Isolation of the protein and RNA content of active sites of transcription from mammalian cells

Svitlana Melnik[1,5,6], Maïwen Caudron-Herger[2,5], Lilija Brant[3], Ian M Carr[4], Karsten Rippe[2], Peter R Cook[1] & Argyris Papantonis[3]

[1]Sir William Dunn School of Pathology, University of Oxford, Oxford, UK. [2]German Cancer Research Center (DKFZ) and BioQuant, Heidelberg, Germany. [3]Center for Molecular Medicine, University of Cologne, Cologne, Germany. [4]Institute of Biomedical and Clinical Sciences, University of Leeds, Leeds, UK. [5]These authors contributed equally to this work. [6]Present address: DNA Vectors Research, DKFZ, Heidelberg, Germany. Correspondence should be addressed to A.P. (argyris.papantonis@uni-koeln.de).

Published online 25 February 2016; doi:10.1038/nprot.2016.032

Mammalian cell nuclei contain three RNA polymerases (RNAP I, RNAP II and RNAP III), which transcribe different gene subsets, and whose active forms are contained in supramolecular complexes known as 'transcription factories.' These complexes are difficult to isolate because they are embedded in the 3D structure of the nucleus. Factories exchange components with the soluble nucleoplasmic pool over time as gene expression programs change during development or disease. Analysis of their content can provide information on the nascent transcriptome and its regulators. Here we describe a protocol for the isolation of large factory fragments under isotonic salt concentrations in <72 h. It relies on DNase I–mediated detachment of chromatin from the nuclear substructure of freshly isolated, unfixed cells, followed by caspase treatment to release multi-megadalton factory complexes. These complexes retain transcriptional activity, and isolation of their contents is compatible with downstream analyses by mass spectrometry (MS) or RNA-sequencing (RNA-seq) to catalog the proteins and RNA associated with sites of active transcription.

INTRODUCTION

Eukaryotic cell nuclei contain three RNAPs (RNAP I, RNAP II and RNAP III)[1,2], which are part of larger complexes; for example, RNAP II associates with complexes carrying enzymatic activities involved in initiation, capping, splicing and polyadenylation[3,4]. These complexes may be organized into even larger structures that harbor most of the machinery required for transcript production known as transcription factories[5,6]. Factories are typically defined as sites containing at least two different active transcription units (thus distinguishing them from sites in which two polymerizing complexes are active on one template), and they harbor more than 95% of total transcriptional activity in a mammalian cell[5]. The biochemical isolation of such factories would enable researchers to determine the protein content of the active sites of transcription and to identify the majority of nascent transcripts being produced at any given moment. However, their isolation has proved to be challenging because of their large mass and tight association with the nuclear substructure. Furthermore, the failure to isolate transcription factories had raised concerns as to whether factories actually exist[7]. A few years ago, we addressed these issues by introducing an approach for purifying large fragments of factories with an apparent mass larger than 8 MDa (the size of the largest protein marker commercially available), and we went on to analyze the proteomes of different complexes containing RNAP I, II or III (ref. 8). More recently, we adapted this method to catalog the RNA content of transcription factories, which revealed a strong enrichment of nascent transcripts[9]. This protocol is based on these previous publications, and it presents a unified workflow for the isolation of transcription factories and subsequent purification of their protein and/or RNA components (**Fig. 1**).

Development and overview of the protocol and its applications

There exists a long-standing research interest in complexes involved in the regulation of gene expression, and a growing volume of literature points to transcription factories as important contributors in this regulation (reviewed by Papantonis and Cook[5]). Still, until recently, neither the protein composition of factories nor their primary products—the nascent, factory-associated transcripts—could be directly interrogated.

Two factors make the purification of mammalian polymerases engaged on endogenous templates difficult. First, active enzymes represent only a quarter of the total number of all polymerases; the remaining proteins form a rapidly diffusing soluble pool that tends to aggregate in nonisotonic buffers[8,10,11]. Therefore, we use isotonic conditions to isolate the active enzymes, while removing the inactive protein fraction that remains soluble under these conditions. Second, engaged polymerases plus their templates and transcripts are housed in factories that are bound to the underlying nuclear scaffold[8,10–12]. As caspases (a family of cysteine proteases) deconstruct nuclei during apoptosis, we reasoned that they might be used to also release factories.

In brief, the present approach relies on the isolation of nuclei from living (unfixed) cells in close-to-physiological salt concentrations. Subsequently, DNase I is used to detach the majority of nuclear chromatin, and factories are disengaged from the substructure using a mixture of group III caspases in a 'native lysis buffer' (NLB). After pelleting nuclear debris, large fragments of the factories (>8 MDa) are found in the supernatant. Next, 2D electrophoretic analysis of proteins followed by MS can be applied to catalog the content of factories in a given cell type[8]. In this case, the complexes that contain the different RNAPs are separated, and their relative positions may be traced by staining with SYBR Green to reveal the boundaries of regions rich in RNA or—if nascent RNA has been ^{32}P-labeled (**Box 1**)—by autoradiography. Alternatively, following standard TRIzol-based purification of RNA, high-throughput sequencing can be applied to comprehensively catalog the transcripts associated with factories[9] (**Fig. 1**).

Overall, the current protocol is straightforward to implement and requires limited hands-on time, and its implementation

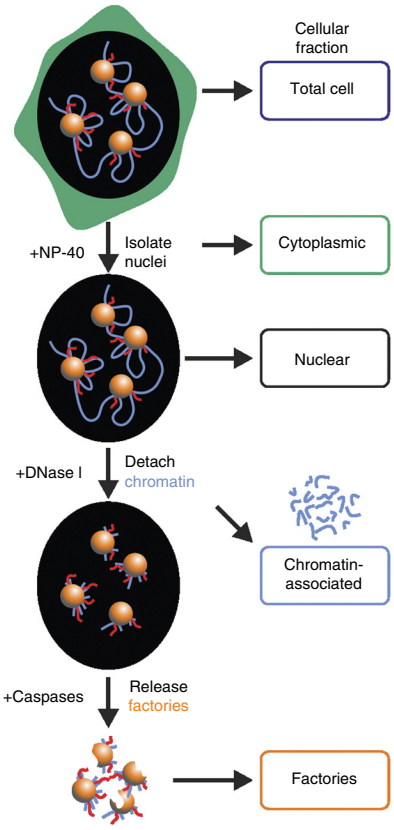

Figure 1 | Overview of the factory-isolation strategy. Live (not cross-linked) cells are collected in PB, and then nuclei are isolated after treating cells with NP-40. Next, whole nuclei (black oval) are treated with DNase I in a step in which the majority of chromatin (blue lines) is detached and can be separated from the rest of the nucleus by centrifugation. Finally, transcription factories (orange spheres), as well as any attached DNA (blue) and RNA (red), are released into the supernatant by digestion with a mixture of active caspases. As a result, different subcellular fractions (boxed areas) may be isolated along the course of this procedure.

offers reproducible results. The whole process takes <72 h (not including the time needed for MS or next-generation sequencing (NGS)), and it is particularly suited to the study of cellular responses to extracellular cues. For example, one could monitor the changes in the protein and RNA content of the active transcription compartment when transcription is reprogrammed by different cytokines.

Limitations of the protocol and comparison with other approaches

Although the content of mature and nascent proteins in large subcellular complexes such as nucleoli has been analyzed[13,14], factory-associated proteins have been difficult to interrogate. Our approach facilitates proteomic analysis of isolated factories (**Fig. 1**), but it still requires large numbers of cells (on the order of $\sim 5 \times 10^7$), especially if the complexes housing each of the RNAPs (I, II or III) need to be separated electrophoretically[8]. In addition, treatment with caspases 6, 8, 9 and 10 inevitably hinders the detection of those proteins that are most targeted by the specific proteases used (but all core subunits of the three RNAPs lack sites recognized by these caspases, with the exception of RPB9).

Cataloging nascent transcriptomes is challenging (in contrast to the now-routine mapping of mRNAs) owing to their short half-lives and low abundance, and because of the variety of processing pathways. A number of methods exist for isolating and characterizing newly synthesized RNA—for example, global run-on sequencing (GRO-seq)[15], mammalian native elongating transcript sequencing (NET-seq)[16,17], chromatin RNA-seq[18], poly(A)-depleted RNA-seq[19], nascent-seq[20] or the metabolic tagging of newly synthesized RNA using 4-thiouridine[21,22].

Box 1 | Labeling nascent RNA by run-on ● TIMING 2 h

Every time a new cell type is used for isolating factories, we recommend performing a run-on experiment to check the extent to which transcriptional activity is retained in PB. The instructions detailed below also enable researchers to track ^{32}P-labeled nascent RNA during the course of the procedure.

Procedure

1. Wash the cells twice in PB, making sure that cells stay in PB, on ice, for at least 5 min per wash.
2. Pellet the cells by centrifuging at 400*g* for 5 min at 4 °C.
3. Resuspend the cells in PB and permeabilize them for 5 min on ice by adding saponin to the cell suspension to a final concentration of 170 µg/ml.

▲ **CRITICAL STEP** The saponin concentration required for permeabilization needs to be established during preliminary experiments for each cell type and batch of saponin. For this purpose, incubate the cells in increasing concentrations of saponin (50–300 µg/ml) for 5 min on ice, and define the optimal concentration as that which gives ~98% permeabilized cells without extensive cell lysis (assessed using trypan blue), by observing an aliquot of the preparation under a light microscope.

4. Wash permeabilized cells four times in PB (as in steps 1 and 2, making sure that cells stay in PB, on ice, for at least 5 min per wash), and exploit the pauses during the washes to prepare the run-on mix, which, as mentioned in Reagent Setup, must be prepared freshly before use.
5. Prewarm permeabilized cells in PB at 33 °C for 3–5 min.
6. Add the run-on mix to the prewarmed cells, so that it just covers cells (e.g., 2.5 ml for a 15-cm culture dish, or 3–5 ml for a cell pellet). Incubate at 33 °C for 10 min.
7. Wash the cells twice in PB as in steps 1 and 2 to remove the mix. These cells will now carry labeled RNA.
8. Test the labeled samples for ^{32}P incorporation by scintillation counting; these samples can now be incorporated into the main PROCEDURE at Step 3.

? TROUBLESHOOTING

TABLE 1 | Comparison of factory RNA-seq to other methods that capture nascent RNA.

RNA-seq method	Applied in	Advantages	Limitations
Metabolic RNA 4sU-tagging[21,22]	Human and mouse	Straightforward 4sU incorporation; robust intronic RNA coverage	4sU pulse labeling might perturb transcript biogenesis and distribution; difficult-to-adjust pulse length so as to avoid labeling mRNAs
Factory RNA-seq[9]	Human, mouse and Chinese hamster	Straightforward and fast protocol; robust intronic and eRNA coverage; low sequencing depth required; separates chromatin from factory-associated RNA	Side effects of caspase treatment are not clear; not yet adapted to fixed cells
GRO-seq[15]	Human and mouse	Useful for identifying paused RNAPs; robust coverage of enhancer and promoter-associated RNAs	Laborious; relying on antibody affinity for the pull-down; artificial induction of RNAP elongation; potential overestimation of transcript levels
Nascent-/chromatin-seq[18,20]	Human, mouse and fruit fly	Good coverage of intronic RNA; strong mRNA depletion	Extraction under nonphysiological conditions; high sequencing depth required; long-lived chromatin-associated and nascent RNAs mixed
(m)NET-seq[16,17]	Human and yeast	Single-nucleotide resolution; strong mRNA depletion	Extraction under nonphysiological conditions; chromatin-associated and nascent RNAs mixed; eRNA coverage unclear; the 'pull-down' variant[13] relies on antibody affinity
Poly(A)-minus RNA-seq[19]	Human, mouse, fruit fly and yeast	Straightforward, based on total RNA isolation	mRNA depletion is not quantitative; high sequencing depth required; chromatin-associated and nascent RNAs mixed

Each of these approaches has its own particular shortcomings, and all are relatively laborious and/or require a high sequencing depth; most of them also focus on particular parts of the transcriptome. **Table 1** summarizes the advantages and disadvantages of each approach—including factory RNA-seq, which is the approach covered in this protocol. Nonetheless, analysis of nascent RNA can be highly informative, especially when studying short-term responses to stimuli and the associated changes in the levels of messenger, noncoding and enhancer RNAs (eRNAs[23,24]).

In particular, analysis of factory-associated transcripts offers a simple and low-cost way for assessing such transcriptional changes at the level of nascent RNA. However, the current protocol requires freshly isolated cells, and analogous methods that can be applied to fixed cells have not been developed yet. Nevertheless, eRNAs (and other short-lived transcripts) can be detected with high sensitivity (see Caudron-Herger *et al.*[9] for an example), and most chromatin-associated RNA can be studied as a separate fraction, despite factory- and chromatin-associated fractions inevitably sharing components.

Experimental design

Suitable cell types. As already noted, the isolation of transcription factories is performed on living cells collected in a 'physiological' buffer. To date, we have successfully implemented this protocol on human, mouse and Chinese hamster cells, either pluripotent or differentiated; the list includes HeLa cells[8], human umbilical vein epithelial cells (HUVECs)[9], IMR-90s cells, human osteosarcoma (U2OS) cells, human induced pluripotent stem cells and cardiac progenitors, mouse E14 stem cells and embryonic fibroblasts, and Chinese hamster ovary (CHO) cells (M.C.-H., L.B., K.R. and A.P., unpublished data). For the minimum cell numbers required for proteomic or transcriptome analysis, see the PROCEDURE.

Nuclear isolation and run-on experiments. The general workflow for the isolation of factories relies on the efficient isolation of nuclei, which should be optimized for each cell type. Although labeling nascent RNA is not required in the protocol, it may still be included (**Fig. 2**), and it is especially recommended when first setting up the technique in order to quantify retention of transcriptional activity and nascent RNA (**Box 1**). To achieve this, cells are permeabilized using saponin, polymerases are allowed to extend their transcripts (by 'running-on') by fewer than 40 nt in a tagged precursor (e.g., Br-UTP or [^{32}P]UTP) and the amount of label incorporated into RNA is measured (e.g., using anti-BrU antisera and immunofluorescence or scintillation counting[8]). After the run-on, the factory fraction is isolated from labeled nuclei and analyzed exactly as for unlabeled ones.

Proteomic analysis. Isolated factory complexes can be resolved electrophoretically using a 'native' (preparative) 2D acrylamide-agarose composite gel; this step separates the three RNA polymerizing complexes into partially overlapping regions of the gel. Separating the three complexes enables the independent interrogation of the contents that are specifically associated with each polymerase. If such separation of the different RNA polymerizing complexes is not required (e.g., if the bulk of the transcription-related proteome is investigated), the user can omit this step and proceed directly to proteome analysis. Omitting this separation step largely alleviates a major limitation of this method: the need for large numbers of cells (typically more than 5×10^7). Consequently, we envisage the use of quantitative stable isotope

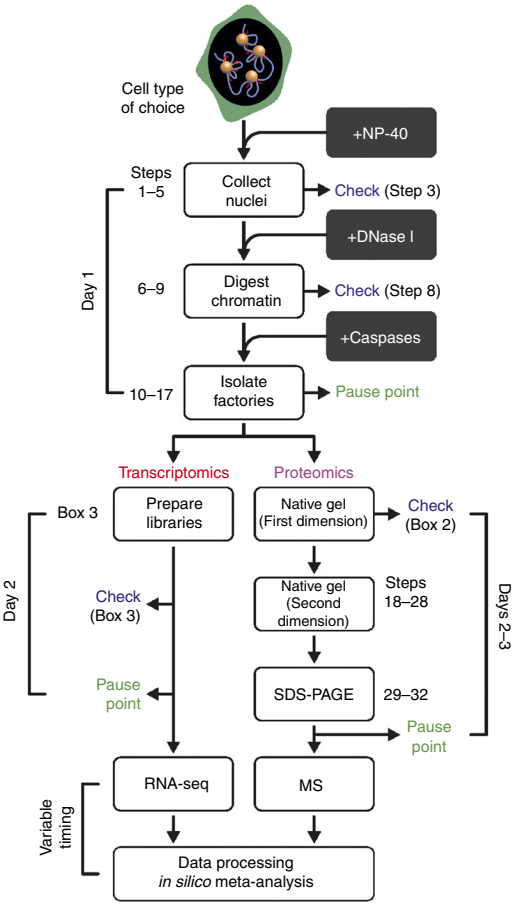

Figure 2 | Workflow for the isolation of transcription factories. This procedure can be applied to perform either an analysis of RNA (transcriptomics; left 'leg') or of proteins (proteomics; right 'leg') associated with transcription factories. The general workflow is presented with an annotation of possible pause points (green), steps at which quality checks are performed (blue) and timing.

labeling of amino acids in cell culture (SILAC)-based[25,26] approaches becoming far more cost-effective, while also offering a means for additional controls (quantitatively assessing contaminants between replicas and/or conditions and so on).

Transcriptome analysis. RNA is isolated from the factory-containing supernatant using TRIzol, and cDNA libraries are prepared for NGS (in our case, on an Illumina platform). For such an analysis, substantially lower cell counts can be used, and mRNAs and eRNAs can be robustly detected using as few as 0.5×10^6 cells (M.C.-H., K.R. and A.P., unpublished data). However, the user should carefully plan analysis to focus on the RNA subfraction of interest. To comprehensively catalog most factory-associated transcripts—including intronic, long noncoding and eRNAs—a sequencing depth of ~35 million reads per replicate suffices[9]; however, if, for instance, only small RNAs are of interest, the user should adapt RNA isolation and sequencing for molecules of smaller size (e.g., using the Illumina TruSeq small RNA library preparation kit). Note also that in the course of isolating factories, additional fractions can easily be retained at the appropriate stages so that total cell, cytoplasmic, nuclear and/or chromatin-associated RNAs can be isolated (**Fig. 1**); subsequent comparison of their contents can provide additional insight into the organization of mammalian transcription. To assess the relative enrichment of sequences mapping to nascent RNA, it is advisable to perform side-by-side analysis of the factory and total (or poly(A)$^+$-enriched) RNA from the same cell type and set of conditions.

Controls. The fractionation achieved using the protocol described here is not perfect, and it may vary slightly between cell types. Hence, some controls are required to monitor the distribution and recovery of protein and/or RNA components of transcription factories along the procedure. It is expected that all three RNAPs will be mostly retained in the final fraction (Step 15; **Fig. 2**), in which no cytosolic or mitochondrial contaminants should be detected. This can be assessed by comparatively analyzing whole-cell and factory preparations of protein and RNA by western blotting and RT-qPCR, respectively; proteins such as MacroH2A.2 and Hsp70, as well as mitochondrial transcripts and tRNAs, will typically not be found in factories. Similarly, if separation of the three RNAP-containing complexes is performed (via 2D electrophoresis), small aliquots of the sample should be used in a set of identical (analytical) mini-gels in order to assess the quality of the preparation (**Box 2**). In addition, the relative positions of the different polymerizing complexes can be traced in these analytical gels, complexes can be excised and subjected to TRIzol-based RNA isolation and RT-qPCR can be used to assess the extent of transcript contamination between complexes (e.g., RNAP III transcripts present in the RNAP II–containing complexes; for an example of such analysis, see Melnik *et al.*[8]).

Optimization of nuclear isolation and enzymatic digestion steps. The protocol described here was designed and optimized using HeLa cells and HUVECs[8,9]. Different cell types may require optimization for the isolation of cell nuclei and incubation times and concentrations for enzymatic digestion using DNase I or caspases.

First, when releasing cell nuclei using physiological buffer (PB) supplemented with NP-40, the different cell types will respond variably. It is advisable to define optimal release conditions by titrating NP-40 concentration (0.2–0.4% (vol/vol)) and by performing 2–5 × 10-min washes on ice using 0.5 million cells in 1 ml of PB. Assess the fraction of lysed cells after each wash by placing a drop of the cell suspension stained with trypan blue on a hemocytometer and examining it under a light microscope.

Second, the optimal conditions for DNase I digestion have to be defined. Divide isolated cell nuclei equally into aliquots and treat them with 0–20 units of DNase I per 10^7 cells for 0–30 min at 33 °C. Stop the reactions by adding EDTA to a final concentration of 2.5 mM, isolate genomic DNA using standard methods and analyze the size of isolated DNA in a 1.5% (wt/vol) agarose gel. Optimal conditions are those that give a smear lacking very large (>5 kbp) and very small (<50 bp) fragments, which are optimally centered around sizes of ~200 bp using the least quantity of enzyme for the shortest incubation time.

Finally, digestion using group III caspases should be optimized by testing increasing enzyme concentrations (0–20 units of caspases per 10^7 cells) and incubation times (0–60 min at 33 °C) on equal aliquots of isolated, DNase I–treated nuclei, in which nascent RNA is ^{32}P-labeled using the run-on procedure (**Box 1**). After stopping the reactions and pelleting by centrifugation,

Box 2 | Native 2D electrophoresis, analytical ● TIMING 3–4 h

To assess the quality of the sample and to determine the relative positions of the complexes containing each RNAP and of the associated RNA (which can be traced, if labeled through the run-on procedure described in **Box 1**), it is worth separating samples in 2D analytical gels before performing the preparative gel described in Steps 18–28 of the main PROCEDURE. This analytical 2D electrophoresis experiment is performed as follows.

1. Cast three identical composite gels side-by-side (e.g., using the Mini-PROTEAN 3 system with gel dimensions of 8.6 × 6.7 × 0.10/0.15 cm). Cast the gels exactly as described for preparative 2D gels (see Steps 18–28 of the main PROCEDURE) using 1-mm-thick spacers for the first-dimension gel and 1.5-mm ones for the second-dimension gel.
2. Load a sample equivalent to 2 × 10^6 cells in each gel.
3. Run gels at a constant voltage of 100 V, until the xylene cyanol dye added to the loading buffer has migrated about three-quarters of the gel's length and the bromophenol blue dye is lost.
4. After conclusion of the electrophoresis experiments, stain one of the gels using Imperial protein stain, according to the manufacturer's instructions, to assess the separation efficiency of the complexes. Place another gel on a Whatman paper, dry it and visualize nascent [^{32}P]RNA by autoradiography using Hyperfilm MP. In both cases, three major stained spots with good separation should be observed (**Fig. 5a**,**b**).
5. Transfer the complexes on the third gel onto a nitrocellulose membrane using the iBlot system (implement program no. 3 for 7 min as detailed in the manufacturer's manual). Use the membrane in western blots with antibodies targeting major subunits of the three mammalian RNAPs (anti-RPA194, anti-7C2 and anti-RPC62, which recognize the RNAP I, II and III subunits, respectively), or other proteins of interest.

▲ **CRITICAL STEP** We recommend confirming efficient protein transfer to the nitrocellulose membrane by staining the membrane using 0.1% (wt/vol) Ponceau S in 5% (vol/vol) acetic acid for 5 min at room temperature. Destain the membrane by immersing in Milli-Q water for an additional 5 min.

▲ **CRITICAL STEP** If the antibody used for the western blot produces a high background noise (as we experienced with the best available antibody to detect RNAP III, chicken anti-RPC62), we recommend blocking membranes using protein-free blocking buffer, rather than 5% nonfat milk in TBS-T (for 30 min at room temperature).

▲ **CRITICAL STEP** After performing chemiluminescence detection, we recommend imaging membranes on a Fuji LAS-4000 scanner (rather than exposing a film). The membrane can be successively probed with antibodies recognizing subunits of RNAP I, II or III; please note that membranes can be stripped up to four times using the RestorePlus buffer for successive blotting with another antibody.

the supernatant is divided into two aliquots: in one aliquot, scintillation counting is used to estimate the effectiveness of nascent RNA release (under optimal digestion settings, it is expected that at least one-third of the total [^{32}P] titers incorporated in the whole-cell RNA should be recovered in this final step), and the other aliquot is analyzed by western blotting using antisera against each of the three RNAPs (e.g., under optimal conditions, western blots using the 7C2 antibody should reveal a prevalent presence of phosphorylated RNAP II isoforms in the factory fraction; **Box 2**).

MATERIALS

REAGENTS

General reagents and chemicals

- Cell type of choice plus appropriate culture media. See Experimental design for a discussion of already-tested cell types ! **CAUTION** All cell lines used should be regularly checked to ensure that they are authentic and not infected with *Mycoplasma*.
- [^{32}P]UTP (3,000 Ci/mmol; PerkinElmer, cat. no. BLU507H250UC)
- [methyl-^{3}H]thymidine (50 Ci/mmol; PerkinElmer, cat. no. NET027W)
- Acrylamide/bis-acrylamide solution, 30% (wt/vol) (37.5:1; Bio-Rad, cat. no. 1610158)
- 5-Bromouridine 5′-triphosphate sodium salt (5-BrUTP; Sigma-Aldrich, cat. no. B7166)
- 6-Aminocaproic acid (Sigma-Aldrich, cat. no. A7824)
- Acetic acid (Sigma-Aldrich, cat. no. A6283)
- Ammonium persulfate (APS; Sigma-Aldrich, cat. no. A3678)
- ATP disodium salt (Na_2ATP; Sigma-Aldrich, cat. no. A2383)
- Blue carrier immunogenic protein (Thermo Scientific, cat. no. PI77661)
- Bromophenol blue (Sigma-Aldrich, cat. no. 32712)
- $CaCl_2$·$2H_2O$ (Sigma-Aldrich, cat. no. C3306)
- Coomassie blue G-250 (Serva, cat. no. 17524)
- Diethyl pyrocarbonate (DEPC; Sigma-Aldrich, cat. no. D5758)
- DMSO (Sigma-Aldrich, cat. no. D5879)
- Direct-zol RNA MiniPrep kit (Zymo Research, cat. no. R2050)
- DTT (Sigma-Aldrich, cat. no. D9779)
- EDTA disodium salt (Sigma-Aldrich, cat. no. 4884)
- Glycine (Sigma-Aldrich, cat. no. G8898)
- Hydrochloric acid (HCl), 37% (wt/vol) (Sigma-Aldrich, cat. no. H1758)
- Hyperfilm MP (GE Healthcare, cat. no. 28-9068-42)
- iBlot nitrocellulose transfer stacks, regular size (Life Technologies, cat. no. IB23001)
- Igepal (NP-40; Sigma-Aldrich, cat. no. I8896)
- Imperial protein stain (Thermo Scientific, cat. no. PI24615)
- KCl (Sigma-Aldrich, cat. no. P9541)
- KH_2PO_4 (Sigma-Aldrich, cat. no. P9791)
- Laemmli sample buffer, 4× (Bio-Rad, cat. no. 1610747) ▲ **CRITICAL** Add β-mercaptoethanol to 1/10 of buffer volume before use.
- $MgCl_2$·$6H_2O$ (Sigma-Aldrich, cat. no. M2670)
- NaCl (Sigma-Aldrich, cat. no. S3014)
- Na_2HPO_4 (Sigma-Aldrich, cat. no. S3264)
- Na_3VO_4 (Sigma-Aldrich, cat. no. S6508)
- NaF (Sigma-Aldrich, cat. no. 30105)
- NEBNext Ultra directional RNA library prep kit (New England BioLabs, cat. no. E7420S)
- NTPs (100 mM each; Roche, cat. no. 11277057001)
- PBS, pH 7.4 (Thermo Fisher Scientific, cat. no. 1001023)
- Ponceau S (Sigma-Aldrich, cat. no. P3504)
- Potassium acetate (Sigma-Aldrich, cat. no. P1190)
- Protein-free Tris-buffered saline (TBS) blocking buffer (Thermo Scientific, cat. no. PI37570)
- RestorePlus western blot stripping buffer (Thermo Scientific, cat. no. PI46428)

- Ribo-Zero rRNA removal kit (Illumina, cat. no. MRZH116)
- Saponin (Sigma-Aldrich, cat. no. 47036)
- SeaKem Gold agarose (Lonza, cat. no. 50152) SDS (Sigma-Aldrich, cat. no. L3771)
- Sucrose (Sigma-Aldrich, cat. no. S0389)
- SYBR Green II nucleic acids gel stain (Life Technologies, cat. no. S33102)
- Tetra-methyl-ethylenediamine (TEMED; Sigma-Aldrich, cat. no. T22500)
- Tris (Sigma-Aldrich, cat. no. T1503)
- Triton X-100 (Sigma-Aldrich, cat. no. T8787)
- TRIzol LS (Life Technologies, cat. no. 10296-010) **! CAUTION** TRIzol-based solutions are toxic, and they must be handled and disposed of according to established national and institutional regulations.
- TruSeq Small RNA library preparation kit (Illumina, cat. no. RS-200-0012)
- Trypan blue solution, 0.4%, wt/vol (Sigma-Aldrich, cat. no. T8154)
- Tween 20 (Sigma-Aldrich, cat. no. P9416)
- Whatman filter paper, grade 1 (Sigma-Aldrich, cat. no. Z240087)
- Xylene cyanol (Sigma-Aldrich, cat. no. X4126)
- β-Glycerophosphate (Sigma-Aldrich, cat. no. G9422)
- β-Mercaptoethanol (Sigma-Aldrich, cat. no. M3148)

Enzymes and inhibitors

- Active human caspases, group III (6,8,9,10) mix (BioCat, cat. no. K243-4-25-BV)
- Complete protease inhibitor cocktail, EDTA-free (PIC; Roche, cat. no. 6538282001)
- DNase I (Worthington, cat. no. LS006331) **▲ CRITICAL** If another brand is used, check for enzyme activity in PB.
- Group III caspase inhibitor (Calbiochem, cat. no. 2188745)
- RNase inhibitor (RiboLock; Thermo Scientific, cat. no. E00384)

Antibodies

- Chicken polyclonal anti-RPC62 (Abcam, cat. no. ab26185; used at 1/1,000 dilution); this antibody is used to detect RNAP III
- Mouse monoclonal anti-RPA194 (Santa-Cruz Biotechnology, cat. no. sc-48385; used at 1/100 dilution); this antibody is used to detect RNAP I
- Mouse monoclonal anti-RPB1 (7C2 (ref. 27), used at 1/10,000 dilution); this antibody is used to detect the (phosphorylated and nonphosphorylated) C-terminal domain of the largest catalytical subunit of RNAP II

EQUIPMENT

- For proteomic analysis of factory fractions: access to a mass spectrometer. For our experiments, we used a LTQ Orbitrap mass spectrometer (Thermo Scientific) coupled to a U3000 nanoHPLC system (Dionex);
- For transcriptome analysis of factory fractions: access to a NGS sequencer. For our experiments, we used a HiSeq 2000 sequencer (Illumina)
- Cooling microcentrifuge (Hermel, cat. no. Z216MK)
- iBlot dry blotting system (Life Technologies, cat. no. IB1001)
- Milli-Q integral water purification (type 2) system (Merk Millipore, cat. no. NF-C72876)
- Mini-PROTEAN 3 electrophoresis cell (Bio-Rad, cat. no. 1658000)
- NanoDrop spectrophotometer (LabTech, ND-1000)
- PROTEAN II xi electrophoresis cell (Bio-Rad, cat. no. 1651801) with 1.0- (cat. no. 1651843) and 1.5-mm (cat. no. 1651844) spacers, plus 1-mm ten-well (cat. no. 1651877) and, 1.5-mm five-well (cat. no. 1651883) combs.
- Thermomixer (HLC, cat. no. DITA15001)
- Disposable stainless steel scalpels (Thermo Scientific, cat. no. 31-200-32)

REAGENT SETUP

▲ CRITICAL Prepare all solutions using Milli-Q water or diethyl pyrocarbonate (DEPC)–treated water in glassware decontaminated by baking (>180 °C, overnight) or in nuclease-free plasticware.

Complete protease inhibitor, EDTA-free (PIC), 50× Dissolve one tablet of PIC in 1 ml of PBS to create a 50× working solution. Store the solution at 20 °C for up to 3 months. **▲ CRITICAL** Avoid repeated freeze-thaw cycles.

Tris acetate buffer, 1 M (pH 7.4) Dissolve 121.14 g of Tris in 800 ml of DEPC-treated or Milli-Q water and titrate the pH to 7.4 using glacial acetic acid; adjust the volume to 1 liter and autoclave the solution. Store this buffer at 4 °C for a maximum of 12 months.

Tris-HCl buffer, 1 M (pH 7.4) Dissolve 121.14 g of Tris in 800 ml of DEPC-treated or Milli-Q water and titrate the pH to 7.4 using 37% (wt/vol) hydrochloric acid; adjust the volume to 1 liter and autoclave the buffer. Store this buffer at 4 °C for a maximum of 12 months.

Complete PB This method of preparation is adapted from Jackson *et al.*[10]. Prepare a solution of basal PB, which contains 100 mM potassium acetate, 30 mM KCl, 10 mM Na_2HPO_4 and 1 mM $MgCl_2$. Before each experiment, supplement the basal PB solution with 1 mM Na_2ATP, 1 mM DTT, 25 units/ml RiboLock, 10 mM β-glycerophosphate, 10 mM NaF, 0.2 mM Na_3VO_4 and a 1/1,000 dilution of PIC. Always use ice-cold PB. Basal PB (not containing ATP) can be stored at 4 °C for up to 6 months.

▲ CRITICAL As the acidity of each ATP batch varies, use 100 mM KH_2PO_4 to adjust the pH to 7.4 with the highest possible precision.

NLB, pH 7.4 This method of preparation is modified from Novakova *et al.*[28]. Prepare a solution containing 40 mM Tris-acetate, 2 M 6-aminocaproic acid, 7% (wt/vol) sucrose, 1/1,000 dilution of PIC and 50 units/ml RiboLock. Before addition of PIC and RiboLock, this solution can be stored at −20 °C for up to 6 months. **▲ CRITICAL** Always add PIC and RiboLock immediately before use.

'Run-on' mix Immediately before use, add enough components to PB to obtain a solution of 100 μM ATP, 100 μM CTP, 100 μM GTP, 0.1 μM UTP, 50 μCi/ml [^{32}P]UT,P or 80 μM 5-bromo-UTP, and then add $MgCl_2$ to reach a final concentration of Mg^{2+} ions that is equimolar to that of all triphosphates. **▲ CRITICAL** Always prepare this mix immediately before use.

Composite gel buffer, 5× This method of preparation is modified from Nadano *et al.*[29]. Prepare a solution containing 200 mM Tris-acetate (pH 7.4), 35% (wt/vol) sucrose and 0.05% (vol/vol) Triton X-100. Store this buffer at 4 °C for a maximum of 6 months.

Composite gel running buffer, 10× Prepare a solution containing 400 mM Tris-acetate (pH 7.4). Store this buffer at 4 °C for up to 12 months.

Coomassie solution Prepare a solution containing 500 mM 6-aminocaproic acid and 2% (wt/vol) Coomassie blue G-250. Store this solution at 4 °C for a maximum of 6 months.

TBST buffer Prepare a solution containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.05% (vol/vol) Tween 20. Store this solution at room temperature (22 °C) for a maximum of 12 months.

Tris-glycine running buffer, 10× Prepare a solution containing Tris (pH 8.3), 1.92 M glycine and 1% (wt/vol) SDS. Store this buffer at 4 °C for a maximum of 6 months.

Group III caspase inhibitor solution Dissolve the inhibitor in DMSO to prepare a 20 mM stock solution. Store it at −20 °C for a maximum of 3 months, and protect it from moisture.

PROCEDURE

Preparation of cells ● TIMING 1 h

1| Grow the cells of choice in suspension (e.g., HeLa cells in spinner culture[8]) or as a monolayer on 15-cm culture dishes (e.g., HUVECs[9]) using the appropriate medium and culture conditions. Note that cells grown in suspension ensure a smaller loss than those grown as a monolayer, as they do not need to be scraped or trypsinized off the surface. We typically start with 5 × 10^7 cells for proteomic analyses and 4–8 × 10^6 cells for nascent RNA sequencing (see **Fig. 3** for data on the use of lower cell counts).

2| Collect cells according to either option A or option B, depending on whether the cells were cultured as a monolayer or in suspension, respectively. If DNA recovery needs to be monitored, scintillation counting can be used; for this purpose, add to the culture medium [methyl-^{3}H]thymidine to a final concentration of 0.25 μCi/ml a day before collecting the cells to achieve uniform labeling by incorporation into cellular DNA[8].

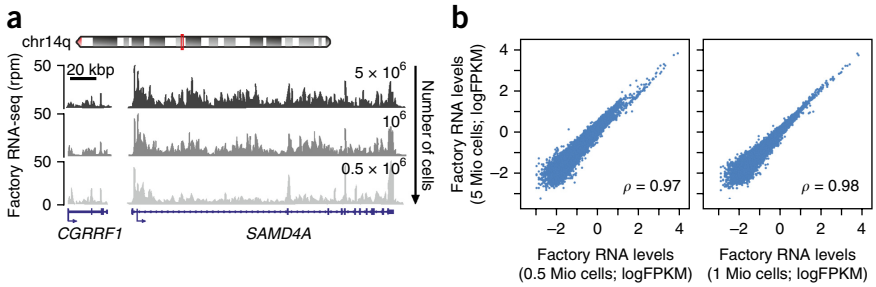

Figure 3 | Factory RNA sequencing performed on low cell counts. (**a**) Typical genome-browser profiles (in reads per million; rpm) along two neighboring genes on human chromosome 14, *CGRRF1* and *SAMD4A*, obtained by deep-sequencing factory RNA prepared from 5 (black), 1 (gray) and 0.5 million (light gray) osteosarcoma U2OS cells. (**b**) Correlation plots of gene expression levels obtained using factory RNA from 0.5–5 million (Mio) cells. Calculated Spearman's correlation coefficients (ρ) indicate the reproducibility of the results despite using decreasing cell counts.

(A) Collection of cells grown as a monolayer

(i) Split cells the day before collection in a ratio that will ensure that they reach 75–85% confluence at the point of collection. If you intend to measure DNA recovery, add [methyl-^{3}H]thymidine to the culture medium when splitting the cells.

(ii) On the next day, aspirate the medium from plates and immediately place them on ice.

(iii) Add 2.5 ml of ice-cold PB per plate. Gently scrape the cells and transfer them to a 50-ml Falcon tube.
▲ **CRITICAL STEP** To minimize cell loss and cell damage, we strongly recommend using soft rubber scrapers for this step.

(iv) Collect cells by centrifugation of the Falcon tube at 400*g* for 5 min at 4 °C. Discard the supernatant, and place the tubes with cell pellets on ice. Please note that we have verified that the immediate transfer of cells to ice 'freezes' the polymerizing activity during the response to the inflammatory cytokine tumor necrosis factor-α (TNF-α) at the desired time point (see Caudron-Herger *et al.*[9]).

(B) Collection of cells grown in suspension

(i) Collect the cells by centrifugation at 400*g* for 5 min at 4 °C.

(ii) Discard the supernatant, and place the tubes with cell pellets on ice. As above, the immediate transfer of cells to ice 'freezes' the polymerizing activity during the response to TNF-α at the desired time point (see Caudron-Herger *et al.*[9]).

Isolation of nuclei and chromatin digestion ● TIMING 1–1.5 h

3| Resuspend cells in 5 ml of ice-cold PB containing 0.4% (vol/vol) NP-40; incubate the resulting mixture on ice for 20 min. If desired, collect an aliquot equivalent to ~2.5 × 10^5 cells (directly upon resuspension) to use for whole-cell RNA isolation. This can serve as a control for assessing the relative enrichment of nascent RNA at the end of the PROCEDURE (see ANTICIPATED RESULTS).
▲ **CRITICAL STEP** Different cell types respond variably to this treatment. When releasing HUVEC nuclei, for example, two 10-min washes with PB + 0.4% (vol/vol) NP-40 are optimal. Assess the fraction of lysed cells by placing a drop of the cell suspension after each wash on a hemocytometer and examining under a light microscope (see Experimental design).

4| Centrifuge the tubes containing the resuspended cells at 400*g* for 5 min at 4 °C to obtain a pellet of the nuclei, and remove the supernatant.

5| Wash the pellets briefly once more in PB + 0.4% (vol/vol) NP-40, and pellet the nuclei once more as described in Step 4.

6| Gently resuspend the pellet in PB + 0.4% (vol/vol) NP-40 supplemented with 0.5 mM $CaCl_2$ (use 100 µl of the solution for every 10^7 cells), and transfer it to a 2-ml tube.
▲ **CRITICAL STEP** From this step onward, use 2-ml round-bottom (nuclease-free) tubes.

7| Add to the resuspended cell nuclei DNase I (10 units per 10^7 cells), and then incubate the resulting mixture for 30 min at 33 °C.
▲ **CRITICAL STEP** Make sure that incubation is carried out without any shaking to avoid aggregation.
? TROUBLESHOOTING

8| Quench the reaction by adding EDTA to a final concentration of 2.5 mM and cooling on ice. Isolate DNA from an aliquot of ~2.5 × 10^5 cells to verify the extent of DNA fragmentation by DNase I (a smear centered around ~200 bp is expected; **Fig. 4a**).
? TROUBLESHOOTING

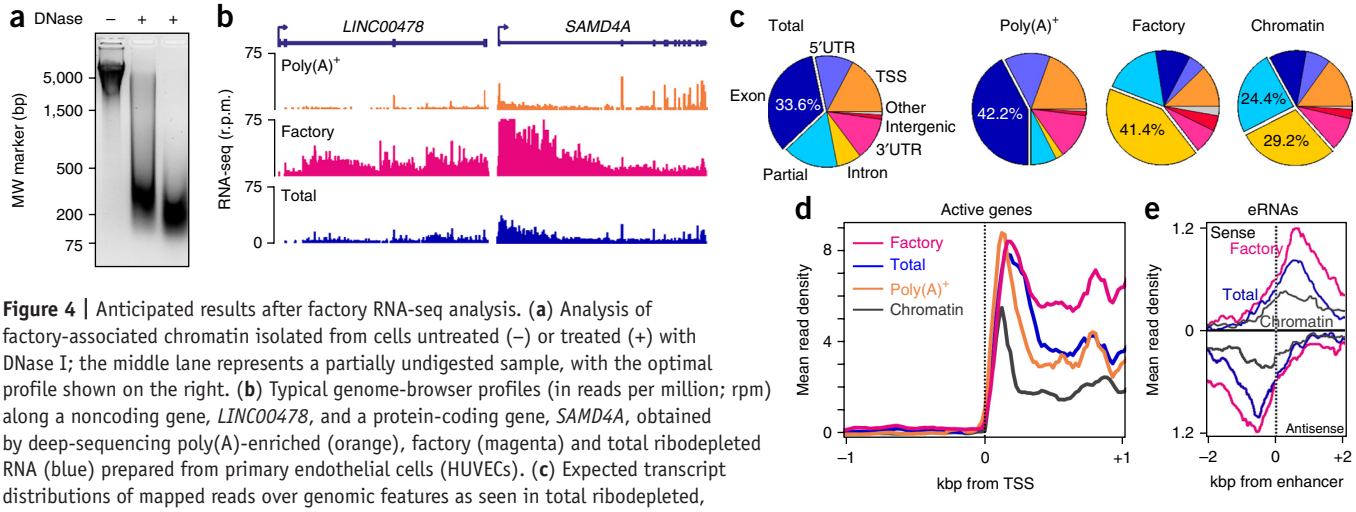

Figure 4 | Anticipated results after factory RNA-seq analysis. (**a**) Analysis of factory-associated chromatin isolated from cells untreated (–) or treated (+) with DNase I; the middle lane represents a partially undigested sample, with the optimal profile shown on the right. (**b**) Typical genome-browser profiles (in reads per million; rpm) along a noncoding gene, *LINC00478*, and a protein-coding gene, *SAMD4A*, obtained by deep-sequencing poly(A)-enriched (orange), factory (magenta) and total ribodepleted RNA (blue) prepared from primary endothelial cells (HUVECs). (**c**) Expected transcript distributions of mapped reads over genomic features as seen in total ribodepleted, poly(A)-enriched, factory and chromatin-associated RNA. The percentage of reads belonging to the prevailing feature per experiment is shown. (**d**) Average read density for 1 kbp upstream and downstream of the transcription start sites (TSS) of ~8,000 active genes plotted using RNA-seq data from total ribodepleted (blue), poly(A)-enriched (orange), factory-associated (magenta) and chromatin-associated (black) RNA. (**e**) Average eRNA read density for 2 kbp upstream and downstream of ~1,300 active HUVEC enhancers plotted using RNA-seq data from total ribodepleted (blue), factory-associated (magenta) and chromatin-associated (black) RNA. All data for **a**–**e** were generated using HUVECs. Adapted from Caudron-Herger *et al.*, *Nucleic Acids Res.* **43**, e95 (2015) (ref. 9), by permission of Oxford University Press.

9| Centrifuge the mixture containing the digested nuclei at 600*g* for 5 min at 4 °C. Separate the supernatant and pellet it, but keep both samples. The pelleted nuclei are used in subsequent steps; the supernatant contains 'chromatin-associated' RNA and can be processed for subsequent RNA-seq by implementing the instructions in **Box 3**, or it can be stored at –80 °C for a maximum of 6 months.

Caspase digestion and factory release ● TIMING 1–1.5 h

10| Resuspend the nuclei in NLB. Although we recommend using 50 µl of NLB solution per 10^7 cells, larger volumes may be used if efficient resuspension of the nuclei cannot be achieved with the recommended volume. Incubate the resulting mixture on ice for 20 min.

▲ **CRITICAL STEP** Vortex the pellet vigorously at this step to disaggregate clusters of nuclei. Always use the smallest possible volume of NLB.

11| Allow the vials containing lyophilized caspases 6, 8, 9 and 10 to thaw on ice, and resuspend them in PB + 0.4% (vol/vol) NP-40 to achieve a final concentration of 2 units of each caspase per µl of the buffer. Snap-freeze small aliquots (10 µl) of the caspases mix in liquid nitrogen and store them at –80 °C for a maximum of 12 months. For each experiment, always use a freshly thawed aliquot.

12| Add the freshly prepared solution of caspases to the sample prepared in Step 10: 2 units of caspase mix per 10^7 cells (or the analogous amount optimized per cell type; cell numbers refer to the starting material) and incubate the sample at 33 °C for 30 min in a Thermomixer with vigorous (~900 r.p.m.) shaking.

13| Place the reaction tube on ice and quench the reaction by adding group III caspase inhibitor to a final concentration of 0.2 mM.

14| After a brief (1–2 min) incubation on ice, centrifuge the mixtures from Step 13 at 600*g* for 5 min at 4 °C.

15| Collect the supernatant and treat it with DNase I, as detailed in Steps 7 and 8; this is the fraction that contains large fragments of factories.

16| Assess the quality of the large fragments by performing a 2D native electrophoresis in (small) analytical gels according to the directions in **Box 2**.

? TROUBLESHOOTING

17| If you are planning to sequence factory-associated RNA, proceed with the directions provided in **Box 3**; for proteomic analysis, implement Steps 18–32 of the PROCEDURE.

Box 3 | Preparation of cDNA libraries and RNA sequencing ● TIMING 5 h

The RNA obtained as a result of the protocol described here can be analyzed using high-throughput NGS. Exemplary steps for this are presented below.

1. Directly mix the supernatant from Step 15 (factory fraction) or Step 9 (chromatin fraction) of the main PROCEDURE with 0.75–1 ml of TRIzol, and allow the sample to sit at room temperature for 5 min.

■ **PAUSE POINT** Samples in TRIzol can be snap-frozen in liquid nitrogen and stored at –80 °C for a maximum of 6 months.

2. Purify factory RNA (and, if desired, chromatin-associated RNA), by implementing the TRIzol reagent manufacturer's instructions or using a kit. We typically use the Direct-zol RNA MiniPrep kit, including a DNase I treatment step.

▲ **CRITICAL STEP** If you are planning to only examine small RNAs, e.g., miRNAs of 20–22 nt, this can be achieved using this same purification kit and a downstream dedicated cDNA library construction kit (e.g., TruSeq small RNA library preparation kit).

▲ **CRITICAL STEP** Before ribodepletion (step 3), the RNA purified in this step can also be used in standard RT-qPCR tests to verify its enrichment for nascent (intronic) RNA in comparison with total cell RNA preparations. For this, we typically use primers spanning exon-intron junctions of actively transcribed genes in the cell type of interest (and exon-exon junctions as controls, although these should be largely depleted from factory RNA preparations).

? TROUBLESHOOTING

3. Deplete 1–5 µg of purified RNA of rRNA species using the Ribo-Zero rRNA removal as per the manufacturer's instructions.

4. Use ~1 µg of rRNA-depleted RNA from step 3 as template for the NEBNext Ultra directional RNA library prep kit to generate barcoded cDNA libraries for sequencing on an Illumina platform as per the manufacturer's instructions. If desired, 'spike' RNA preparations destined for sequencing with exogenous RNA of known concentrations to facilitate the quantitative comparison of RNA-seq data across cell types and conditions[34].

5. Multiplex and sequence cDNA libraries on a HiSeq 2000 platform (Illumina). For most purposes, 30–50 million read pairs (each read is 75–100 bp in length) per samples, in at least two biological replicates, should suffice to obtain robust genome-wide coverage of intronic RNA and the detection of low-level, unstable, species, such as eRNAs[19,20]. Although RNA-seq data may be analyzed using standard tools (see Caudron-Herger *et al.*[9]), there now exist dedicated *in silico* pipelines[35,36] for assessing changes in gene expression based only on intronic RNA levels (instead of exonic levels, as used by standard algorithms), and we anticipate that their implementation will further enhance the sensitivity of factory RNA-seq.

(Optional) native 2D electrophoresis, preparative ● TIMING 36–48 h

▲ **CRITICAL** Depending on the aims of the experiment, it is possible to skip Steps 18–28 to proceed directly to MS (see Steps 29–32) of the supernatant isolated in Step 15, without separating the different RNAP-containing complexes by 2D electrophoresis; this approach can reduce the starting amount of cells required for analysis.

▲ **CRITICAL** Please note that the conditions reported below for electrophoresis in a native 2D gel are modified from those reported previously[28,29].

18| Cast first-dimension gel (0.7% (wt/vol) agarose, 1.5% (wt/vol) acrylamide in 40-mM Tris acetate, pH 7.4, 7% (wt/vol) sucrose and 0.01% (vol/vol) Triton X-100), for example, using the PROTEAN IIxi system (gel dimensions of 16 × 20 × 0.15 cm) using a 1.5-mm, five-well wide-tooth comb. Assemble the casting plates and, using water, measure the exact volume required for one such gel, and prepare it as detailed in the in-text table below (example for a 10-ml gel is given; scale amounts up proportionately for larger gel volumes).

Component	Amount	Final concentration (wt/vol)
Agarose (SeaKem Gold)	0.07 g (in 7.425 ml water)	0.7%
Acrylamide (30%, wt/vol; 37.5:1)	0.5 ml	1.5%
Composite gel buffer (5×)	2.0 ml	1×
TEMED (100% (wt/vol))	15 µl	0.15%
APS (25% (wt/vol))	60 µl	0.15%
Total	10 ml	–

▲ **CRITICAL STEP** Weigh the 0.07 g of agarose in a glass beaker, add 7.425 ml of Milli-Q water and weigh the beaker before melting the agarose in a microwave oven. After slowly bringing the gel to boiling, weigh the beaker again and add back the amount of water that evaporated. In addition, preheat the casting plates to 37 °C before pouring the gel. Pour the gel and allow it to polymerize at room temperature for at least 1 h before conducting the electrophoresis.

19| Load the preparative gels with a sample equivalent to that isolated from 4–5 × 10^7 cells (not radioactively labeled). Add to the sample bromophenol blue to a final concentration of 0.004% (wt/vol) and xylene cyanol to a final concentration of 0.004% (wt/vol) to monitor migration.

20| (Optional) If you are planning to separate and/or isolate complexes containing RNAPs II and III, add Coomassie blue solution (1/1,000 dilution) to the sample to be loaded onto the gel, incubate for 30 min on ice and add Coomassie solution into the 1× cathode buffer at a 1/100 dilution.
▲ **CRITICAL STEP** If you are intending to separate and/or isolate complexes containing RNAP I (nucleolar factories), do not add Coomassie blue solution to the sample, as Coomassie blue affects migration of the RNAP I complexes; in this case, omit Step 20 and proceed directly to sample loading (Step 21).

21| Load the sample onto the first-dimension gel. Use blue carrier immunogenic protein (200 ng in NLB with 1/1,000 Coomassie solution) as a protein marker (8 MDa).

22| Run the gel at a constant electric potential of 100 V in 1× composite running buffer until the sample has fully entered the gel. Then, reduce the potential to 40 V and carry out the electrophoresis experiment overnight in a cold room (4–8 °C). At the end of the run, xylene cyanol will have migrated ~3/4 of the length of the gel, and bromophenol blue will be lost.

23| After electrophoresis, cut a sample-containing strip out of the gel using a clean scalpel.

24| Cast the second-dimension gel using 2-mm spacers; please note that no comb is used to prepare the second-dimension gel. The composition and running conditions of the second-dimension gel are the same as for the first-dimension gel (see Step 18).

25| To polymerize the sample-containing strip to the second-dimension gel, prepare a small amount of composite gel and pour it over the sample strip, as it lays on top of the second-dimension gel. This procedure takes place with the gel placed horizontally; make sure that all gaps between the strip and the already-polymerized gel are filled.
▲ **CRITICAL STEP** Avoid air bubbles between the sample strip and the second-dimension gel. To use blue carrier immunogenic protein as a marker, insert a 1.5-mm-thick spacer vertically next to the sample strip to cast a single well. Cover the top of the gel with Milli-Q water, and let it cool for at least 1 h.

26| Run the second-dimension gel at 40 V for 8–12 h in a cold room.

27| After electrophoresis, stain the gel by adding to it SYBR Green II nucleic acids gel stain according to the manufacturer's instructions, and rock the resulting mixture gently for 20–30 min at room temperature.

28| Locate the complexes separated in the stained gel by comparison of the relative positions of the different RNAP complexes with those obtained after running the analytical gels at Step 16. Excise the appropriate complex(es) from the gel using a clean scalpel.
? TROUBLESHOOTING

Preparation of samples for MS ● TIMING 1–2 h
29| Cast a standard 7% (wt/vol) SDS-polyacrylamide gel (gel dimensions of 8.6 × 6.7 × 0.15 cm) and 1.5-mm five-well comb.

30| Equilibrate each excised gel piece from the second-dimension run in 1× Tris-glycine running buffer for 10 min at room temperature. Repeat this step once more. Note that if Steps 18–28 are omitted, the equilibrated gel pieces are substituted by the sample obtained at Step 15.

31| Load each equilibrated gel piece or each sample obtained at Step 15 into a separate well of the SDS-polyacrylamide gel, together with a small volume of 1× Laemmli sample buffer to enable monitoring of migration. Carry out electrophoresis for 10–15 min after the dye has entered the separating part of the gel at a constant electric potential of 100 V to ensure that all samples have completely migrated from the stacking into the separating gel.

32| Once electrophoresis concludes, excise the whole band containing the dye from the gel that now contains the concentrated proteins in resolving gel (10- to 15-mm-long gel slice). The excised gel slice can now be cut in small pieces with a clean scalpel, and it can be subjected to trypsin digestions and extraction[30] for MS according to the specifications of the instrument used.
▲ **CRITICAL STEP** To avoid excessive contamination by skin keratin, it is advisable to handle all related equipment and consumables while wearing gloves.

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 2**.

TABLE 2 | Troubleshooting table.

Step	Problem	Possible reason	Solution
7	Aggregation of isolated nuclei, as determined by inspection under a light microscope using a hemocytometer	This is a cell type–specific effect, and it can be aggravated by DNA leakage	In Step 6, optimize the detergent (NP-40) concentration and treatment times for each cell type (see Experimental design). Dilute nuclei into larger volumes; add PB to the nuclei in a stepwise manner
8	Poor DNase I digestion, as assessed by DNA electrophoresis in agarose gels (see Experimental design)	The coating of cell culture containers is interfering with the reaction or the brand of DNase I used is not compatible with the PB buffer	We have observed that DNase I from different manufacturers display variable cutting efficiencies when treating cells grown on coated cell culture containers (e.g., mouse embryonic stem cells on gelatin-coated plates). In these cases, trypsinization is preferred to scraping and cells should then be washed twice in PB. DNase I performance is also pH-dependent, and it can be inefficient if the PB buffer has a pH >7.4
16	Very low yield in factory fraction	Poor DNase I and/or caspase digestion	Isolate RNA from the pellet after Step 15 and quantify it relative to the released factory fraction. If the protocol is not being applied to mammalian cells, investigate the efficiency of caspase digestion in your system. It is also critical to get complete solubilization of the pellet in NLB buffer; no pellet clumps should be observed. For poor DNase I digestion, see Troubleshooting advice for Step 8
28	Large overlap between the three different RNAP complexes	Poor DNase I and/or caspase digestion, or poor electrophoretic separation	See suggestions above. Also check that 2D gel concentration and electrophoresis parameters are exactly as described and that the gels are well polymerized
Box 1, step 8	Poor run-on signal	Cells not in the exponential growth phase	Split cells the day before the experiment
Box 3, step 2	Highly degraded RNA profiles	High titers of free bivalent cations in PB, or RNase contamination	Do not exceed the proposed 0.5 mM of $CaCl_2$ for DNase digestion. Always use buffers prepared with DEPC-treated water, RNase inhibitors and RNase-free tubes and glassware

● TIMING

Steps 1 and 2, preparation of cells: 1 h
Steps 3–9, isolation of nuclei and chromatin digestion: 1–1.5 h
Steps 10–17, caspase digestion and factory release: 1–1.5 h
Steps 18–28, (optional) native 2D electrophoresis, preparative: 36–48 h
Steps 29–32, sample preparation MS: 1–2 h plus MS and analysis
Box 1, labeling nascent RNA by run-on: 2 h
Box 2, native 2D electrophoresis, analytical: 3–4 h
Box 3, preparation of cDNA libraries for RNA sequencing: 5 h plus high-throughput sequencing

ANTICIPATED RESULTS

This protocol generates purified RNA and/or proteins associated with large fragments of transcription factories. Depending on the isolated fraction and its desired analysis, anticipated results and quality controls differ.

The efficiency of the isolation of factory-associated, nascent, RNA can be assessed as follows. First, electrophoretic analysis of the RNA isolated using this protocol may not display the typical profile of whole-cell RNA with its prominent rRNA bands (note that it is strongly recommended to isolate whole-cell RNA from the starting material to use as control; see Step 3).

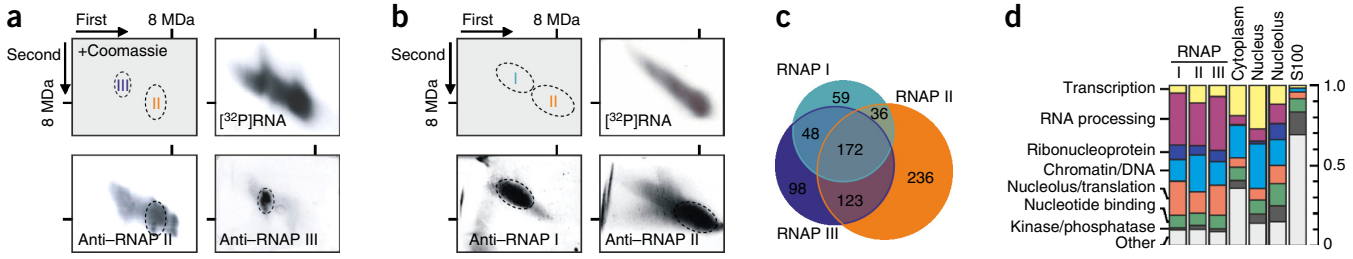

Figure 5 | Anticipated results after factory proteomic analysis. (**a**) Resolving RNAP complexes II and III by native 2D electrophoresis using Coomassie blue in the first dimension. The cartoon (top left) shows regions excised for MS. An autoradiograph (top right) of the gel shows overlapping spots of (nascent) [^{32}P]RNA along the diagonal. The region with prominent signal contained ~0.03, ~0.8 and ~5% of the protein, DNA and [^{32}P]RNA initially present, respectively. The protein sample was next transferred onto nitrocellulose and immunoblotted successively for RNAP II (bottom left) and III (bottom right). (**b**) Resolving RNAP I complexes by native 2D electrophoresis in the absence of Coomassie blue in the first dimension. All details are as in **a**, except for immunoblotting for RNAP I (bottom left) and RNAP II (bottom right). (**c**) The Venn diagram shows the number of proteins detected per RNAP complex and their overlap. (**d**) The bar plot shows a comparison of Gene Ontology (GO) terms associated with the proteins identified in each RNAP fraction, as well as the cytoplasmic, nuclear and nucleolar fractions, plus those in S100 cytosolic extracts. All data for **a** and **d** were generated using HeLa cells. Adapted with permission from Melnik *et al.*[8].

This discrepancy should be ignored, unless excessive degradation is observed—i.e., if the majority of RNA fragments form a smear at around 50–100 nt. Next, RT-qPCR can be used to assess enrichment of the factory RNA fraction for intronic nascent transcripts and/or eRNA (in comparison with whole-cell RNA); similarly, exonic RNA levels should be depleted (but not absent) in factory versus whole-cell preparations. Of course, as high-throughput transcriptomics are becoming routine, end-users can also perform quality controls directly on the data generated after RNA sequencing. In this case, we also recommend performing side-by-side RNA-seq of the factory and total or poly(A)$^+$-enriched fractions from the same cell type and set of conditions to facilitate comparisons. After sequencing of the respective libraries and standard mapping to the reference genome of interest[31], RNA levels between replicates should correlate well, as is indicative of high reproducibility (even if using increasingly lower cell counts; **Fig. 3b**). Moreover, <10% of mapped reads should map to rDNA regions (given that RNA preparations are ribodepleted; **Box 3**), and these reads should be excluded from downstream analyses. Visual inspection of the data on a genome browser should reveal a prevalence of signal from introns, intron-exon junctions and intergenic space in factory versus total/poly(A)$^+$ RNA-seq data (**Fig. 4b**). In addition, analysis of the read distribution profiles from each RNA-seq approach should confirm this prevalence (**Fig. 4c**,**d**). Similarly, as a key advantage of the approach is the high sensitivity of eRNA detection, examination of their levels should also reveal an enrichment in the factory RNA-seq samples (**Fig. 4e**).

The quality of the isolation of proteins associated with factories may be assessed as follows. If RNAP I, II and III complexes are separated by native 2D electrophoresis (Steps 18–28), the degree of overlap on the gel should be such that each complex can be excised with minimal contamination (**Fig. 5a**,**b**). This is monitored by western blots using antisera targeting major RNAP I, II or III subunits, and/or by visualization of the RNA associated with these complexes by staining with SYBR Green or by [^{32}P]UTP incorporation and autoradiography (**Box 2**). In addition, nonoverlapping parts of the three complexes can be excised from 2D gels or RNA-purified using TRIzol, and RT-qPCR can be performed to determine relative enrichments—e.g., intronic RNA of an active protein-coding gene should be detected in the RNAP II complex, but not in the RNAP I or III complexes[8]. Such controls apply to any cell type of interest.

If complexes of each of the three polymerases are not resolved (i.e., if proteins are directly purified from the sample collected at Step 15), western blots can be used to verify that the fraction isolated is rich for RNAP subunits, and not for proteins belonging to the cytoplasmic or mitochondrial compartment, in comparison with whole-cell lysates. Again, as proteomic analyses are now becoming routine and cost-effective, the users may choose to directly interrogate their preparations using MS. Depending on the instrument used and the implementation (or not) of metabolic protein labeling[25], results and analyses will vary. However, some quality assessments can be universally implemented. For example, we previously[8] used an established pipeline[32,33] to analyze complexes resolved in native 2D gels. First, the quality of the preparations is estimated from the extent of overlap between the RNAP I complex components[8] and the proteome of isolated nucleoli[13] (≥80% overlap was observed). Next, biological replicates of the experiment are expected to display at least 75% overlap in the proteins identified per complex. Finally, each complex should contain proteins known to uniquely associate with the respective polymerase (e.g., respective RNAP subunits, general transcription factors such as TFIIH for RNAP II, or the Lupus La antigen for RNAP III), as well as a considerable number of shared protein components (for an overview, see **Fig. 5c**). Finally, Gene Ontology (GO) terms, which concisely describe the contents of each complex, can be used to assess the relative enrichment of factory fractions in terms of 'cytoplasmic', 'nuclear' or 'S100 cytosolic' extracts (**Fig. 5d**).

ACKNOWLEDGMENTS We thank B. Deng and J. Bartlett for their help, and the Sir William Dunn School of Pathology (SWDSOP) proteomics and BioQuant/DKFZ sequencing facilities for the high-throughput analyses. This work was supported by a Medical Research Council grant (to P.R.C.), by the ERASysBio+/FP7 initiative (to K.R. and P.R.C.), by a DKFZ intramural grant (to M.C.-H.) and by Center for Molecular Medicine core funding (to A.P.).

AUTHOR CONTRIBUTIONS S.M. and P.R.C. conceived and developed the factory isolation procedure. S.M., A.P., I.M.C. and P.R.C. implemented and validated the procedure. M.C.-H., L.B., K.R. and A.P. adapted and implemented the protocol for nascent RNA isolation. All authors analyzed data and wrote the manuscript.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

1. Roeder, R.G. The eukaryotic transcriptional machinery: complexities and mechanisms unforeseen. *Nat. Med.* **9**, 1239–1244 (2003).
2. Cramer, P. *et al.* Structure of eukaryotic RNA polymerases. *Annu. Rev. Biophys.* **37**, 337–352 (2008).
3. Das, R. *et al.* SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol. Cell* **26**, 867–881 (2007).
4. Shi, Y. *et al.* Molecular architecture of the human pre-mRNA 3′ processing complex. *Mol. Cell* **33**, 365–376 (2009).
5. Papantonis, A. & Cook, P.R. Transcription factories: genome organization and gene regulation. *Chem. Rev.* **113**, 8683–8705 (2013).
6. Chakalova, L. & Fraser, P. Organization of transcription. *Cold Spring Harb. Perspect. Biol.* **2**, a000729 (2010).
7. Sutherland, H. & Bickmore, W.A. Transcription factories: gene expression in unions? *Nat. Rev. Genet.* **10**, 457–466 (2009).
8. Melnik, S. *et al.* The proteomes of transcription factories containing polymerases I, II, or III. *Nat. Methods* **28**, 963–968 (2011).
9. Caudron-Herger, M., Cook, P.R., Rippe, K. & Papantonis, A. Dissecting the nascent human transcriptome by analyzing the RNA content of transcription factories. *Nucleic Acids Res.* **43**, e95 (2015).
10. Jackson, D.A., Iborra, F.J., Manders, E.M.M. & Cook, P.R. Numbers and organization of RNA polymerases, nascent transcripts and transcription units in HeLa nuclei. *Mol. Biol. Cell* **9**, 1523–1536 (1998).
11. Kimura, H., Tao, Y., Roeder, R.G. & Cook, P.R. Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. *Mol. Cell. Biol.* **19**, 5383–5392 (1999).
12. Jackson, D.A. & Cook, P.R. Transcription occurs at a nucleoskeleton. *EMBO J.* **4**, 919–925 (1985).
13. Ahmad, Y., Boisvert, F.M., Gregor, P., Cobley, A. & Lamond, A.I. NOPdb: Nucleolar Proteome Database—2008 update. *Nucleic Acids Res.* **37**, D181–D184 (2009).
14. Aviner, R., Geiger, T. & Elroy-Stein, O. Novel proteomic approach (PUNCH-P) reveals cell cycle-specific fluctuations in mRNA translation. *Genes Dev.* **27**, 1834–1844 (2013).
15. Core, L.J., Waterfall, J.J. & Lis, J.T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848 (2008).
16. Mayer, A. *et al.* Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell* **161**, 541–554 (2015).
17. Nojima, T. *et al.* Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* **161**, 526–540 (2015).
18. Bhatt, D.M. *et al.* Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell* **150**, 279–290 (2012).
19. Ameur, A. *et al.* Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat. Struct. Mol. Biol.* **18**, 1435–1440 (2011).
20. Khodor, Y.L. *et al.* Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in *Drosophila*. *Genes Dev.* **25**, 2502–2512 (2011).
21. Rabani, M. *et al.* Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nat. Biotechnol.* **29**, 436–442 (2011).
22. Windhager, L. *et al.* Ultrashort and progressive 4sU-tagging reveals key characteristics of RNA processing at nucleotide resolution. *Genome Res.* **22**, 2031–2042 (2012).
23. Kim, T.K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182–187 (2010).
24. Kaikkonen, M.U. *et al.* Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol. Cell* **51**, 310–325 (2013).
25. Mann, M. Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell Biol.* **7**, 952–958 (2006).
26. Cox, J. *et al.* A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* **4**, 698–705 (2009).
27. Besse, S., Vigneron, M., Pichard, E. & Puvion-Dutilleul, F. Synthesis and maturation of viral transcripts in herpes simplex virus type 1 infected HeLa cells: the role of interchromatin granules. *Gene Expr.* **4**, 143–161 (1995).
28. Novakova, Z., Man, P., Novak, P., Hozak, P. & Hodny, Z. Separation of nuclear protein complexes by blue native polyacrylamide gel electrophoresis. *Electrophoresis* **2**, 1277–1287 (2006).
29. Nadano, D., Aoki, C., Yoshinaka, T., Irie, S. & Sato, T.A. Electrophoretic characterization of ribosomal subunits and proteins in apoptosis: specific downregulation of S11 in staurosporine-treated human breast carcinoma cells. *Biochemistry* **40**, 15184–15193 (2001).
30. Raposo, R.A.S., Thomas, B., Ridlova, G. & James, W. Proteomic-based identification of CD4-interacting proteins in human primary macrophages. *PLoS ONE* **6**, e18690 (2011).
31. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
32. Trudgian, D.C. *et al.* CPFP: a central proteomics facilities pipeline. *Bioinformatics* **26**, 1131–1132 (2010).
33. Shteynberg, D. *et al.* iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. *Mol. Cell. Proteomics* **10**, M111.007690 (2011).
34. Lovén, J. *et al.* Revisiting global gene expression analysis. *Cell* **151**, 476–482 (2012).
35. Madsen, J.G. *et al.* iRNA-seq: computational method for genome-wide assessment of acute transcriptional regulation from total RNA-seq data. *Nucleic Acids Res.* **43**, e40 (2015).
36. Gaidatzis, D., Burger, L., Florescu, M. & Stadler, M.B. Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation. *Nat. Biotechnol.* **33**, 722–729 (2015).