2B), and the paused Pol II was released faster into elongation upon heat shock (12.5-min inset in Figure 2B). At *Mt1*, prominent Pol II pausing was detected upon 12.5 min of heat shock only in unconditioned cells, although in preconditioned cells, it was actively elongating at all time points (Figure S5B). Noteworthy is that unconditioned cells also gained efficient Pol II pause release and high heat-induced transcription after 12.5 min of heat shock (Figures 2B, S5A, and S5B), indicating that preconditioning accelerated the onset of heat-shock-induced transcription.

## Faster pause release accelerates gene induction in preconditioned cells

More than 400 heat-activated genes displayed an accelerated induction after preconditioning, measured as a significant increase in productive elongation upon 12.5 min of heat shock (Figure S5A). To investigate whether the increased Pol II density on the gene bodies could be explained by changes in initiation, pausing, or pause release, we monitored Pol II progression through the promoter-proximal region. At genes with accelerated induction, the average Pol II pausing was similar between unconditioned and preconditioned cells upon 12.5 min of heat shock, but more Pol II had escaped into productive elongation in preconditioned cells (Figures 2C and S5A). In comparison, preconditioning did not change Pol II progression through the pause at genes that were highly (Figure 2C) or early (Figure S5D) heat induced in unconditioned cells. The faster progression of Pol II through the promoter-proximal region at a subset of genes revealed that preconditioning produced a transcription memory that primed a selected set of genes for a more-rapid heat activation.

A faster entry of Pol II into productive elongation can be accomplished by an accelerated onset of *trans*-activation, as demonstrated at *Ubc* (Figure 2B) and *Mt1* (Figure S5B), or by a faster moving Pol II. At over a 100-kb-long *vinculin* (*Vcl*) gene (Figure 2D), the wave of productive elongation extended tens of kb both in unconditioned and preconditioned cells upon a 12.5-min heat shock, showing an instant *trans*-activation, regardless of the preconditioning. Intriguingly, the elongation wave had proceeded farther at *Vcl* in preconditioned cells (Figure 2D), indicative of a faster moving Pol II. In agreement, Pol II density at the pause of *Vcl* (insets in Figure 2D) was lower in preconditioned cells, which demonstrates a shorter residence time of Pol II at the pause region before entering into productive elongation. Regardless of whether a gene gained accelerated induction due to a faster onset of *trans*-activation, faster moving Pol II through promoter-proximal region and gene body, or their combination, our results uncover the promoter-proximal pause regulation as a mechanistic step for enabling an accelerated heat induction.

## Single heat shock preconditioning accelerates induction of quality control genes

Genes that gained a faster heat induction by preconditioning were enriched for lysosomal, autophagocytosis, and membrane-associated functions (Figure S5E). These genes encode a machinery for clearing damaged organelles and proteins through lysosomal degradation (reviewed in (Kroemer et al., 2010). In comparison, genes that were highly or early induced, regardless of the preconditioning, encoded chaperones, cytoskeletal components, and negative regulators of transcription (Figure S5E). Hence, preconditioning MEFs with a single heat shock primed the lysosomal pathway of quality control for instant transcriptional activation, a pathway that complements the chaperone-mediated combating of proteotoxic stress.

## Human K562 cancer cells restore basal and heat-induced transcription after a single heat shock

The proliferation and integrity of cancer cells are regularly challenged by both intracellular and extracellular stresses (reviewed in Hanahan and Weinberg, 2011; Chen and Xie, 2018). To understand the transcriptional mechanisms by which cancer cells adapt to stress, we moved from stress-sensitive untransformed MEFs to human K562 erythroleukemia cells. K562 cells are a patient-derived malignant cancer cell line (Lozzio and Lozzio, 1975; Koeffler and Golde, 1980), known to tolerate extended heat treatments and develop thermotolerance (Mivechi, 1989; Vihervaara et al., 2013). Preconditioning K562 cells with a single heat shock recapitulated the instant heat-induced reprogramming of transcription (Figure S6A) and the precise restoration of cell-type-specific transcription program upon a 48-h recovery (Figure S6B), alike MEFs (Figure 1A). Furthermore, the daughters of cells exposed to a single heat stress displayed an unaltered stress response by inducing and repressing virtually the same set of genes (Figures S6C and S7A) and with strikingly similar kinetics (Figures S7A and S7B) as their parental cells. The similar stress responses in unconditioned and singly preconditioned K562 cells may reflect the constitutive stress response in cancer cells (Mivechi 1989; Leppä et al., 2001; Chatterjee and Burns, 2017; Klimczak et al., 2019).

Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007





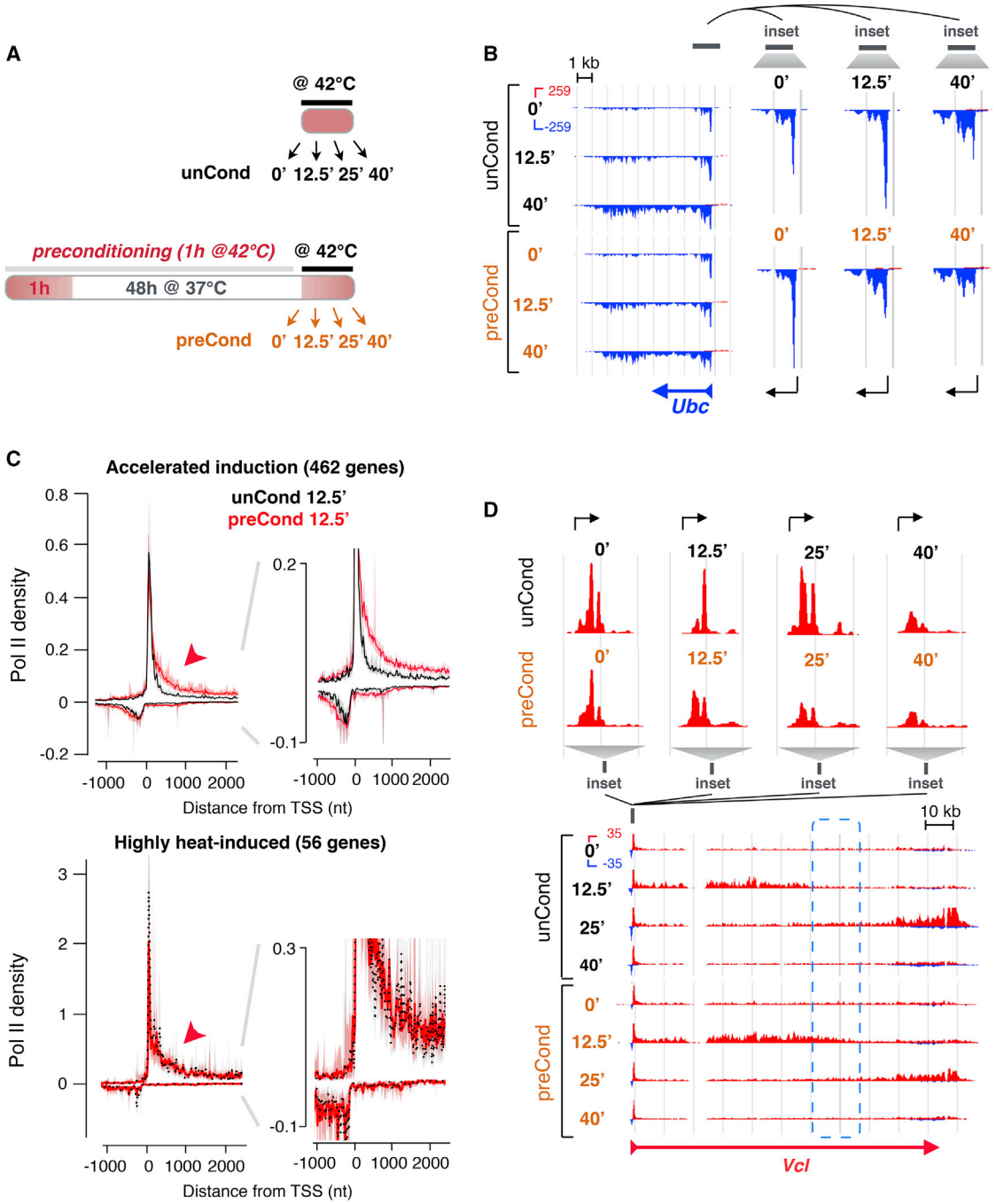



**Figure 2. A single heat shock primes accelerated gene induction over mitotic divisions**

(A) Experimental setup for measuring transcription kinetics in MEFs. Transcription was analyzed upon heat shock in unconditioned cells (unCond, upper panel) and in cells that were preconditioned with a 1-h heat shock and 48-h recovery (preCond, lower panel).

(B) Nascent transcription at *Ubc* in unconditioned and preconditioned cells. Insets show Pol II density in promoter-proximal region in unconditioned (upper panels) and preconditioned (lower panels) cells.

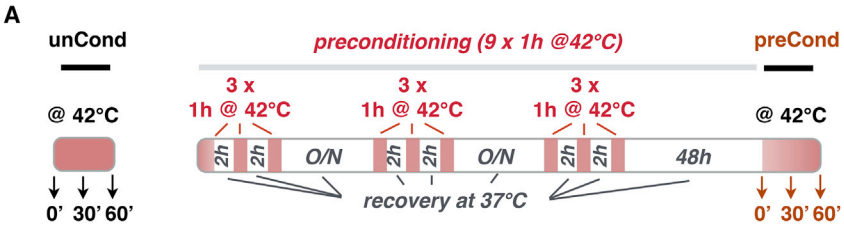
(C) Average intensity of promoter-proximally engaged Pol II upon a 12.5-min heat shock at genes that gain a faster heat induction by preconditioning (upper panel) or at genes that are highly heat-induced regardless of preconditioning (lower panels). Pol II density after the pause release is indicated with an arrowhead. Shaded area: 12.5%–87.5% confidence interval.

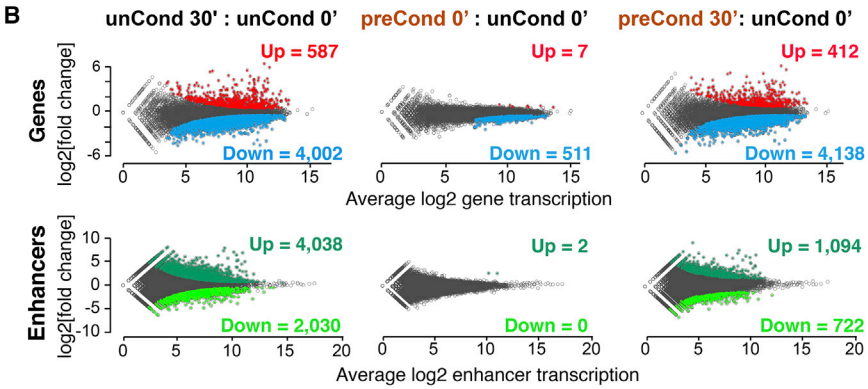
(D) Heat-induced wave of transcription along *Vcl* gene. The blue dashed region indicates an advancing wave of transcription that has proceeded farther in preconditioned than in unconditioned cells upon a 12.5-min heat shock. Insets show promoter-proximal Pol II density in unconditioned (upper panels) and preconditioned (lower panels) cells. Gray vertical lines in insets mark 100-nt intervals.

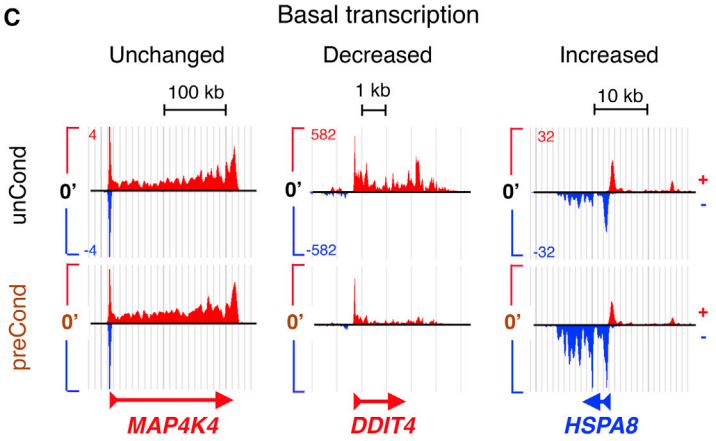
See also Figures S1, S4–S6, and S16.

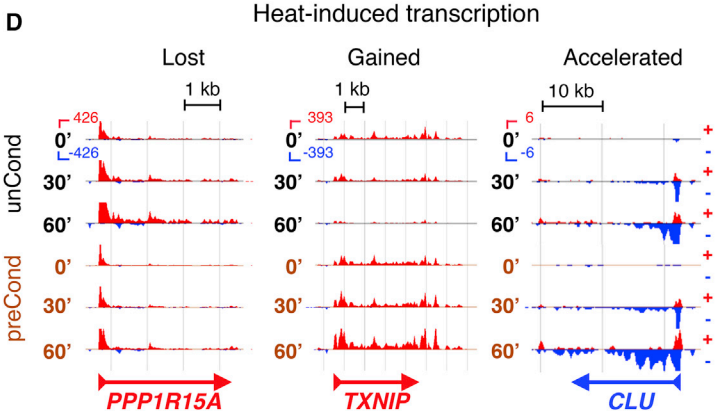
Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007

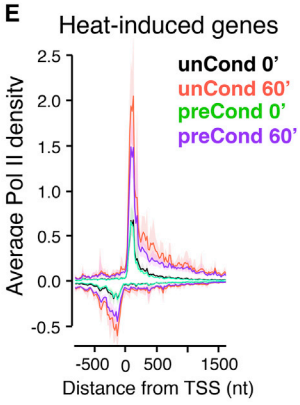





(legend on next page)

## Multiple heat shocks reprogram basal transcription in a cancer cell line

Pathophysiological stresses caused by cancer and neurodegeneration are often sustained or repeated. To investigate whether repeatedly encountered stress affects gene and enhancer transcription, we preconditioned K562 cells with a total of nine 1-h heat shocks during 3 consecutive days. After a 48-h recovery, the basal transcription in daughter cells and their transcriptional response to an additional single heat shock was measured (Figure 3A). K562 cells proliferated throughout the 6 days of preconditioning, recovery, and additional heat shock (Figure S8A) without showing signs of apoptosis or increased polyploidy (Figure S8B). PRO-seq datasets were normalized using whole-genome spike-in (Figures S8C and S8D), and Pol II protein levels were verified to remain constant during the experiments (Figure S8E).

Following the recovery from nine heat shocks, the vast majority of genes and virtually every enhancer had restored their transcription to a level detected in unstressed cells (Figures 3B and S9A), including lineage-specific regulators GATA-binding factor (GATA) and TAL BHLH transcription factor 1 (TAL1) (Fujiwara et al., 2009; Wu et al., 2014; Huang et al., 2016). However, preconditioning with several heat shocks caused elevated synthesis of seven genes and reduced synthesis of over 500 genes (Figures 3B and 3C). The most prominent increase in basal transcription was detected for *HSPA8* (Figure 3C) that encodes HSP70 cognate (HSC70), a constitutively expressed chaperone important for protein homeostasis (Ingolia and Craig, 1982; Kampinga et al., 2009). Genes with repressed basal transcription encode regulators of protein production and maturation (Figures S9B and S9C; Data S1), suggesting a slower protein production in the daughters of repeatedly stressed cancer cells.

## Repeated stress rewires heat inducibility

Subjecting the daughters of repeatedly stressed cells to an additional heat shock revealed that some genes had lost, gained, or accelerated heat induction due to preconditioning (Figure 3D). One of the genes that had lost heat induction encodes protein phosphatase 1 regulatory subunit 15A (PPP1R15A alias GADD34; Figure 3D), which is a key regulator of translation and maintains protein production in stressed cells (Harding et al., 2009; Walter and Ron, 2011). Genes with accelerated heat induction included *clusterin* (*CLU*) (Figure 3D), a glycosylated chaperone that facilitates autophagy, ameliorates endoplasmic reticulum (ER) stress, and enhances cancer cell survival (Zhang et al., 2014). The few genes that had gained heat induction encode proteins with functions in cell survival and growth arrest (Data S1). We did not detect activation of apoptotic pathways or changes in cell cycle regulators (Figure S9D; Data S1). This underscores the survival potential of K562 cancer cells throughout the series of protein-damaging stress (Figures S8A and S8B), an adaptation that involves altering the transcription program to maintain homeostasis.

## Repeated heat shocks reduce initiation and prolong termination over mitotic divisions

The most striking change in transcription in the daughters of repeatedly preconditioned cells was a global reduction in Pol II density at the promoter-proximal regions of heat-activated genes (Figures 3E and 4A). In PRO-seq, the 3′ end of each read reports the genomic position of transcribing Pol II, and it is used for mapping the active sites of transcription. Instead, the 5′ ends of PRO-seq reads are enriched at the initiating base of each transcript, providing a readout for the usage of transcription start sites (TSSs). Visualizing the 5′ ends of PRO-seq reads revealed that initiation was severely declined at heat-induced genes after preconditioning (Figure 4B). In comparison, distribution of the 3′ ends of PRO-seq reads showed that both the pausing and the pause release followed a similar course (Figure 4B). The reduction in heat-induced initiation in preconditioned cells occurred concurrently with an increase in Pol II density at the termination window (Figures 4A–4D). Indeed, the more actively the gene was transcribed upon heat shock, the more Pol II accumulated at the termination window (rho 0.55) and the less Pol II was engaged at the gene's promoter-proximal region (rho −0.43) in preconditioned cells (Figure 4C). The increased Pol II density in preconditioned cells was confined to 5,000 nt downstream of the cleavage and polyadenylation site (CPS) (Figure 4D). This local confinement of Pol II at the termination window differs from previously described run-through transcription that has been detected under stress conditions (Vilborg et al., 2017). Although the run-through transcription can extend tens of kb downstream of CPS (median 8.9 kb) and does not locally confine Pol II to CPS (Vilborg et al., 2017; Figure S9D), the daughters of repeatedly preconditioned cells accumulated Pol II at the termination window (Figures 4A–4D).

## Chromatin accessibility spreads from primed promoters to heat-induced genes

Promoter architecture primes genes for heat activation (reviewed in Vihervaara et al., 2018), and changes in epigenetic landscape have been coupled to transcriptional memory (reviewed in D'Urso and Brickner, 2017). To study whether the compromised Pol II progression through genes in repeatedly

**Figure 3. Repeated heat shocks refine basal and inducible transcription in daughter cells**

(A) Experimental setup for preconditioning human K562 cells with multiple heat shocks. Nascent transcription was measured upon heat shock in unconditioned cells (unCond, left panel), and in cells that were pre-exposed to nine 1-h heat shocks during 3 consecutive days and allowed to recover for 48 h prior to an additional 1-h heat shock (preCond, right panel).

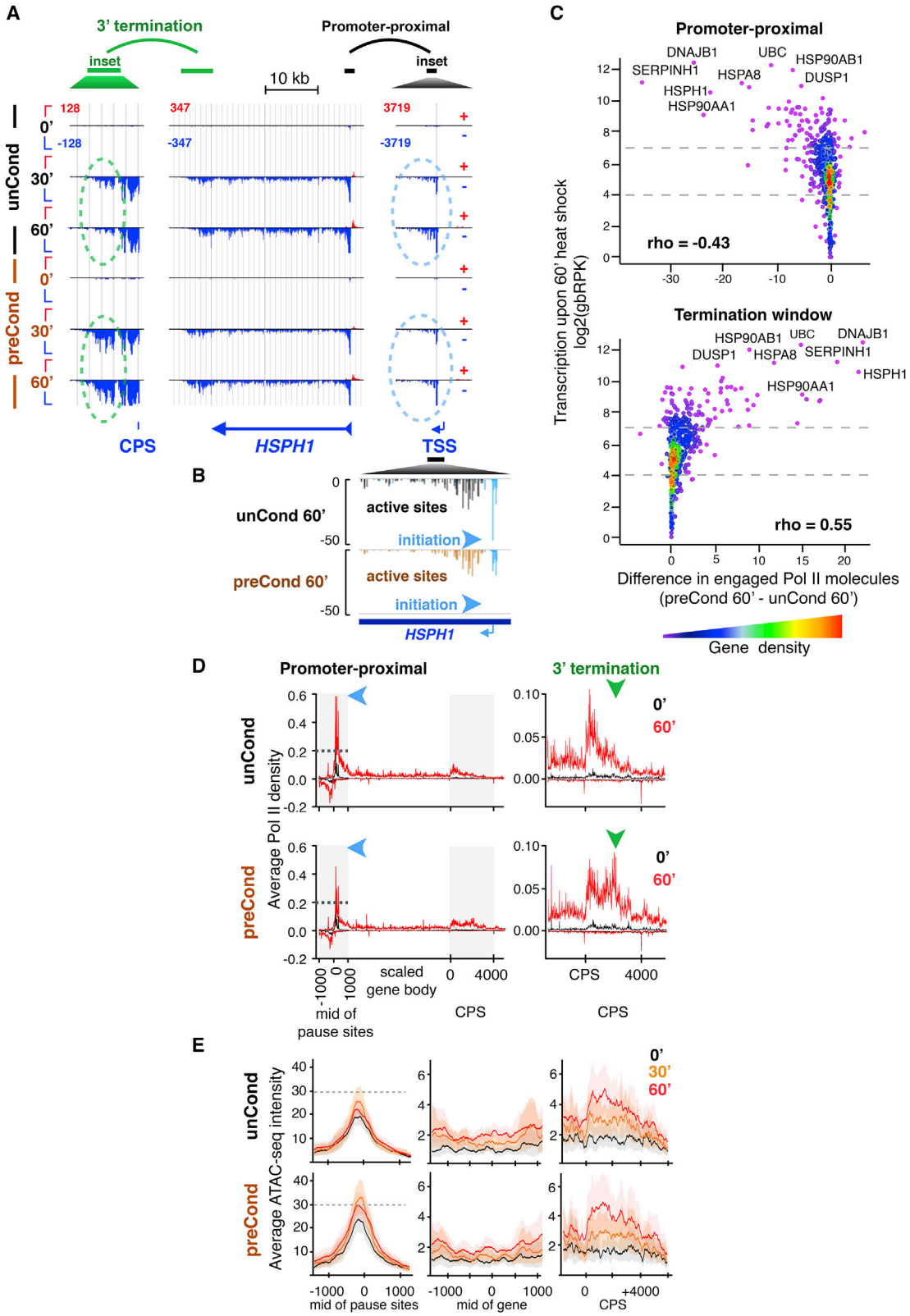
(B) Differential gene and enhancer transcription. Up and down denote numbers of genes (upper panels) and enhancers (lower panels) with significantly heat-induced or heat-repressed Pol II density, as measured against non-stressed cells (unCond 0').

(C) Transcriptional profiles of genes with unchanged (left panel), decreased (middle panel), or increased (right panel) basal transcription due to preconditioning.

(D) Transcriptional profiles of genes that had lost (left panel), gained (middle panel), or accelerated (right panel) heat induction as a result of preconditioning.

(E) Average Pol II density at promoter-proximal regions of heat-induced genes. Shaded area: 12.5%–87.5% confidence interval.

See also Figures S3 and S6–S9.




*(legend on next page)*

Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007





stressed cells was coupled to altered chromatin accessibility, we performed ATAC-seq (Buenrostro et al., 2013) in unconditioned, singly preconditioned, and repeatedly preconditioned K562 cells (Figures S10 and S11A). Measuring chromatin accessibility prior to and upon heat shock revealed that chromatin accessibility spread with transcription into heat-induced genes and that, upon recovery, the chromatin accessibility was restored to pre-stress levels (Figures 4E, S11B, S11C, and S12A). However, ATAC-seq found only minor, if any, changes in the chromatin due to preconditioning (Figures 4E, S11B, S11C, and S12A–S12D). Particularly, at genes with the highest heat induction, the difference in Pol II densities was pronounced between unconditioned and repeatedly preconditioned cells (Figure 4D), but the corresponding average ATAC-seq densities showed no significant differences (Figure 4E). Only a few genes with the most remarkable changes in Pol II progression displayed minor changes in chromatin accessibility due to preconditioning (Figures S11B, S11C, S12C, and S12D).

Chromatin state could change without detectable differences in transposase accessibility. Therefore, we performed Micrococcal Nuclease (MNase)-coupled chromatin immunoprecipitation (MNase-ChIP) (Skene and Henikoff, 2015) to quantify the levels of histones H2.AZ, H3, and H4, as well as histone H4 acetylation (H4ac) at the promoters and +1 nucleosomes of *HSPA1A* and *HSPH1*. In accordance with our ATAC-seq results and previous studies (Petesch and Lis, 2008; Mueller et al., 2017), chromatin accessibility increased at the +1 nucleosomes upon stress-induced activation (Figures S12E and S12F). However, we did not find clear differences in the histone levels between unconditioned and repeatedly preconditioned cells either under basal or heat-induced conditions (Figures S12E and S12F).

## Reduced initiation in preconditioned cells occurs in the presence of HSF1

Heat-induced *trans*-activation of primed genes requires strong transcription factors, such as HSF1 (reviewed in Vihervaara et al., 2018). To investigate whether a deficiency in HSF1 reduced initiation at heat-induced genes, we analyzed the expression and DNA-binding ability of HSF1. The transcription (Figure S13A), mRNA expression (Figure S13B), and protein levels (Figure S13C) of HSF1 were comparable in unconditioned and repeatedly preconditioned K562 cells. The binding of HSF1 to the promoters of *HSPA1A* and *HSPH1* was also similar in unconditioned and preconditioned cells (Figures 5A and S13D). Despite the uncompromised capacity of HSF1 to bind to its *cis*-acting elements, the RNA synthesis of *HSPA1A* and *HSPH1* was severely reduced, as were the levels of their corresponding mature mRNAs in preconditioned cells (Figures 5A and S13D). These results coupled the reduced initiation of heat-activated genes (Figures 4A–4D) to their lower mRNA expression (Figures 5A and S13D). Furthermore, the reduced initiation in an open chromatin environment and in the presence of a potent *trans*-activator manifested that the key step for decreased heat activation resided upstream of the promoter architecture and HSF1 binding, i.e., at the level of Pol II recruitment.

## Initiation and chromatin opening are abated at heat-induced enhancers after preconditioning

We depleted K562 cells of HSF1 (Figure 5B) and identified over 200 genes and close to 500 enhancers that were heat induced in an HSF1-dependent manner (Figures 5C–5F, S13E, and S13F; Data S2). In addition to *trans*-activating genes by binding to their promoters (reviewed in Vihervaara and Sistonen, 2014), the ability of HSF1 to *trans*-activate genes from enhancers became evident. At the *Tax1 binding protein 1* (*TAX1BP1*) locus, HSF1 only bound to a divergently transcribed enhancer 4.5 kb upstream of the promoter (Figures 5C and 5D), but it was essential for the heat-induced eRNA transcription and for the release of paused Pol II from the *TAX1BP1* promoter (Figures 5C–5E). Importantly, in repeatedly preconditioned cells, the heat-induced recruitment of Pol II to the HSF1-dependent enhancers was diminished (Figure 5F), uncovering a globally decreased initiation at heat-induced promoters and enhancers (Figures 4C, 4D, and 5C–5F). The reduced transcription at HSF1-dependent enhancers after preconditioning was recapitulated in the ATAC-seq data (Figure 5G); transcription-coupled chromatin opening did not occur at HSF1-activated enhancers in repeatedly preconditioned cells, although it was detected in unconditioned and singly preconditioned cells (Figures 5G and S14). In comparison, highly transcribed enhancers showed similar chromatin accessibility, regardless of preconditioning or heat shock (Figures 5G and S14).

## Pol II accumulates at the termination window of actively transcribed genes

The reduced initiation in preconditioned cells prompted us to quantify the distribution of transcription complexes across the genome. We counted engaged Pol II molecules at distinct genomic regions (Figures S15A and S15B) and found an

**Figure 4. Initiation is reduced and termination prolonged in the daughters of repeatedly stressed K562 cells**

(A) Nascent transcription along *HSPH1* showing reduced density of engaged Pol II at the promoter-proximal region (light blue dashed circle) and increased density downstream of the cleavage and polyadenylation site (CPS) (green dashed circle) in preconditioned cells.

(B) Promoter-proximal region of *HSPH1* showing active sites of transcription (3′ end of each PRO-seq read) and initiation intensity (5′ end of each PRO-seq read). Arrowheads compare the intensity of initiation; the arrow denotes transcription initiation site upon heat shock.

(C) Transcriptional activity of heat-induced genes (n = 587) compared with the change in Pol II density at promoter-proximal region (0 to +1,000 from TSS, upper panel) or termination window (+100 to +6,000 from CPS, lower panel) due to preconditioning. Spearman's rank correlations (rho) and the most affected genes are indicated.

(D) Average density of engaged transcription complexes along highly heat-induced genes in unconditioned (upper panels) and preconditioned (lower panels) cells. Light blue arrowheads indicate the promoter-proximal Pol II; green arrowheads show the site of increased Pol II engagement in preconditioned cells.

(E) Average ATAC-seq density at highly heat-induced genes in unconditioned (upper panels) and preconditioned (lower panels) cells. Shaded area: 12.5%–87.5% confidence interval.

See also Figures S10–S12.

Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007



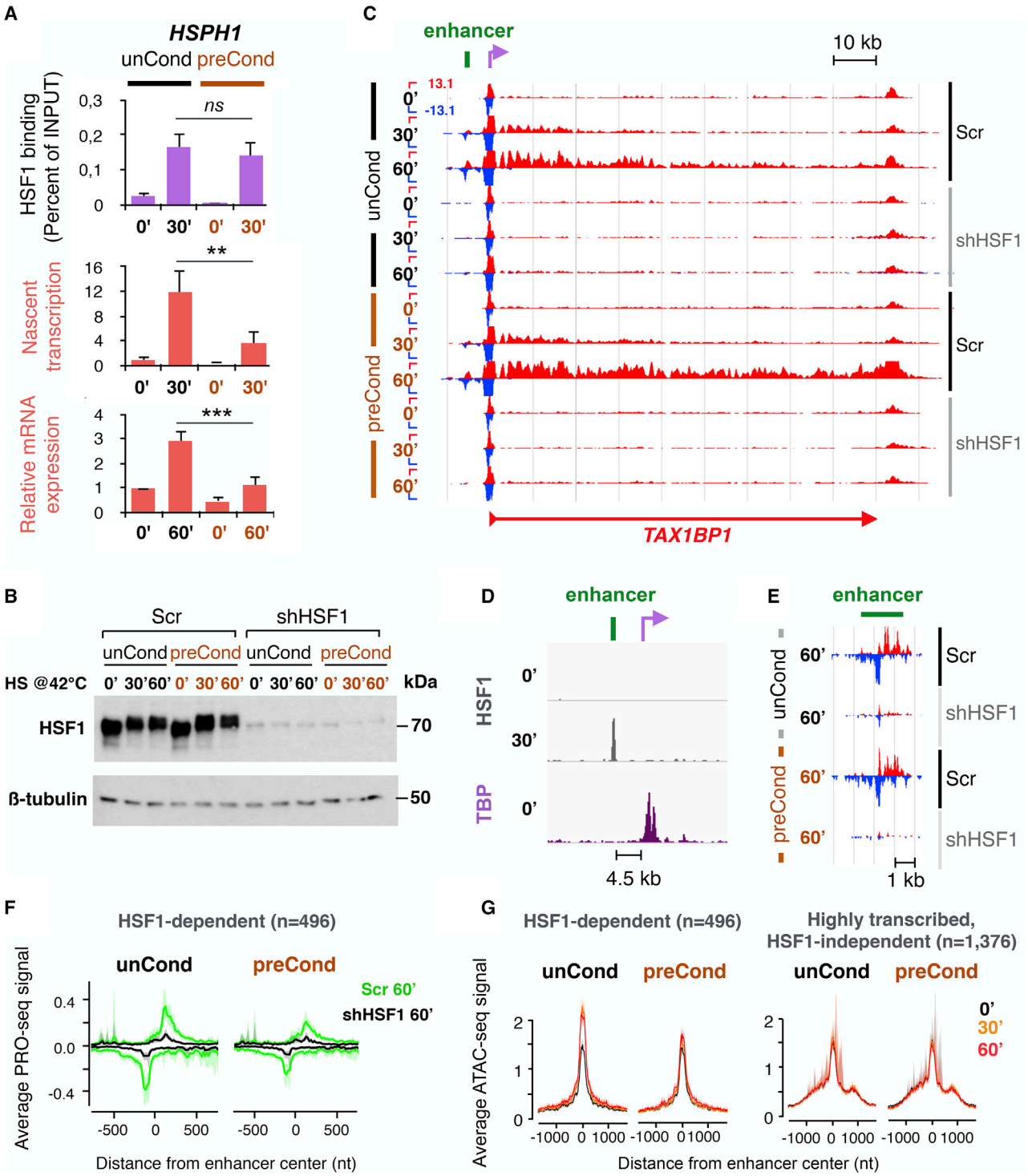



**Figure 5. HSF1 *trans*-activates genes via promoters and enhancers**

(A) HSF1-binding intensity to the *HSPH1* promoter (uppermost panel), nascent transcription of *HSPH1* as measured from the first intron (middle panel), and relative level of polyA-containing *HSPH1* mRNA (bottom panel) in unconditioned and preconditioned K562 cells. **p < 0.05; ***p < 0.005. The error bars indicate standard deviations (n = 3).

(B) HSF1 protein expression in scrambled-transfected (Scr) and HSF1-depleted (shHSF1) K562 cells.

(C) Transcription of *TAX1BP1* and its upstream enhancer in the presence and absence of HSF1.

*(legend continued on next page)*

accumulation of Pol II at the termination window of actively transcribed genes (Figures S15B–S15E). Over 400 genes simultaneously reduced engagement of Pol II at the promoter-proximal region and increased Pol II engagement at the termination window in the daughters of repeatedly stressed cells (Figure S15D). These genes were characterized by high nascent transcription upon heat shock and included many heat-repressed genes that retained active transcription during heat stress (Figures S15B–S13E).

### Repeated heat shocks reduce transcript cleavage and recycling of Pol II

To understand why Pol II accumulated downstream of CPS in preconditioned cells, we examined the processing of transcripts at the termination window. At CPS, the nascent transcript is cleaved, exposing an uncapped 5′ end of the RNA (Figure 6A). The uncapped 5′ end of the nascent transcript is then targeted by 5′-3′ Exoribonuclease 2 (XRN2), which chases down Pol II and terminates transcription (reviewed in Proudfoot, 2016; Wissink et al., 2019). Thus, mapping the 5′ ends of Pol-II-associated transcripts at the termination window can provide a readout for transcript cleavage (Figure 6A). For example, the robustly heat-induced *DNAJB1* gene displayed a clear decrease in initiation and a profound accumulation of Pol II at the termination window after preconditioning (Figures 6B and 6C). In unconditioned cells, the 5′ ends of PRO-seq reads demonstrated a prominent cleavage at the annotated CPS of *DNAJB1* (Figure 6C). In preconditioned cells, the cleavage site had shifted downstream to a single site at the end of the termination window (Figure 6C), and this site occurred at the region of increased Pol II density (Figure 6B).

We investigated whether reduced cleavage at the termination window could cause the global change in Pol II distribution by analyzing initiation and cleavage at the genes that displayed a prominent change in Pol II progression (Figure 6D). Paused Pol II at the promoter-proximal region has transcribed through fewer nucleotides (<60 nt) than the sequenced read length in our PRO-seq data (75 nt). Thus, the TSS-containing reads report both the initiating base (5′ end of the read) and the position of Pol II at the pause region (Rasmussen and Lis, 1993); Nechaev et al., 2010; Tome et al., 2018) and allows deducing whether Pol II resides at the pause or has proceeded into productive elongation (Figure 6D, upper left panel). The decrease in promoter-proximal Pol II in preconditioned cells composed of transcripts with the whole spectrum of PRO-seq read lengths (20–75 nt), which indicates less initiating, pausing, and early elongating Pol II complexes (Figure 6D). This reduction in all promoter-proximal Pol II states corroborates our analyses at individual genes where Pol II recruitment was found as the major rate-limiting step of decreased transcription in preconditioned cells (Figures 4B and 6C). At the region downstream of CPS, the read length provides a measure of transcript cleavage: reads shorter than the maximum sequenced read length contain transcripts that have been cleaved to release the pre-mRNA (Figure 6D, upper right panel). The genome-wide increase in Pol II density at the termination window (Figures 4C, 4D, and S15A–S15D) was composed almost exclusively of reads with the maximum read length (Figure 6D). This selective increase in transcription complexes with no signs of cleavage indicated that the accumulation of Pol II at CPS co-occurred with reduced pre-mRNA processing. Moreover, the reduction in a gene's initiation strongly correlated with the Pol II accumulation at its termination window ($p = -0.51$; $rho = -0.29$; Figure S15F), coupling the prolonged termination to the same gene's lower rate of initiation. Because transcript cleavage is required to release Pol II from the chromatin, a compromised recycling of Pol II from the end of the gene into a new initiation could account for the global change in transcription in preconditioned cells.

### Enhancers with reduced initiation connect to genes with increased Pol II density at the termination window

Enhancers recruit transcription factors and Pol II, and they are brought to physical proximity with the target genes via chromatin looping (reviewed in Field and Adelman, 2020). To analyze recycling of Pol II between genes and enhancers, we identified chromatin loops and measured Pol II density at the connected genes and enhancers. Enhancers that looped to genes with increased Pol II density at the termination window showed a significant reduction in heat-induced Pol II density after preconditioning (Figures 6E and S15G). In contrast, enhancers that looped to genes without a prominent change in the termination displayed similar Pol II densities in unconditioned and preconditioned cells (Figures 6E and S15G). Monitoring the progression of Pol II through the distinct rate-limiting steps of transcription allows us to propose a model (Figure 7) where reduced transcript cleavage at the termination window retains Pol II bound to chromatin and diminishes recycling of the transcription machinery. The limited availability of Pol II in preconditioned cells lowers initiation without the need to change the chromatin state or HSF1 binding. The lower initiation rate, in turn, reduces mRNA production in preconditioned cells. Our model also explains the lower enhancer transcription in preconditioned cells, identifying the affected enhancers to reside in chromatin loops with genes where Pol II accumulates at the termination window.

## DISCUSSION

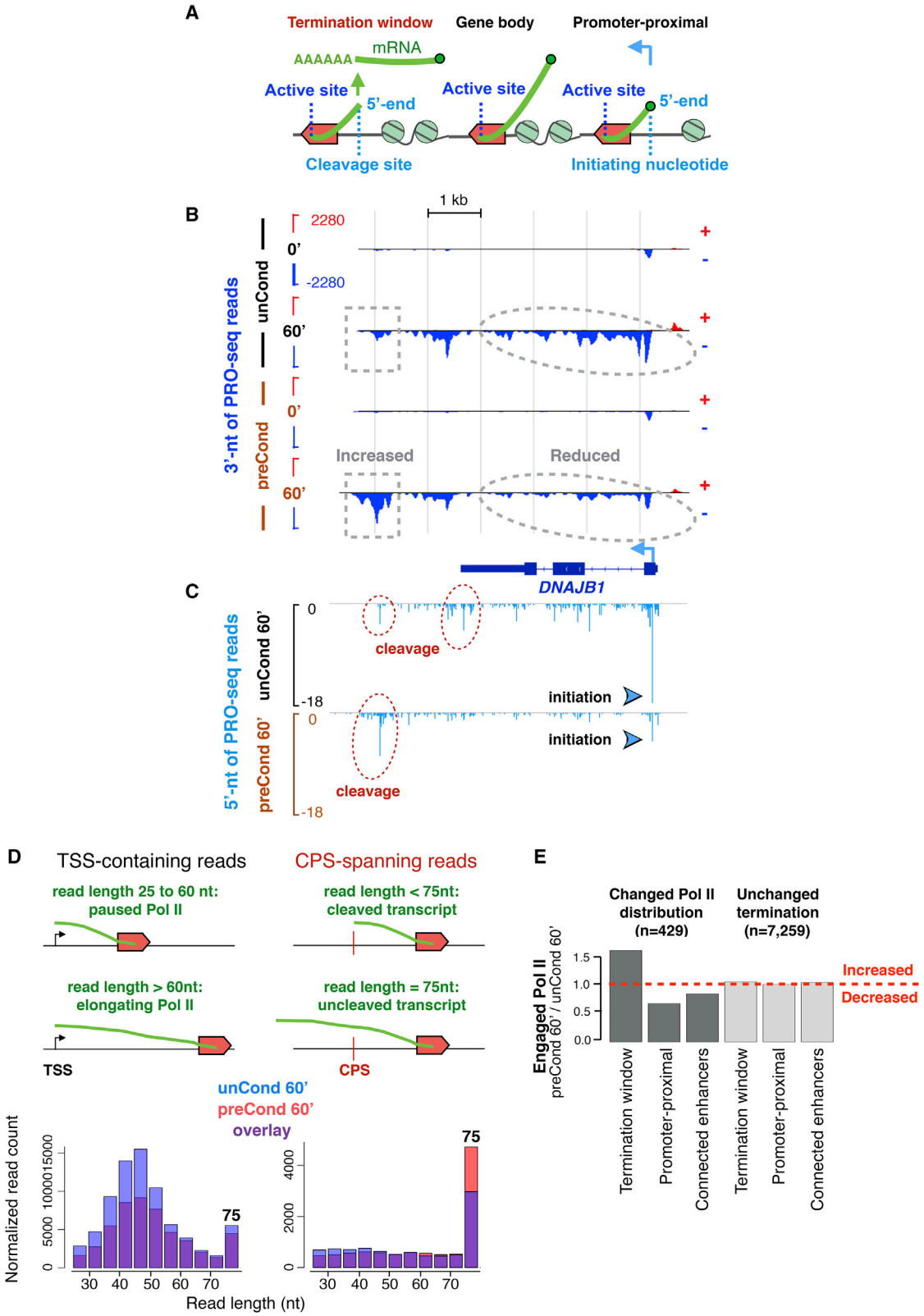
### Control of Pol II pause release enables rapid and reversible transcriptional reprogramming

The groundbreaking model by Conrad Waddington (1957) describes developing cells as marbles that roll down an energy landscape of hills and valleys. While rolling down, cells take different paths and commit to distinct cell types, remodeling their

(D) Inset of *TAX1BP1* enhancer (green bar) and TSS (purple arrow), showing heat-induced HSF1 binding (gray) to the enhancer and TBP binding to the promoter (purple).
(E) Inset showing enhancer transcription in the presence and absence of HSF1.
(F) Average Pol II density at HSF1-dependently heat-induced enhancers.
(G) Average ATAC-seq density at HSF1-dependent and highly transcribed enhancers.
In (F) and (G), the shaded area: 12.5%–87.5% confidence interval. See also Figures S3 and S10–S14.

Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007






*(legend on next page)*

chromatin environment and transcription program (reviewed in Takahashi and Yamanaka, 2015). Reversing from a differentiated to pluripotent cell, instead, requires specific transcription factors that push the cell up the energy landscape, which rarely occurs in nature (Gurdon et al., 1958; Takahashi and Yamanaka 2006). Here, we showed that, after genome-wide reprogramming of transcription by heat shock, cells return to their cell-type-specific transcription program within hours of recovery (Figure 1). In Waddington's landscape, the heat-induced reprogramming would be analogous with the cell transiently occupying a nearby valley but, during recovery, returning to its cell-type-specific basal transcription program. This rapid reprogramming and precise recovery highlight the plasticity of transcription program and imply that the transcriptional heat shock response is truly transient.

The rapid and reversible heat-induced reprogramming can be explained mechanistically by genome-wide control of promoter-proximal Pol II pause release. An important consequence of repressing thousands of genes by preventing the release of Pol II from their promoter-proximal regions is its rapid reversibility; a simple reactivation of the pause release can restore productive gene transcription throughout the genome without extensive chromatin remodeling. In this regard, Pol II pausing can be considered as a memory that marks active genes and maintains open and accessible promoters during their transient repression. Indeed, reprogramming of transcription during differentiation involves gene silencing and activation by remodeling the chromatin (reviewed in Perino and Veenstra, 2016; Gökbuget and Blelloch, 2019). The reported changes in the chromatin upon heat shock (Zobeck et al., 2010; Petesch and Lis, 2012, Niskanen et al., 2015, Mueller et al., 2017; Vihervaara et al., 2017) involve modifications that are likely to transiently compartmentalize distinct gene activities (reviewed in Vihervaara et al., 2018). Moreover, chromatin conformation remains stable upon heat shock (Ray et al., 2019), which implies that the rapid recovery from stress does not require rewiring of the chromatin connectivity. We conclude that, as chromatin architecture is primed for an instantaneous transcriptional response to heat shock (Vihervaara et al., 2017; Ray et al., 2019), the Pol II pausing at heat-repressed genes primes rapid and robust transcriptional recovery, providing a memory of the cell's transcription program.

## Stress-induced control of Pol II is carried over mitotic divisions

A single heat shock, which is unlikely to cause permanent or long-lasting damage to the cell, did not change the basal transcription but increased Pol II pausing (Figure 1). The pausing of Pol II, in turn, can function as a space holder for a rapid signal-responsive regulation. In accordance, the daughter cells of singly preconditioned MEFs were able to accelerate Pol II entry into productive elongation (Figure 2). The faster induction of the machinery that clears damaged proteins and organelles via lysosomal degradation (Figure 7) is likely to raise another instant cytoprotective arm next to the rapidly heat-induced chaperone expression.

Cancer cells are highly stress tolerant (Hanahan and Weinberg, 2011). Accordingly, human K562 erythroleukemia cells proliferated through multiple heat shocks and adapted nascent transcription program to support survival. Two mitotic divisions after nine heat exposures, transcription of certain pro-survival genes was elevated, expression of genes that maintain protein production was decreased (Figure 3), and processing of transcripts at the 3′ ends of active genes was decelerated (Figure 6). In cells with decreased protein synthesis, the decelerated transcription termination likely serves to reduce the mRNA load as fewer Pol II molecules become available for new rounds of heat-induced transcription (Figures 6 and 7). The increased association of uncleaved transcripts at the 3′ ends of genes could provide a reservoir of pre-mRNAs that are rapidly processed to mature mRNAs once the cell restores its protein synthesis. Our results demonstrate that priming a faster gene activation and refining transcription over mitotic divisions can occur via regulation of Pol II (Figures 2, 3, and 4), without involving major changes in chromatin accessibility or binding activity of HSF1 (Figures 4 and 5). Taken together, cells exposed to stress can establish a memory by regulating the key rate-limiting steps of transcription.

## Limitations of study

This study tracks the process of nascent transcription at genes and enhancers across the genome and identifies the rate-limiting steps involved in establishing a transcriptional memory of cellular stress. Nevertheless, the factors that execute the increased Pol II pausing and trigger a faster release of the paused Pol II in the daughters of stress-exposed cells remain to be identified. Likewise, the molecular machinery at the termination window that is involved in

**Figure 6. Prolonged termination co-occurs with decreased RNA cleavage and reduced initiation at the gene's promoter and connected enhancers**

(A) Schematic: 3′ nt of PRO-seq reads report the active sites of transcription, and 5′ nt provide a readout for initiation (promoter-proximal region) and transcript cleavage (termination window). Pol II is depicted as a red rocket going from right to left. Green sphere in the end of the RNA molecule indicates 5′ cap that protects the transcript from exonucleosomal degradation.

(B) Active sites of transcription at *DNAJB1* gene. The boxed areas compare heat-induced transcription in unconditioned and repeatedly preconditioned K562 cells.

(C) 5′ nt of PRO-seq reads along *DNAJB1*. Initiation is indicated with light blue arrowheads; transcript cleavage sites are denoted with dashed red circles.

(D) Upper left panel: schematic presentation of TSS-overlapping PRO-seq reads. Paused Pol II associates with 25- to 60-nt-long reads, although productively elongating Pol II has proceeded beyond the +60 nt from TSS. Lower left panel: lengths of TSS-overlapping PRO-seq reads at genes where Pol II progression changes due to preconditioning are shown (n = 429). Upper right panel: schematic representation of CPS-overlapping PRO-seq reads is shown. Reads shorter than the maximum read length (75 nt) report events of transcript cleavage. Lower right panel: lengths of CPS-spanning reads at genes with changed Pol II progression are shown (n = 429). Accumulated Pol II molecules in termination windows of repeatedly preconditioned cells associate with uncleaved transcripts.

(E) Fold change of engaged Pol II in preconditioned over unconditioned cells at termination windows, promoter-proximal regions, and connected enhancers. The red dashed line indicates fold change 1. Increased and decreased denote higher and lower Pol II density after preconditioning.

See also Figures S3, S12, and S15.

Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007



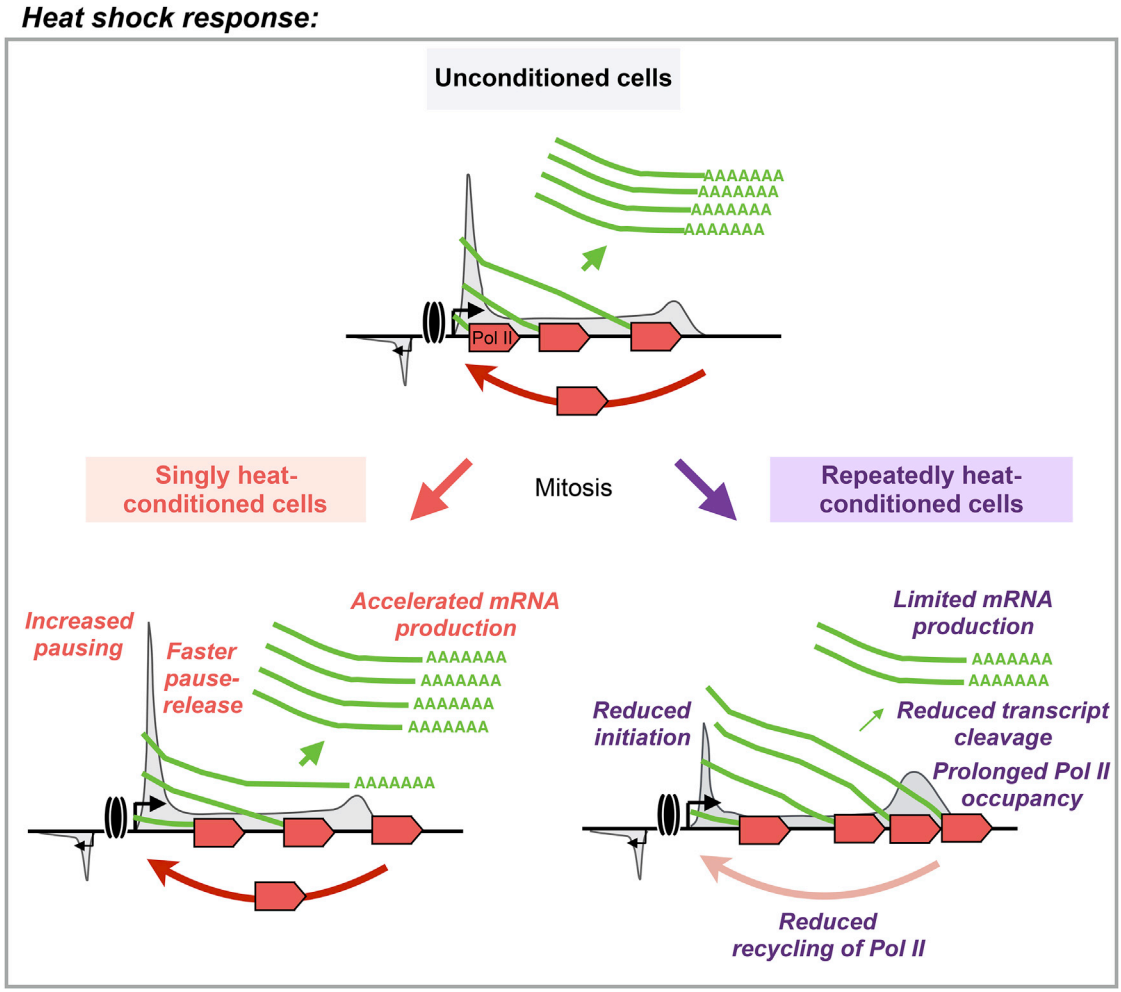



**Figure 7. Model for heat-induced transcriptional memory accelerating promoter-proximal pause release and decelerating termination over cell divisions**

In unconditioned cells (upper panel), paused Pol II is rapidly released from the promoters of heat-induced genes into elongation, and it efficiently proceeds through the gene. A single heat shock exposure (lower left panel) primes an additional set of genes for instant heat induction in the daughter cells by increasing Pol II pausing. Multiple heat shocks (lower right panel) cause reduced transcript cleavage at the 3′ end of active genes, which decelerates termination and decreases recycling of the transcription machinery to heat-activated genes and enhancers.

retaining Pol II associated with the nascent transcript are currently unknown. Increased residency of Pol II at the termination window correlated with reduced initiation at the gene's TSS and connected enhancers. The movements of Pol II between genes and enhancers remain to be shown. The cell models in this study are limited to untransformed MEFs and human K562 erythroleukemia cells. The memory-induced changes in Pol II regulation can occur without major changes in chromatin accessibility, but our results do not exclude the involvement of transcriptional regulators in priming a faster transcriptional response to stress or coordinating prolonged termination in the daughters of heat-shocked cells.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Cell Culture, Heat Treatments and Cell Cycle Profiling
  - Depletion of HSF1 with Short Hairpin RNA
  - MNase-coupled quantitative ChIP
  - Quantitative Reverse Transcription PCR
  - Western Blotting
  - PRO-seq
  - PRO-qPCR
  - Omni-ATAC-seq
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Computational Analyses of PRO-seq Data

- Normalization of PRO-seq Data
- Quantifying Gene Transcription
- Identification of Transcribed Enhancers
- Identifying Gene-Enhancer Loops
- Analyses of Differential Gene and Enhancer Transcription
- Analyses of HSF1-Dependent Transcription at Genes and Enhancers
- Visualizing Transcriptionally Engaged Pol II in Genome Browsers and as Composite Profiles
- Identification of Genes with Compromised Pol II Progression
- Quantifying Engaged Pol II Molecules in Distinct Genomic Regions
- Additional Datasets Used

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2021.03.007.

## ACKNOWLEDGMENTS

We thank the members of the Sistonen and the Lis laboratories for valuable advice during the manuscript preparation. This work was financially supported by the Sigrid Jusélius Foundation (A.V. and L.S.), Academy of Finland (A.V. and L.S.), SciLifeLab (A.V.), Svenska Tekniska Vetenskapsakademin i Finland (A.V.), South-West Finland's Cancer Foundation (A.V.), Borg Memory Foundation (A.V.), Åbo Akademi University (A.V. and L.S.), The Finnish Cultural Foundation (A.V.), Alfred Kordelin Foundation (S.V.H.), Cancer Foundation Finland (L.S.), Magnus Ehrnrooth Foundation (L.S.), and NIH grants RO1-GM25232 and HG009393 (J.T.L.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

## AUTHOR CONTRIBUTIONS

A.V., J.T.L., and L.S. conceived and designed the study. A.V., D.B.M., S.V.H., and M.A.H.B. conducted the laboratory work, and A.V. and D.B.M. performed the computational data analyses. All the authors interpreted the results. A.V., J.T.L., and L.S. wrote the manuscript with edits from D.B.M., S.V.H., and M.A.H.B.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 15, 2019
Revised: August 14, 2020
Accepted: March 4, 2021
Published: March 29, 2021

## REFERENCES

Ahn, S.G., Liu, P.C., Klyachko, K., Morimoto, R.I., and Thiele, D.J. (2001). The loop domain of heat shock transcription factor 1 dictates DNA-binding specificity and responses to heat stress. Genes Dev. *15*, 2134–2145.

Azofeifa, J.G., and Dowell, R.D. (2017). A generative model for the behavior of RNA polymerase. Bioinformatics *33*, 227–234.

Boettiger, A.N., and Levine, M. (2009). Synchronous and stochastic patterns of gene activation in the Drosophila embryo. Science *325*, 471–473.

Booth, G.T., Parua, P.K., Sansó, M., Fisher, R.P., and Lis, J.T. (2018). Cdk9 regulates a promoter-proximal checkpoint to modulate RNA polymerase II elongation rate in fission yeast. Nat. Commun. *9*, 543.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Cardiello, J.F., Sanchez, G.J., Allen, M.A., and Dowell, R.D. (2020). Lessons from eRNAs: understanding transcriptional regulation through the lens of nascent RNAs. Transcription *11*, 3–18.

Chatterjee, S., and Burns, T.F. (2017). Targeting heat shock proteins in cancer: a promising therapeutic approach. Int. J. Mol. Sci. *18*, E1978.

Chen, M., and Xie, S. (2018). Therapeutic targeting of cellular stress responses in cancer. Thorac. Cancer *9*, 1575–1582.

Chu, T., Rice, E.J., Booth, G.T., Salamanca, H.H., Wang, Z., Core, L.J., Longo, S.L., Corona, R.J., Chin, L.S., Lis, J.T., et al. (2018). Chromatin run-on and sequencing maps the transcriptional regulatory landscape of glioblastoma multiforme. Nat. Genet. *50*, 1553–1564.

Consortium, E.P.; ENCODE Project Consortium (2011). A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol. *9*, e1001046.

Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J., et al. (2016). Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat. Genet. *48*, 1193–1203.

Core, L.J., Martins, A.L., Danko, C.G., Waters, C.T., Siepel, A., and Lis, J.T. (2014). Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. Nat. Genet. *46*, 1311–1320.

D'Urso, A., and Brickner, J.H. (2017). Epigenetic transcriptional memory. Curr. Genet. *63*, 435–439.

D'Urso, A., Takahashi, Y.H., Xiong, B., Marone, J., Coukos, R., Randise-Hinchliff, C., Wang, J.P., Shilatifard, A., and Brickner, J.H. (2016). Set1/COMPASS and Mediator are repurposed to promote epigenetic transcriptional memory. eLife *5*, 5.

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. *4*, 3.

Duarte, F.M., Fuda, N.J., Mahat, D.B., Core, L.J., Guertin, M.J., and Lis, J.T. (2016). Transcription factors GAF and HSF act at distinct regulatory steps to modulate stress-induced gene activation. Genes Dev. *30*, 1731–1746.

Elsing, A.N., Aspelin, C., Björk, J.K., Bergman, H.A., Himanen, S.V., Kallio, M.J., Roos-Mattjus, P., and Sistonen, L. (2014). Expression of HSF2 decreases in mitosis to enable stress-inducible transcription and cell survival. J. Cell Biol. *206*, 735–749.

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nat. Protoc. *7*, 1728–1740.

Field, A., and Adelman, K. (2020). Evaluating enhancer function and transcription. Annu. Rev. Biochem. *89*, 213–234.

Fujiwara, T., O'Geen, H., Keles, S., Blahnik, K., Linnemann, A.K., Kang, Y.A., Choi, K., Farnham, P.J., and Bresnick, E.H. (2009). Discovering hematopoietic mechanisms through genome-wide analysis of GATA factor chromatin occupancy. Mol. Cell *36*, 667–681.

Fulco, C.P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S.R., Perez, E.M., Kane, M., Cleary, B., Lander, E.S., and Engreitz, J.M. (2016). Systematic mapping of functional enhancer-promoter connections with CRISPR interference. Science *354*, 769–773.

Gerner, E.W., and Schneider, M.J. (1975). Induced thermal resistance in HeLa cells. Nature *256*, 500–502.

Gökbuget, D., and Blelloch, R. (2019). Epigenetic control of transcriptional regulation in pluripotency and early differentiation. Development *146*, 164772.

Guan, Z., Giustetto, M., Lomvardas, S., Kim, J.H., Miniaci, M.C., Schwartz, J.H., Thanos, D., and Kandel, E.R. (2002). Integration of long-term-memory-related synaptic plasticity involves bidirectional regulation of gene expression and chromatin structure. Cell *111*, 483–493.

Guertin, M.J., and Lis, J.T. (2010). Chromatin landscape dictates HSF binding to target DNA elements. PLoS Genet. *6*, e1001114.

Gurdon, J.B., Elsdale, T.R., and Fischberg, M. (1958). Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. Nature *182*, 64–65.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Harding, H.P., Zhang, Y., Scheuner, D., Chen, J.J., Kaufman, R.J., and Ron, D. (2009). Ppp1r15 gene knockout reveals an essential role for translation initiation factor 2 alpha (eIF2alpha) dephosphorylation in mammalian development. Proc. Natl. Acad. Sci. USA *106*, 1832–1837.

Heard, E., and Martienssen, R.A. (2014). Transgenerational epigenetic inheritance: myths and mechanisms. Cell *157*, 95–109.

Henriques, T., Scruggs, B.S., Inouye, M.O., Muse, G.W., Williams, L.H., Burkholder, A.B., Lavender, C.A., Fargo, D.C., and Adelman, K. (2018). Widespread transcriptional pausing and elongation control at enhancers. Genes Dev. *32*, 26–41.

Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., Trompouki, E., Bowman, T.V., Zon, L.I., Yuan, G.C., et al. (2016). Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. Dev. Cell *36*, 9–23.

Ingolia, T.D., and Craig, E.A. (1982). Drosophila gene related to the major heat shock-induced gene is transcribed at normal temperatures and not induced by heat shock. Proc. Natl. Acad. Sci. USA *79*, 525–529.

Kaati, G., Bygren, L.O., and Edvinsson, S. (2002). Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. Eur. J. Hum. Genet. *10*, 682–688.

Kampinga, H.H., Hageman, J., Vos, M.J., Kubota, H., Tanguay, R.M., Bruford, E.A., Cheetham, M.E., Chen, B., and Hightower, L.E. (2009). Guidelines for the nomenclature of the human heat shock proteins. Cell Stress Chaperones *14*, 105–111.

Klimczak, M., Biecek, P., Zylicz, A., and Zylicz, M. (2019). Heat shock proteins create a signature to predict the clinical outcome in breast cancer. Sci. Rep. *9*, 7507.

Koeffler, H.P., and Golde, D.W. (1980). Human myeloid leukemia cell lines: a review. Blood *56*, 344–350.

Kroemer, G., Mariño, G., and Levine, B. (2010). Autophagy and the integrated stress response. Molecular Cell *40*, 280–293.

Kwak, H., Fuda, N.J., Core, L.J., and Lis, J.T. (2013). Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science *339*, 950–953.

Lämke, J., Brzezinka, K., and Bäurle, I. (2016). HSFA2 orchestrates transcriptional dynamics after heat stress in Arabidopsis thaliana. Transcription *7*, 111–114.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Leppä, S., Kajanne, R., Arminen, L., and Sistonen, L. (2001). Differential induction of Hsp70-encoding genes in human hematopoietic cells. J. Biol. Chem. *276*, 31713–31719.

Li, Q., Peterson, K.R., Fang, X., and Stamatoyannopoulos, G. (2002). Locus control regions. Blood *100*, 3077–3086.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lozzio, C.B., and Lozzio, B.B. (1975). Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. Blood *45*, 321–334.

Luft, J.C., Benjamin, I.J., Mestril, R., and Dix, D.J. (2001). Heat shock factor 1-mediated thermotolerance prevents cell death and results in G2/M cell cycle arrest. Cell Stress Chaperones *6*, 326–336.

Mahat, D.B., and Lis, J.T. (2017). Use of conditioned media is critical for studies of regulation in response to rapid heat shock. Cell Stress Chaperones *22*, 155–162.

Mahat, D.B., Salamanca, H.H., Duarte, F.M., Danko, C.G., and Lis, J.T. (2016). Mammalian heat shock response and mechanisms underlying its genome-wide transcriptional regulation. Mol. Cell *62*, 63–78.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. *17*, 10–12.

Maytin, E.V., Wimberly, J.M., and Anderson, R.R. (1990). Thermotolerance and the heat shock response in normal human keratinocytes in culture. J. Invest. Dermatol. *95*, 635–642.

McMillan, D.R., Xiao, X., Shao, L., Graves, K., and Benjamin, I.J. (1998). Targeted disruption of heat shock transcription factor 1 abolishes thermotolerance and protection against heat-inducible apoptosis. J. Biol. Chem. *273*, 7523–7528.

Melgar, M.F., Collins, F.S., and Sethupathy, P. (2011). Discovery of active enhancers through bidirectional expression of short transcripts. Genome Biol. *12*, R113.

Mikhaylichenko, O., Bondarenko, V., Harnett, D., Schor, I.E., Males, M., Viales, R.R., and Furlong, E.E.M. (2018). The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. Genes Dev. *32*, 42–57.

Mivechi, N.F. (1989). Heat sensitivity, thermotolerance, and profile of heat shock protein synthesis of human myelogenous leukemias. Cancer Res. *49*, 1954–1958.

Mueller, B., Mieczkowski, J., Kundu, S., Wang, P., Sadreyev, R., Tolstorukov, M.Y., and Kingston, R.E. (2017). Widespread changes in nucleosome accessibility without changes in nucleosome occupancy during a rapid transcriptional induction. Genes Dev. *31*, 451–462.

Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y., and Adelman, K. (2010). Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. Science *327*, 335–338.

Niskanen, E.A., Malinen, M., Sutinen, P., Toropainen, S., Paakinaho, V., Vihervaara, A., Joutsen, J., Kaikkonen, M.U., Sistonen, L., and Palvimo, J.J. (2015). Global SUMOylation on active chromatin is an acute heat stress response restricting transcription. Genome Biol. *16*, 153.

Östling, P., Björk, J.K., Roos-Mattjus, P., Mezger, V., and Sistonen, L. (2007). Heat shock factor 2 (HSF2) contributes to inducible expression of hsp genes through interplay with HSF1. J. Biol. Chem. *282*, 7077–7086.

Perez, M.F., and Lehner, B. (2019). Intergenerational and transgenerational epigenetic inheritance in animals. Nat. Cell Biol. *21*, 143–151.

Perino, M., and Veenstra, G.J. (2016). Chromatin control of developmental dynamics and plasticity. Dev. Cell *38*, 610–620.

Petesch, S.J., and Lis, J.T. (2008). Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. Cell *134*, 74–84.

Petesch, S.J., and Lis, J.T. (2012). Activator-induced spread of poly(ADP-ribose) polymerase promotes nucleosome loss at Hsp70. Mol. Cell *45*, 64–74.

Proudfoot, N.J. (2016). Transcriptional termination in mammals: stopping the RNA polymerase II juggernaut. Science *352*, aad9926.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rasmussen, E.B., and Lis, J.T. (1993). In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. Proc. Natl. Acad. Sci. USA *90*, 7923–7927.

Ray, J., Munn, P.R., Vihervaara, A., Lewis, J.J., Ozer, A., Danko, C.G., and Lis, J.T. (2019). Chromatin conformation remains stable upon extensive transcriptional changes driven by heat shock. Proc. Natl. Acad. Sci. USA *116*, 19431–19439.

Rougvie, A.E., and Lis, J.T. (1988). The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. Cell *54*, 795–804.

Sailaja, B.S., Cohen-Carmon, D., Zimmerman, G., Soreq, H., and Meshorer, E. (2012). Stress-induced epigenetic transcriptional memory of acetylcholinesterase by HDAC4. Proc. Natl. Acad. Sci. USA *109*, E3687–E3695.

Skene, P.J., and Henikoff, S. (2015). A simple method for generating high-resolution maps of genome-wide protein binding. eLife *4*, e09225.

Song, S.H., Hou, C., and Dean, A. (2007). A positive role for NLI/Ldb1 in long-range beta-globin locus control region function. Mol. Cell *28*, 810–822.

Spektor, R., Tippens, N.D., Mimoso, C.A., and Soloway, P.D. (2019). methyl-ATAC-seq measures DNA methylation at accessible chromatin. Genome Res. *29*, 969–977.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Takahashi, K., and Yamanaka, S. (2015). A developmental framework for induced pluripotency. Development *142*, 3274–3285.

Tetievsky, A., and Horowitz, M. (2010). Posttranslational modifications in histones underlie heat acclimation-mediated cytoprotective memory. J. Appl. Physiol. (1985) *109*, 1552–1561.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.

Tippens, N.D., Vihervaara, A., and Lis, J.T. (2018). Enhancer transcription: what, where, when, and why? Genes Dev. *32*, 1–3.

Tippens, N.D., Liang, J., Leung, A.K., Wierbowski, S.D., Ozer, A., Booth, J.G., Lis, J.T., and Yu, H. (2020). Transcription imparts architecture, function and logic to enhancer units. Nat. Genet. *52*, 1067–1075.

Tome, J.M., Tippens, N.D., and Lis, J.T. (2018). Single-molecule nascent RNA sequencing identifies regulatory domain architecture at promoters and enhancers. Nat. Genet. *50*, 1533–1541.

Vihervaara, A., and Sistonen, L. (2014). HSF1 at a glance. J. Cell Sci. *127*, 261–266.

Vihervaara, A., Sergelius, C., Vasara, J., Blom, M.A.H., Elsing, A.N., Roos-Mattjus, P., and Sistonen, L. (2013). Transcriptional response to stress in the dynamic chromatin environment of cycling and mitotic cells. Proc. Natl. Acad. Sci. USA *110*, E3388–E3397.

Vihervaara, A., Mahat, D.B., Guertin, M.J., Chu, T., Danko, C.G., Lis, J.T., and Sistonen, L. (2017). Transcriptional response to stress is pre-wired by promoter and enhancer architecture. Nat. Commun. *8*, 255.

Vihervaara, A., Duarte, F.M., and Lis, J.T. (2018). Molecular mechanisms driving transcriptional stress responses. Nat. Rev. Genet. *19*, 385–397.

Vilborg, A., Sabath, N., Wiesel, Y., Nathans, J., Levy-Adam, F., Yario, T.A., Steitz, J.A., and Shalgi, R. (2017). Comparative analysis reveals genomic features of stress-induced transcriptional readthrough. Proc. Natl. Acad. Sci. USA *114*, E8362–E8371.

Waddington, C.H. (1957). The Strategy of the Genes. A Discussion of Some Aspects of Theoretical Biology (George Allen & Unwin).

Walter, P., and Ron, D. (2011). The unfolded protein response: from stress pathway to homeostatic regulation. Science *334*, 1081–1086.

Wang, Z., Chu, T., Choate, L.A., and Danko, C.G. (2019). Identification of regulatory elements from nascent transcription using dREG. Genome Res. *29*, 293–303.

Wei, Y., Schatten, H., and Sun, Q.Y. (2015). Environmental epigenetic inheritance through gametes and implications for human reproduction. Hum. Reprod. Update *21*, 194–208.

Wissink, E.M., Vihervaara, A., Tippens, N.D., and Lis, J.T. (2019). Nascent RNA analyses: tracking transcription and its regulation. Nat. Rev. Genet. *20*, 705–723.

Wu, W., Morrissey, C.S., Keller, C.A., Mishra, T., Pimkin, M., Blobel, G.A., Weiss, M.J., and Hardison, R.C. (2014). Dynamic shifts in occupancy by TAL1 are guided by GATA factors and drive large-scale reprogramming of gene expression during hematopoiesis. Genome Res. *24*, 1945–1962.

Yost, H.J., and Lindquist, S. (1991). Heat shock proteins affect RNA processing during the heat shock response of Saccharomyces cerevisiae. Mol. Cell. Biol. *11*, 1062–1068.

Zhang, F., Kumano, M., Beraldi, E., Fazli, L., Du, C., Moore, S., Sorensen, P., Zoubeidi, A., and Gleave, M.E. (2014). Clusterin facilitates stress-induced lipidation of LC3 and autophagosome biogenesis to enhance cancer cell survival. Nat. Commun. *5*, 5775.

Zobeck, K.L., Buckley, M.S., Zipfel, W.R., and Lis, J.T. (2010). Recruitment timing and dynamics of transcription factors at the Hsp70 loci in living cells. Mol. Cell *40*, 965–975.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Antibodies** | | |
| HSF1 | Enzo Life Sciences | SPA-901; RRID: AB_2120247 |
| Pol II | Abcam | 8WG16; RRID: AB_2268549 |
| β-tubulin | Abcam | ab6046; RRID: AB_2210370 |
| **Deposited data** | | |
| PRO-seq, K562 | This study | GSE127844 |
| PRO-seq, K562 | This study | GSE154746 |
| PRO-seq, MEFs | This study | GSE128160 |
| ATAC-seq, K562 cells | This study | GSE154744 |
| PRO-seq, K562 | Vihervaara et al., 2017 | GSE89230 |
| PRO-seq, MEFs | Mahat et al., 2016 | GSE71708 |
| HSF1 ChIP-seq, K562 | Vihervaara et al., 2013 | GSE43579 |
| ChIA-PET, Pol II, K562 | ENCODE | EGSM970213 |
| TBP ChIP-seq, K562 | ENCODE | GSM935495 |
| GATA1 ChIP-seq, K562 | ENCODE | GSM935540 |
| GATA2 ChIP-seq, K562 | ENCODE | GSM935373 |
| H3K9me1 | ENCODE | GSM733777 |
| H3K27ac | ENCODE | GSM733656 |
| H3K4me1 | ENCODE | GSM733692 |
| H3K4me3 | ENCODE | GSM733680 |
| DNaseI hypersensitive sites, K562 | ENCODE | GSM736629 |
| MNase-seq, K562 | ENCODE | GSM920557 |
| Western Blot raw images | Mendeley | https://doi.org/10.17632/gycj6tnw6v.1 |
| **Experimental models: Cell lines** | | |
| Mouse embryonic fibroblasts | McMillan et al., 1998 | WT |
| Human K562 erythroleukemia cells | ATCC | CCL-243 |
| **Oligonucleotides** | | |
| Primer and probe sequences for MNase-ChIP-qPCR, mRNA-qPCR and run-on-qPCR are listed in Table S1 (Table S1). | | N/A |
| Plasmid: shScr (human): GCG CGC TTT GTA GGA TTC G in pSUPER vector | Östling et al., 2007 | shScr |
| Plasmid: shHSF1 (human): GCT CAT TCA GTT CCT GAT C in pSUPER vector | Östling et al., 2007 | shHSF1 |
| **Software and algorithms** | | |
| Fastx toolkit | Cold Spring Harbor Laboratories, Hannon lab | http://hannonlab.cshl.edu/fastx_toolkit/ |
| MACS2 | Feng et al., 2012 | https://github.com/macs3-project/MACS |
| IGV | Thorvaldsdóttir et al., 2013 | http://software.broadinstitute.org/software/igv/ |
| Bowtie2 | Langmead and Salzberg, 2012 | http://bowtie-bio.sourceforge.net/bowtie2/index.shtml |
| dREG | Wang et al., 2019 | https://dreg.dnasequence.org |
| Bedtools | Quinlan and Hall, 2010 | https://bedtools.readthedocs.io/en/latest/ |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| DAVID | Dennis et al., 2003 | https://david.ncifcrf.gov/ |
| DESeq2 | Love et al., 2014 | https://bioconductor.org/packages/release/bioc/html/DESeq2.html |
| bigWig package | Andre Martins | https://github.com/andrelmartins/bigWig/ |
| Cutadapt | Martin, 2011 | https://cutadapt.readthedocs.io/en/stable/ |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed and will be fulfilled by the Lead Contact, Prof. Lea Sistonen (lea.sistonen@abo.fi).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

The PRO-seq and ATAC-seq datasets generated in this study have been deposited to Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/), and are available as raw and processed files through accession numbers GSE127844, GSE154746, GSE128160 and GSE154744. Original figures for Western Blotting images presented in this paper are available in Mendeley (https://doi.org/10.17632/gycj6tnw6v.1). Computational analyses have been performed using Unix, R and Python languages. Custom made scripts can be made available upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

In this study, human K562 erythroleukemia cells and mouse embryonic fibroblasts (MEFs) were used. The K562 cell line originated from ATCC. The immortalized MEFs originate from wild-type mouse (McMillan et al., 1998), and were obtained from Ivor Benjamin laboratory.

## METHOD DETAILS

### Cell Culture, Heat Treatments and Cell Cycle Profiling

Cells were maintained at 37°C in a humidified 5% $CO_2$ atmosphere. MEFs (McMillan et al., 1998) were cultured in Dulbecco's modified medium (GIBCO), and K562 cells in RPMI (Sigma), supplemented with 10% FCS, 2 mM L-glutamate, and streptomycin/penicillin (Mahat and Lis, 2017; Vihervaara et al., 2017). The 30-min and 60-min heat shock treatments were conducted by submerging the cell culture into a 42°C water bath (Vihervaara et al., 2017). The 12.5-min, 25-min, and 40-min heat shock treatments were instantly provoked by replacing the 37°C media with pre-warmed pre-conditioned media (Mahat and Lis 2017). In adherent MEFs, inducing or terminating heat shock does not require pelleting the cells. In K562 suspension cells, the 37°C media was removed after centrifugation (1000 rpm, 4 min), and the heat shock initiated by re-suspending the cells in pre-warmed (42°C) pre-conditioned media. The heat shock in K562 cells was terminated by placing the 10 mL of heat shock cell suspension into 35 mL of ice-cold PBS, followed by centrifugation (1000 rpm, 4 min) at 4°C. The non-heat-shocked control cells were retained in similar confluence, and subjected to same treatments, but only exposed to media and conditions at 37°C. Recovery from the heat shock(s) was conducted by placing the cells at 37°C in a humidified 5% $CO_2$ atmosphere. DNA content of the cells was determined by propidium iodide (PI) staining (40 μg/ml; Sigma), and progression of the cell cycle monitored by fluorescence-mediated counting (FACSCalibur, BD Biosciences). The FACS histograms were generated using Cell Quest Pro-6.0 (BD Biosciences) and Flowing Software 2.5 (Turku Bioscience Centre). The error bars in statistical analyses indicate standard deviations.

### Depletion of HSF1 with Short Hairpin RNA

HSF1 was depleted from K562 cells as previously described (Östling et al., 2007; Vihervaara et al., 2013) using shRNA constructs ligated into pSUPER vectors (Oligoengine). The vector-encoded oligonucleotides recognized HSF1 mRNA (GCT CAT TCA GTT CCT GAT C), or contained a scrambled sequence (GCG CGC TTT GTA GGA TTC G) that is not predicted to bind any sequence encoded by the human genome. The shRNA constructs were transfected into cells by electroporation (970 μF, 220 mV) 24 h prior to the first heat treatment.

Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007



### MNase-coupled quantitative ChIP

ChIP was performed as previously described (Östling et al., 2007; Vihervaara et al., 2013), with the following modifications to digest the unshielded chromatin with endo- and exonuclease MNase (Skene and Henikoff, 2015). After cross-linking protein-DNA interactions with 1% formaldehyde for 5 min on ice, the samples were quenched with 0.125 M glycine and washed with PBS. The pellets were resuspended in TM2 buffer (10 mM Tris, pH 7.5, 2 mM $MgCl_2$, 1x proteinase inhibitors cocktail from Roche, 1 mM DTT), and supplemented with 1.5% NP-40 to permeabilize the cells. The chromatin was fragmented using 6.3 U/μl MNase (New England Biolabs, NEB) for 10 min at 37°C in MNase buffer (10 mM Tris, pH 7.5, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 1x proteinase inhibitors cocktail from Roche, 1 mM DTT). The reaction was terminated in final concentration of 1% SDS and 10 mM EGTA. The digested chromatin was diluted in ChIP buffer (150 mM NaCl, 20 mM Tris-HCl pH 8.0, 1% Triton-X, 1x protease inhibitors from Roche), and pre-cleared with uncoupled protein G coated Sepharose beads (GE Heathcare). Immunoprecipitation was carried over night at 4°C using the following ChIP-verified antibodies: HSF1 (Spa-901, Enzo), H2.AZ (Abcam, ab4174), H3 (Merck Millipore, 06-755), H4 (Merck Millipore, 05-858) and AcH4 (06-866, Upstate). Proteins were degraded with proteinase K (Thermo Fisher) and RNA with RNase A (Invitrogen), and the cross-links reversed at 65°C overnight. The DNA was purified with phenol:chloroform extraction and ethanol precipitation, and amplified with primers and probes designed to match the exact +1 nucleosomes and promoters of *HSPA1A* and *HSPH1*. The primers and probes (Table S1) were as follows. *HSPA1A* promoter: forward: CTGGCCTCTGATTGGTCCAA; reverse: CACGGAGACCCGCCTTTT; probe: 5′-FAM-CGGGAGGCGAAACCCCTGGAA-BHQ-3′. *HSPA1A* +1 nucleosome: forward: CGGAAGGACCGAGCTCTT; reverse: GGCTCCGCTCTGAGATTG; probe: #47 (universal probe library, Roche). *HSPH1* promoter: forward: GAGGCAGGTTTGAGCCAAT; reverse: CGAGCCTTCTGGAAAGATTC; probe: #44 (universal probe library, Roche). *HSPH1* +1 nucleosome: forward: GGAAAGTTCTGATCAGTGCGATA; reverse: TGAACTACCGACCCAAAAGG; probe #73 (universal probe library, Roche). The enriched chromatin was quantified using TaqMan chemistry (Applied Biosystems), and the signal intensity in each sample was normalized against the respective total MNase-digested DNA (input).

### Quantitative Reverse Transcription PCR

For analyzing polyadenylated mRNA, RNA over 200 nt was isolated using RNeasy kit (QIAGEN). Subsequently, 1 μg of RNA was treated with DNase I (Promega) and reverse transcribed with Moloney murine leukemia virus reverse transcriptase RNase H(–) (Promega) using oligoT primer. Quantitative PCR (qPCR) reactions were run using ABI Prism 7900 (Applied Biosystems) with HSPA1A, HSPH1 and GAPDH primers (Oligomer) and probes (Oligomer or Roche Applied Science) reported in Table S1, and in Vihervaara et al. (2013) and Elsing et al. (2014). The forward primer for HSF1 mRNA is CAAGCTGTGGACCCTCGT, the reverse TCGAACACGTGGAAGCTGT, and the probe #67 (universal probe library, Roche). HSP and HSF1 mRNA levels were normalized to mRNA of GAPDH, and fold inductions calculated against non-treated (unCond 0') cells. All reactions were made in triplicate for samples derived from at least three biological replicates. Standard deviations were calculated and are shown in the graphs.

### Western Blotting

Cells were lysed in buffer C (25% glycerol, 20 mM HEPES pH 7.4, 1.5 mM $MgCl_2$, 0.42 M NaCl, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT), and protein concentration in the soluble fraction was measured using Bradford analysis. 20 μg of total soluble protein was boiled in Laemmli sample buffer, subjected to SDS-PAGE and transferred to nitrocellulose membrane (Protran nitrocellulose; Schleicher & Schuell). Proteins were analyzed with primary antibodies against HSF1 (Spa-901, Enzo), Pol II (Abcam, 8WG16) and β-tubulin (Abcam, ab6046). The secondary antibodies were HRP conjugated (GE Healthcare), and the blots were developed using an enhanced chemiluminescence method (ECL kit; GE Healthcare).

### PRO-seq

PRO-seq was performed as previously described (Kwak et al., 2013; Mahat et al., 2016; Vihervaara et al., 2017) with minor modifications. Nuclei of K562 cells were isolated in buffer A (10 mM Tris-Cl pH 8.0, 300 mM sucrose, 3 mM $CaCl_2$, 2 mM $MgAc_2$, 0.1% Triton X-100, 0.5 mM DTT) using Wheaton homogenizer (#357546, loose pestle). MEFs were incubated in permeabilization buffer (10 mM Tris-Cl, pH 7.5, 10 mM KCl, 250 mM sucrose, 5 mM $MgCl_2$, 1 mM EGTA, 0.05% Tween-20, 0.5 mM DTT, 1x protease inhibitors from Roche, 0.4 u/μl RNase inhibitor Superase In, Thermo Fisher). The nuclei or permeabilized cells were flash-frozen and stored at −80°C (10 mM Tris-HCl pH 8.0, 25% glycerol, 5 mM $MgAc_2$, 0.1 mM EDTA, 5 mM DTT). Before run-on reaction, an equal amount of untreated *Drosophila* S2 cells was spiked into each sample, counted to account for 1% of the total DNA in each run-on reaction. The following run-on reaction was performed at 37°C for 3 min in the presence of biotinylated nucleotides (5 mM Tris-HCl pH 8.0, 150 mM KCl, 0.5% Sarkosyl, 2.5 mM $MgCl_2$, 0.5 mM DTT, 0.05 mM biotin-A/C/G/UTP from Perkin Elmer, 0.4 u/μl RNase inhibitor). The total RNA was isolated with Trizol LS (Invitrogen). After EtOH-precipitation, the RNA was air-dried, base hydrolyzed with 0.1 N NaOH for 5 min on ice, and the NaOH was neutralized with Tris-HCl (pH 6.8). Unincorporated nucleotides were removed using P-30 columns (Bio-Rad), and the biotinylated nascent transcripts were isolated in a total of three rounds of streptavidin-coated magnetic bead (M-280, Invitrogen) purifications. Each bead binding was followed by Trizol extraction and EtOH-precipitation of the transcripts. The 5′-cap was removed with RNA 5′ Pyrophosphohydrolase (Rpph, NEB), and the 5′-hydroxyl group was repaired with T4 polynucleotide kinase (NEB). The libraries were generated using TruSeq small-RNA adaptors and sequenced using NextSeq500 (Illumina).

### PRO-qPCR

To quantify nascent RNA synthesis from selected heat-responsive genes, we modified PRO-seq to perform qPCR after the 3′-adaptor ligation. In brief, run-on reactions were conducted in the presence of both unlabeled (200 μM A/C/G/UTP) and biotinylated (50 μM biotin-A/C/G/UTP) nucleotides during a 5-minute run-on reaction at 37°C. Total RNA isolation, base hydrolysis, and 3′ adaptor ligation were conducted as described for PRO-seq. After the second bead binding, reverse transcription was performed using a primer against the 3′ adaptor, and qPCR reactions run with ABI Prism 7900 (Applied Biosystems). Primers (Oligomer) and probes (Oligomer and Roche Applied Sciences) were (Table S1): *HSPH1* forward: AGCAGGCGGATTGTTGTTAG; *HSPH1* reverse: AAAGAGGTGGGCTAATCTTTCA; *HSPH1* probe: #38 (universal probe library, Roche); *HSPA1A* forward: GCCGAGAAGGACGAGTTTGA; *HSPA1A* reverse: CCTGGTACAGTCCGCTGATGA; *HSPA1A* probe: FAM- TTACACACCTGCTCCAGCTCCTTCCTCTT-BHQ1; *MED26* forward: ATTCCAGATGACCCGCTAAG; MED26 reverse: CGGATCACTACCACACCAGA; *MED26* probe: #21 (universal probe library, Roche). The nascent transcription of *HSPA1A* and *HSPH1* was compared against nascent transcription of *Mediator subunit 26* (*MED26*), a gene and a region in the gene that was actively transcribed and unchanged upon heat shock (Vihervaara et al., 2017).

### Omni-ATAC-seq

ATAC-seq was performed as previously described (Corces et al., 2016; Spektor et al., 2019) using 100,000 human K562 cells as starting material. Instantly after the treatments, the cells were washed with ice-cold PBS, and incubated 3 min in 100 μL ice-cold lysis buffer [10 mM Tris-Cl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% (vol/vol) NP-40, 1x protease inhibitor cocktail, Roche]. After centrifugation (600 g, 10 min, 4°C), the samples were re-suspended in 50 μL tagmentation buffer [10 mM Tris-Cl, pH 7.4, 10% (vol/vol) dimethyl formamide, 5 mM $MgCl_2$], and tagmentation performed with 1 μL Tn5 transposase (described in Spektor et al., 2019) for 30 min at 37°C. DNA was isolated with phonol:chloroform extraction and ethanol precipitation using GlycoBlue (Invitrogen #AM9516) as a carrier. The correct size distribution in each library was verified by test amplifying 1/10 of the material in a dilution series, followed by visualization of the DNA in a 5% polyacrylamide gel using SYBR Gold (ThermoFisher). Half of each library was amplified 12 cycles with barcoded Nextera primers (Illumina) and Q5 DNA polymerase (NEB). After amplification, the DNA fragments were size selected with Ampure XP beads (Beckman Coulter #A63880), incubating the samples first in 0.5X beads, and subsequently, in 1.8X beads. The barcoded samples were pooled, verified with Bioanalyzer, and sequenced using Illumina NexSeq500. The sequenced reads were trimmed with fastx toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). To match the genomic coordinates of datasets generated in this study to the genomic coordinates of previously aligned datasets (Consortium and ENCODE Project Consortium, 2011; Vihervaara et al., 2013, 2017; Fulco et al., 2016; Mahat et al., 2016; Ray et al., 2019), we aligned the reads to the human genome version GRCh37/hg19. Each dataset was density-normalized (fragments per million mapped fragments, FPM). Correlation of replicate pairs was assessed by first calling ATAC-seq peaks (enriched loci of chromatin accessibility) with MACS2 (Feng et al., 2012) using a combined bam file from all the samples in this study. Next, the count of fragments at every MACS2-called peak was measured in each sample, and replicate correlation analyzed with Spearman's rank correlation. After ensuring accurate correlation, the replicates were combined, and three types of FPM-normalized bigwig files generated, reporting the whole released fragment, the middle 20 nt of each fragment, and 10 nt at both ends of each fragment, respectively (Figure S11A). The complete raw ATAC-seq datasets (GSE154744) are available through Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/).

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Computational Analyses of PRO-seq Data

The PRO-seq reads were adaptor-clipped using cutadapt (Martin, 2011) and trimmed and filtered with fastx toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Due to usage of external spike-in material from *Drosophila* S2 cells, we combined the human (GRCh37/hg19) and *Drosophila* (dm3) genomes into a single genome file (hg19-dm3). Likewise, the mouse genome (mm10) was combined with the *Drosophila* genome (dm3) into a distinct genome file (mm10-dm3). In both cases, chromosomes of the dm3 were renamed. Reads from K562 cells were aligned to the hg19-dm3 genome and reads from MEFs to the mm10-dm3 genome, using Bowtie 2 (Langmead and Salzberg, 2012). Reads that uniquely mapped to the chromosomes of the human (hg19) or, respectively, the mouse (mm10) genome, were retained. The reads that uniquely mapped to the dm3 chromosomes provided a count of reads for spike-in derived normalization factors. The complete raw PRO-seq datasets in K562 cells (GSE127844 and GSE154746), and MEFs (GSE128160) are available through Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/).

### Normalization of PRO-seq Data

Mapped reads were processed from bed files to coverage files, retaining only the 3′ end nucleotide (active sites of transcription), or the 5′ end nucleotide (for analyses of initiation and cleavage), of each read. Density normalized bedgraph files were adjusted by sample-specific normalization factors that were derived either from the spike-in read count (Booth et al., 2018) or the count of reads at the ends (+120,000 nt from TSS to −500 nt from CPS) of long (> 150 kb) genes (Mahat et al., 2016; Vihervaara et al., 2017). For samples measuring transcription during recovery from the heat shock, or the effect of multiple heat shocks, only spike-in control was utilized for normalization. When comparing transcription upon short heat shocks (12.5 min, 25 min, and 40 min), we first ensured correct normalization between the unconditioned 0 min and preconditioned 0 min time points with the spike-in-derived normalization factors.

Then, the 3′ ends of over 150 kb long genes were utilized to normalize samples of the rapid heat shock kinetics (unconditioned 12.5 min, 25 min and 40 min normalized against the unconditioned 0 min; preconditioned 12.5 min, 25 min and 40 min against the preconditioned 0 min). This strategy allows for highly sensitive sample normalization between short heat shock time points, and usage of an extrinsic control when normalization regions within samples are not available.

### Quantifying Gene Transcription

Actively transcribed genes and their primarily used isoforms were identified by mapping transcription initiation sites genome-wide using discriminative regulatory elements identification from global run-on data (dREG; https://dreg.dnasequence.org). The most updated version of dREG (Wang et al., 2019) is trained to call transcription initiation sites of genes and enhancers with high sensitivity using their characteristic pattern of divergent transcription (Core et al., 2014; Tome et al., 2018). To identify gene isoforms with active transcription initiation, TSSs of RefSeq-annotated transcripts were intersected (bedtools, Quinlan and Hall, 2010) with dREG-called active regulatory elements. Subsequently, transcripts that harbored dREG-called initiation at the TSS were retained. The level of transcription *per* each annotated transcript was measured from the gene body (+500 nt from TSS to −500 nt from CPS), as described previously (Mahat et al., 2016; Vihervaara et al., 2017). In the downstream analyses, we retained a single transcript *per* gene by selecting the isoform that showed the largest fold change to heat shock, or if called unresponsive to heat stress, had the highest level of transcription in non-stress condition. The analyses of enriched gene annotation categories were performed with Database for Annotation, Visualization and Integrated Discovery (DAVID; Dennis et al., 2003).

### Identification of Transcribed Enhancers

Transcribed enhancers were identified across the genome *de novo* using dREG (Wang et al., 2019; https://dreg.dnasequence.org) that recognizes patterns of transcription at genes and enhancers. Since heat shock changes Pol II progression at regulatory elements (Vihervaara et al., 2017), we identified transcribed regulatory elements individually in each sample, and then unified the coordinates obtained from all samples using bedtools merge with d −100 (Quinlan and Hall, 2010). Subsequently, the dREG-called regulatory elements were intersected with RefSeq-annotated TSSs of genes, and only elements that did not occur at any gene promoter were retained for enhancer analyses. We confirmed that this class of distal regulatory elements robustly captured functionally verified enhancers of MYC (described by Fulco et al., 2016), and of LCR at the beta-globin locus (Li et al., 2002; Song et al., 2007). The occurrence of putative enhancers at sites of physical chromatin connections was investigated from existing Pol II ChIA-PET data (EGSM970213). First, the ChIA-PET-enriched sites of chromatin connections (blocks) were intersected (bedtools, Quinlan and Hall, 2010) with our putative enhancer calls, as well as with annotated TSSs of genes. Subsequently, the chromatin connections from an enhancer to an enhancer, from an enhancer to a promoter, or from an enhancer to any Pol II ChIA-PET enriched region were identified. The percentage of putative enhancers in each of these chromatin connection classes is indicated.

### Identifying Gene-Enhancer Loops

To annotate enhancers to their target genes, we first utilized Pol II ChIA-PET data (EGSM970213) as indicated above, identifying the set of enhancers that connected to each gene's TSS. Since chromatin capture techniques negatively select for short-range interactions, we additionally annotated enhancers within 25 kb from the gene's TSS. Pol II densities were measured at a 1,000 nt span from the dREG-called enhancer midpoint, and the average Pol II densities at connected enhancers are shown for each indicated gene group.

### Analyses of Differential Gene and Enhancer Transcription

To call significant changes in gene and enhancer transcription, we utilized DESeq2 (Love et al., 2014), which uses the variance in biological replicates to assess significant changes between conditions. Differential gene expression was quantified from gene body transcription (+500 nt from TSS to −500 nt from CPS) of each gene. In this gene body window, Pol II has passed the initiation and pause regions and is undergoing productive elongation. Enhancer transcription was quantified along the whole enhancer length, individually for minus and plus strands (Vihervaara et al., 2017). For significantly changed transcription, we required p value < 0.05 (K562) or < 0.001 (MEFs), and fold enrichment > 1.25. The less stringent criterion for K562 cells used in this study, as compared to MEFs and our earlier data on K562 cells (Vihervaara et al., 2017), is due to lower sequencing depth. The heat-induced changes in transcription, as well as the sets of differentially transcribed genes and enhancers are highly similar in our distinct studies of the same cell type. The identification of genes with faster heat-induction or slower heat-repression in preconditioned MEFs is depicted in Figure S16. Highly heat-induced genes displayed FC > 2 in gene body transcription (heat shock / non-heat shock) and dRPK > 200 (heat shock –non heat shock) at least in one of the heat shock time points as compared to non-heat shock condition.

### Analyses of HSF1-Dependent Transcription at Genes and Enhancers

Nascent transcription upon HSF1-knockdown was inferred from a single replicate, chosen by the most prominent downregulation of HSF1 throughout the length of the experiment (Figure 5B). To identify HSF1-dependent genes, we used two approaches. First, we measured the heat-induced gene body transcription for each gene in the presence and absence of HSF1. This comparison of transcription level identified 186 genes whose heat induction in K562 cells depleted of HSF1 remained under 50% of the respective induction in cells expressing intact levels of HSF1. Second, we used the fact that unconditioned and preconditioned cells correlated to

Please cite this article in press as: Vihervaara et al., Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.03.007



the same extent as biological replicates (rho = 0.98) and contained similar levels of gene body transcription (Figure S15B), conducting DESeq2 using the same time point from unconditioned and preconditioned cells as a replicate pair. These analyses showed 227 genes and 496 enhancers to be HSF1 dependent in both unconditioned and preconditioned K562 cells (Figure S13F). To account for a subset of genes changing basal or heat-inducible transcription due to preconditioning, we complemented the DESeq2-analysis to also find genes that showed HSF1-dependency only in unconditioned or preconditioned cells, or that were called insignificant due to changes in basal transcription. Since HSF1-dependency of the 227 DESeq2-called genes ranged from 64.1% to 99.9% (Figure S13F), we queried genes that in either unconditioned or preconditioned cells were HSF1-dependent at least to 64.1%, gained at least two-fold heat-induction, and had a minimum gene body transcription of 50 RPK in any condition. This analysis identified 18 additional HSF1-dependent genes, including *PPP1R15A* that had lost heat-inducibility, and *HSPA8* that had gained higher basal transcription, upon preconditioning. All of the 18 genes were individually verified to be HSF1-dependent by browsing.

### Visualizing Transcriptionally Engaged Pol II in Genome Browsers and as Composite Profiles

Pol II densities as bigWig and bedgraph files were visualized with Integrative Genomics Viewer (IGV; Thorvaldsdóttir et al., 2013) and an in-house browser (Hojoong Kwak, Cornell University, Ithaca, NY, USA). The scale of y axis is equal and linear for tracks across different conditions for an indicated genomic region. To generate composite profiles, the read counts in defined genomic regions were obtained, and composite profiles generated using bigWig package (https://github.com/andrelmartins/bigWig/). The average intensities in composite profiles were queried in 20-nt, 10-nt or 1-nt bins. The shaded areas display 12.5%–87.5% fractions of the data in each queried window. To generate an average profile of gene bodies with different lengths, 1/500 of the gene body length was set to the bin size, after filtering out short genes where the bin would have been less than 1 nt.

### Identification of Genes with Compromised Pol II Progression

To identify genes with decreased Pol II density in 5′- and increased density in 3′-region, genes were first divided into three distinct regions: 1) 5′-coding region comprising 1000 nt downstream of the mid coordinate between Pol II pause sites of divergent transcription, 2) gene body, measured from +1000 nt from the mid of the pause sites to −1000 nt from the CPS, and 3) downstream, +100 nt to +6000 nt, of the CPS. PRO-seq reads in each region were measured, after which the read count in the preconditioned 60-min heat shock sample was deduced from the respective read count in the unconditioned 60-min heat shock sample. Since gene body transcription varies from gene to gene, we compared the change in Pol II progression within each gene. To identify genes with reduced 5′- and increased 3′- Pol II density, we required the reduction at 5′-coding region to be three times larger than the absolute change in the gene body read count. Simultaneously, the increase in read counts downstream of the CPS was required to be three times higher than the absolute change in the gene body read count.

### Quantifying Engaged Pol II Molecules in Distinct Genomic Regions

The mapped reads were sorted to distinct genomic regions by intersecting the 3′-coordinate of the read with the genomic coordinates described in Figure S15A. To avoid double mapping, gene body reads that overlapped with enhancers or pause regions were omitted. Subsequently, the number of reads in a given region was counted as fraction of total uniquely mapping reads in the PRO-seq data.

### Additional Datasets Used

Besides the PRO-seq (GSE127844, GSE128160 and GSE154746) and ATAC-seq (GSE154744) datasets generated in this study, the following datasets have been utilized: HSF1-binding sites in non-stressed and 30-min heat-shocked K562 cells (GSE43579; Vihervaara et al., 2013), binding sites of TBP (GSM935495), GATA1 (GSM935540) and GATA2 (GSM935373) in non-stressed K562 cells (Consortium and ENCODE Project Consortium, 2011); DNase I hypersensitive (GSM736629), MNase resistant (GSM920557), as well as H3K9me1 (GSM733777), H3K27ac (GSM733656), H3K4me1 (GSM733692) and H3K4me3 (GSM733680) enriched loci in non-stressed K562 cells (Consortium and ENCODE Project Consortium, 2011); Pol II ChIA-PET in non-stressed K562 cells (GSM970213); PRO-seq data in non-stressed and 30-min heat-shocked K562 cells for verification purposes (GSE89230; Vihervaara et al., 2017); PRO-seq data in non-stressed and 12.5-min heat-shocked MEFs (GSE71708; Mahat et al., 2016).