ARTICLES
https://doi.org/10.1038/s43587-021-00056-0

nature aging

Check for updates

Aging-associated lncRNAs are evolutionarily conserved and participate in NFκB signaling

Donghong Cai[1,2,3] and Jing-Dong J. Han[1,2] ✉

The transcriptome undergoes global changes during aging, including both protein-coding and noncoding RNAs. Using comparative genomics, we identify aging-associated long noncoding RNAs (lncRNAs) that are under evolutionary constraint and are more conserved than lncRNAs that do not change with age. Aging-associated lncRNAs are enriched for functional elements, including binding sites for RNA-binding proteins and transcription factors, in particular nuclear factor kappa B (NFκB). Using CRISPR screening, we discovered that 13 of the aging-associated lncRNAs were regulators of the NFκB pathway, and we named this family 'NFκB modulating aging-related lncRNAs (NFKBMARLs)'. Further characterization of NFκBMARL-1 reveals it can be traced to 29 Ma before humans and is induced by NFκB during aging, inflammation and senescence. Reciprocally, NFκBMARL-1 directly regulates transcription of the NFκB inhibitor NFKBIZ in *cis* within the same topologically associated domain by binding to the NFKBIZ enhancer and recruiting RELA to the NFKBIZ promoter. These findings reveal many aging-associated lncRNAs are evolutionarily conserved components of the NFκB pathway.

Messenger RNAs with protein-coding ability are highly evolutionarily conserved and strongly constrained by their essential molecular functions[1]. Pervasive transcription also produces many long noncoding RNAs (lncRNAs) in various species that are typically not as conserved[2–4]. Recent studies of global RNA–chromatin interactions have revealed a large number of coding RNAs and lncRNAs prefer to bind active promoters or enhancers[5], and mRNA and lncRNA both play an important role in chromatin remodeling and nuclear architecture. Many lncRNAs regulate transcription as they recruit various regulators, such as transcription factors (TFs), histone modifiers or polymerases, to the targeted gene to activate or repress transcription directly[6,7]. However, the molecular functions of many lncRNAs and how they are regulated are unknown[6]. Cross-species genomic comparison of lncRNAs provides a new strategy to identify the specific activities of lncRNAs, as their functions can be evolutionary constrained even when their sequences diverge[8]. CRISPR can also be used for reliable, high-throughput screening strategies to identify lncRNAs that affect related phenotypes, such as cellular proliferation[9–11].

Aging causes the gradual functional deterioration of whole organisms. Chronic inflammation driven by the NFκB pathway[12,13] likely contributes to the physiological changes of aging, and aging-associated diseases such as neurodegeneration. For example, SIRT6 attenuates RELA binding to the promoters of inflammatory genes and SIRT6-deficient mice have elevated NFκB transcriptional activity and a short life span[14]. In contrast, IκB kinase-β deficiency in mice blunts activation of NFκB, delays aging and extends their life span[15]. Some lncRNAs alter cellular aging, for example, *ANRIL* and *GARDIN* regulate senescence[16]. Whether there are aging-associated lncRNAs that are involved in NFκB signaling is unknown.

We examined the transcriptomes of eleven species from young to older age groups to identify the potential contribution of lncRNAs to the aging process. Surprisingly, we found that aging-associated lncRNAs are more evolutionarily conserved than those expressed independently of age, and are more likely to be functional, as judged by enrichment for cancer survival and disease genome-wide association study (GWAS) loci. Motif enrichment analysis reveals that aging-associated lncRNAs are significantly targeted by the NFκB pathway. CRISPR screening using an NFκB reporter revealed that several of these lncRNAs extensively regulate NFκB signaling in turn. We explored in more detail the molecular mechanisms by which the aging-induced lncRNA *NFKBMARL-1* regulates the NFκB pathway component NFKBIZ to drive a feed-forward loop.

Results

Cross-species comparison of aging-associated lncRNAs reveals signatures of evolutionary constraints. To identify aging-associated lncRNAs, we analyzed public RNA-sequencing (RNA-seq) datasets of eleven species at different ages (human, macaca, green monkey, rat, mouse, pig, wolf, killifish, zebrafish, fly and worm; Fig. 1a and Extended Data Fig. 1a). We constructed new lncRNAs from Ensembl annotation with long transcripts (>200 nucleotides) and removed transcripts with protein-coding potential (Extended Data Fig. 1b). In total, 7,000 to 35,000 lncRNAs were annotated in each species (Fig. 1b). Through syntenic genomic mapping and transcript sequence similarities between species, we annotated the homologous lncRNAs and protein-coding genes (Extended Data Fig. 1c,d,e). Although most lncRNAs are species specific, we identified 9,032 (26%) human lncRNAs as conserved in other species (Extended Data Fig. 2a,b).

To investigate the features of aging-associated lncRNAs, we identified the lncRNAs with significantly differential expression between old and young ages of the worm, zebrafish, rat and pig, and across time-series aging datasets for the wolf, green monkey, fly, macaca, killifish, mouse and human. A much higher proportion of protein-coding genes (7–33%) than lncRNAs (1.7–13%) showed age-dependent expression in each species (Supplementary Table 1). Aging-associated lncRNAs had longer transcripts and more exon numbers (Extended Data Fig. 2c,d), suggesting they may contain

[1]CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China. [2]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Center for Quantitative Biology (CQB), Peking University, Beijing, China. [3]University of Chinese Academy of Sciences, Beijing, China. ✉e-mail: jackie.han@pku.edu.cn

more binding sites for regulators like RNA-binding proteins (RBPs), nucleic acids or splice sites[17].

Evolutionary conservation suggests important functions of lncRNAs. The conservation scores from placental mammal PhastCons20 are higher for the lncRNAs in human brain or blood whose expression is age-associated than those that are unchanged during aging (Extended Data Fig. 2e). Most of the aging-associated protein-coding genes are evolutionarily conserved, with an increased fraction of these found in evolutionarily older gene groups compared to other genes (Extended Data Fig. 2f). Surprisingly, the enrichment of aging-associated lncRNAs in human blood progressively increases with older evolutionary age, all the way to 797 Ma and the beginning of the metazoan evolution. Consistent with a more recent species specification of brain functions, the enrichment of human brain aging-associated lncRNAs goes back to 435 Ma—the beginning of vertebrates (Fig. 1c).

The overrepresentation of antisense transcripts is associated with closer proximity to protein-coding genes (Extended Data Fig. 2g). The intergenic aging-associated lncRNAs have the highest fraction traced to an evolutionary age of 29 Ma of 29.6% in brain and 96 Ma of 26.8% in blood, whereas antisense aging-associated lncRNAs have the highest fraction in 797 Ma (89.5% in brain and 93.9% in blood) compared to other genomic locations (Extended Data Fig. 2h,i).

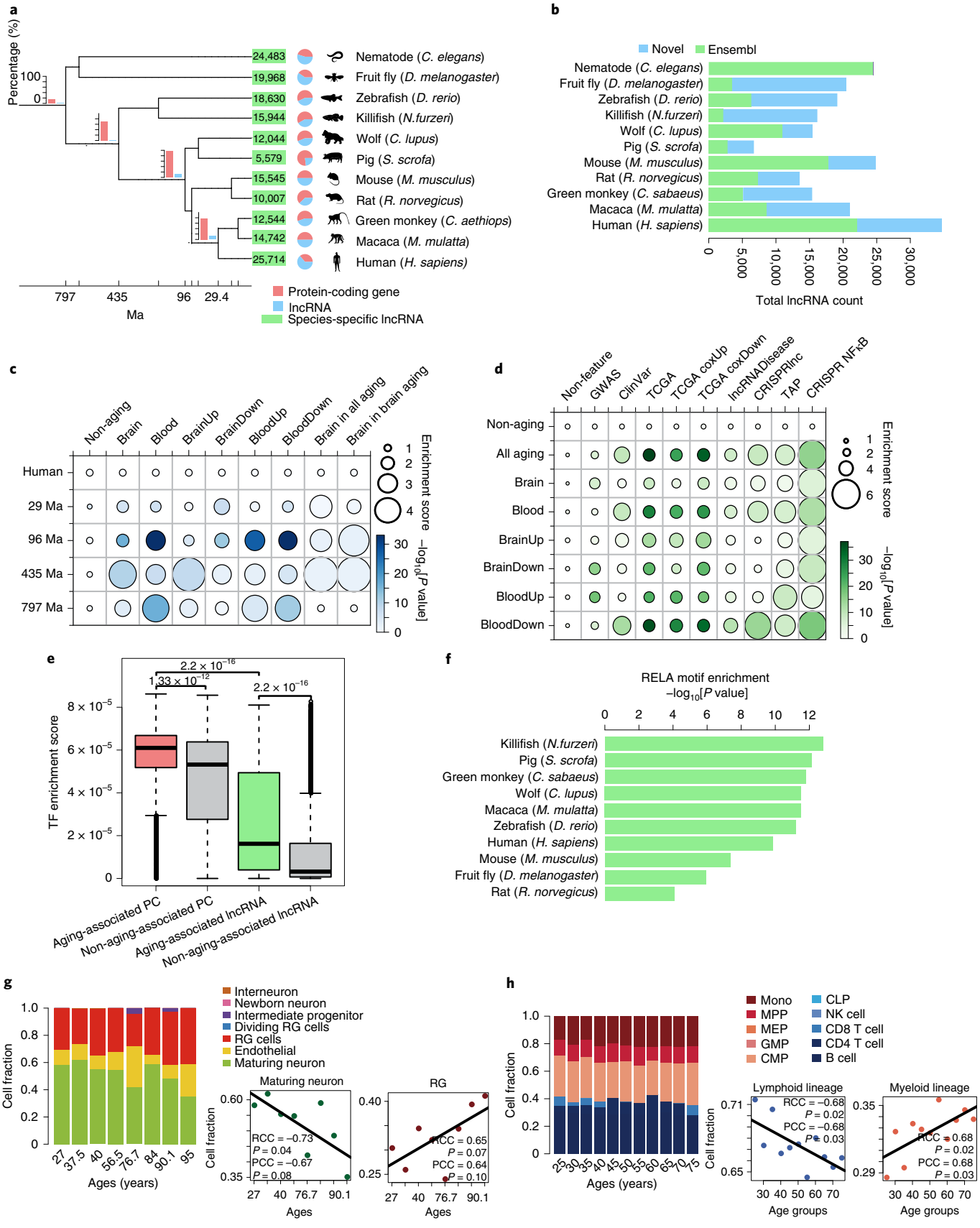
To infer the functional relevance of aging-associated lncRNAs, we curated functional lncRNAs from several databases, including known disease-associated lncRNAs from the lncRNADisease reference database[18]; survival correlated lncRNAs from The Cancer Genome Atlas (TCGA) database; lncRNAs with genetic variants associated with specific traits in the GWAS database; ClinVar[19]; and a set of unbiased CRISPR screen hits from the CRISPRlnc reference database[20]; and topological anchor point (TAP) lncRNAs[21]. Similarly to aging-associated lncRNAs, the lncRNAs with these functional features have older evolutionary ages (Extended Data Fig. 2j). More importantly, aging-associated lncRNAs are more likely to be functional, as judged by enrichment for the above features (Fig. 1d and Extended Data Fig. 2k).

To investigate whether aging-associated genes may be subject to more complex transcriptional regulation, we scanned the TF binding sites on promoters (transcription start site (TSS) ± 250 bp) of human aging-associated genes. While protein-coding genes contained the most TF target sites from the REMAP database, as analyzed by chromatin immunoprecipitation followed by sequencing (ChIP–seq), aging-associated lncRNAs contained more than non-aging-associated lncRNAs, suggesting stronger transcriptional regulation adjusts expression of aging-associated lncRNAs (Fig. 1e). We next estimated the enrichment of TF motifs in the promoters (TSS ± 250 bp) of aging-associated lncRNAs in nine vertebrate species and found 40 significantly enriched TF motifs, including RELA, RUNX1 and BCL6 (Fig. 1f and Extended Data Fig. 2l,m). RBPs are important for RNA maturation[22] and many more RBP binding sites were predicted on aging-associated lncRNAs than non-aging-associated (Extended Data Fig. 3a). Dysregulation of transposon element (TE) expression is associated with the aging process[23,24]. Unexpectedly, the distribution of TEs is increased in lncRNAs of older evolutionary ages (Extended Data Fig. 3b). However, aging-associated lncRNAs contain less TEs than non-aging-associated lncRNAs (Extended Data Fig. 3a). Recently, some inflammation-induced lncRNAs, such as *Umlilo*, were reported to function in *cis* and regulate expression of a neighboring gene in the same topologically associated domain (TAD)[25]. However, whether this is a general phenomenon for many lncRNAs remains unknown. We observed greater correlation between the expression of aging-associated lncRNAs and protein-coding genes in the same TAD than randomly selected lncRNAs. The protein-coding genes that correlated with brain aging-associated lncRNAs in the same TAD were enriched for brain development functions (Extended Data Fig. 3c,d). More aging-associated lncRNAs in the TADs determined by Hi-C technologies were evolutionarily conserved when compared to all aging-associated lncRNAs, which suggests an important connection between aging-associated lncRNAs and TADs (Extended Data Fig. 3e).

The expression of many protein-coding genes changes during aging, which is accompanied by the decline of organ function and increased risk of disease. Interestingly, a greater fraction of lncRNAs (56.7%) are differentially expressed during aging than protein-coding genes (27.9%). In addition, consistent with lncRNAs being highly tissue specific, we found that intergenic lncRNAs exhibit the greatest tissue specificity (Extended Data Fig. 3f,g). These two observations prompted us to investigate whether the cell-type-specific lncRNAs can be used as signatures to deconvolute cell fractions. Indeed, by doing so, we found that during brain aging the fraction of mature neurons decreases and radial glia increase. During blood aging, the fraction of cells of the myeloid lineage increases and the lymphoid lineage cells decrease (Fig. 1g,h), as previously reported by using flow cytometry analysis[26]. This suggests that the tissue specificity of lncRNAs may be exploited to reliably classify tissues and even to deconvolute aging-associated cell-type-composition changes using bulk RNA-seq data.

Fig. 1 | Cross-species comparison of aging-associated lncRNAs shows the signature of evolutionary and functional constraints. a, Phylogenetic tree of species with aging samples used to map one-to-one orthologous lncRNA families (branches) and species-specific lncRNA (leaves). The pie chart shows the fraction of protein-coding genes and lncRNAs. The bar plot shows the percentage of human protein-coding genes and lncRNAs conserved to different evolutionary ages. **b**, Numbers of new and annotated lncRNAs in 11 species. **c**, Enrichment of aging-associated lncRNAs in different evolutionary ages, including brain aging dependent, blood aging dependent, brain aging up or down and blood aging up or down. Brain in 'all aging' is the enrichment of human brain aging-associated lncRNAs in all species aging-associated lncRNAs. Brain in 'brain aging' is the enrichment of human brain aging-associated lncRNAs in all species brain aging-associated lncRNAs. A total of 2,000 randomly selected non-aging-associated lncRNAs were used as controls for comparison. The enrichment analysis was done by two-sided Fisher's exact test. **d**, Enrichment scores of aging-associated lncRNAs for different functional feature groups by two-sided Fisher's exact test. A total of 2,000 randomly selected non-aging-associated lncRNAs served as controls. TAP, topological anchor point. **e**, Enrichment score of TF binding sites per gene on the promoters of protein-coding (PC) genes, aging-associated lncRNAs and non-aging-associated lncRNAs. *P* values were calculated by two-sided Wilcoxon rank-sum test with correction. The center line indicates the median; box limits are the interquartile range of 25th and 75th percentiles; whiskers each extend to 1.5 times the interquartile range of maxima and minima; and dots are outliers. **f**, Enrichment of RELA motif (MA0107) among aging-associated lncRNAs in nine vertebrates, with the dl motif (MA0022, the orthologous motif of Rel homology region factors in insect) shown for fly. **g**, Deconvoluted cell fraction distribution in aging human brain RNA-seq data by cell-type-specific lncRNAs as feature signature matrix. Dot plots represent the dynamic changes in cell fractions of maturing neurons and radial glial (RG) cells during aging. The Spearman's RCC and Pearson's correlation coefficient (PCC) of cell fraction to age are labeled. **h**, Deconvoluted cell fraction distribution in aging human blood RNA-seq data by using cell-type-specific lncRNAs as a feature signature matrix. Dot plots depict the dynamic changes of cell fractions of distribution in lymphoid and myeloid lineages during aging, showing the RCC and PCC of cell fraction to age. MPP, multipotent progenitor; MEP, megakaryocytic erythroid progenitor; GMP, granulocyte monocyte progenitor; CMP, common myeloid progenitor; CLP, common lymphoid progenitor.

Aging-associated lncRNAs regulate, and are regulated by, the NFκB pathway. To assign the putative functions of aging-associated lncRNAs, we applied the guilt-by-association method based on their coexpression with protein-coding genes. The lncRNAs upregulated in brain aging were enriched for immune-related pathways, such as NFκB pathway and cytokine receptor interaction, whereas



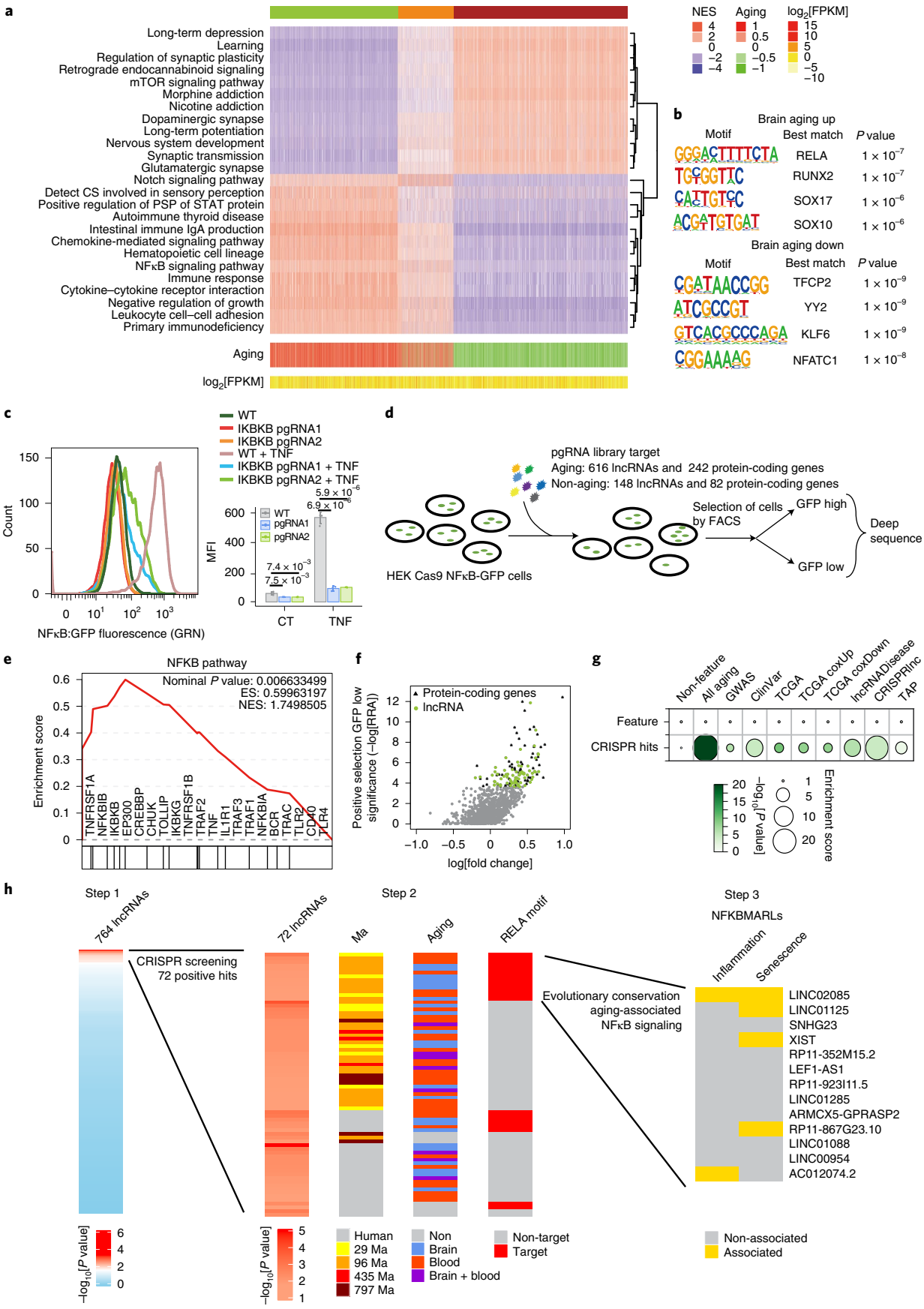
the downregulated lncRNAs were enriched for neuron and synaptic functions (Fig. 2a). The upregulated and downregulated aging-associated lncRNAs in the blood were enriched for the NFκB pathway and mitochondrial function, respectively (Extended Data Fig. 4b). The predicted functions of aging-associated lncRNAs are highly concordant with experimentally validated aging-associated protein-coding genes[12]. We suspected that aging-associated lncRNAs are regulated by similar pathways as protein-coding genes. TF motif analysis showed the lncRNAs upregulated in brain and blood aging were enriched for the RELA motif (Fig. 2b and Extended Data Fig. 4c). The enrichment of NFKB1 and RELA binding in the TSS of lncRNAs upregulated in aging was also revealed by the REMAP ChIP–seq data (Extended Data Fig. 4a). Therefore, RELA is an important factor activated by aging-related inflammation and is predicted to bind the promoters of both aging-associated lncRNAs and protein-coding genes.

To find out whether aging-associated lncRNAs play a role in regulating the NFκB pathway, we performed NFκB reporter CRISPR screening using paired-guide RNAs (pgRNAs) to delete aging-associated lncRNAs in a HEK293T cell line stably expressing Cas9 and a green fluorescent protein (GFP)-tagged NFκB response element (HEK Cas9 NFκB-GFP cells) and validated it by IKBKB knockout (KO), and NFKB1 and RELA knockdown (KD) under serial concentrations of tumor necrosis factor (TNF) induction (Fig. 2c, Extended Data Fig. 4d,e and Methods). We constructed a pgRNA library targeting 764 lncRNAs, including 616 aging-associated lncRNAs and 148 non-aging-associated lncRNAs (Fig. 2d). HEK Cas9 NFκB-GFP reporter cells were transduced with pgRNA library and cultured for 7 d, treated with TNF, and the cells with high and low GFP expression were sorted and sequenced (Extended Data Fig. 4f,g). Unsupervised hierarchical clustering on the normalized reads could indeed group the samples with high and low GFP expression in three independent replicates into two separate clusters (Extended Data Fig. 4h). Also, the negative-control pgRNAs targeting AAVS loci and nontarget control loci showed small variance (Extended Data Fig. 4i,j). We then applied the MAGeCK algorithm to identify the CRISPR screening hits enriched in samples with low GFP expression (Methods). Gene-set enrichment analysis (GSEA) showed that positive-control pgRNAs were enriched in CRISPR screening hits, confirming the effectiveness of the screen in recovering known NFκB pathway regulators (Fig. 2e). The CRISPR screening hits ($P<0.05$) contained 64 protein-coding genes and 72 lncRNAs (Fig. 2f). To validate the CRISPR screening result, we individually transfected the pgRNAs targeting 18 of the lncRNAs from CRISPR screening hits and the AAVS1 loci as a negative control in the reporter cells. All 18 lncRNAs reduced the fluorescence intensity of the NFκB reporter to various degrees (Extended Data Fig. 4k). One of the lncRNAs in the CRISPR screening hits, nuclear paraspeckle assembly transcript 1 (*NEAT1*) KD has previously been reported to repress NFκB pathway[27]. Within the 72 lncRNAs in the CRISPR screening hits, 65 were aging associated and significantly enriched, indicating that aging-associated lncRNAs preferentially regulate the NFκB pathway (Fig. 2g). These lncRNAs were also strongly enriched in features of functional lncRNAs, indicating that lncRNAs that regulate the NFκB pathway are also important in physical health (Fig. 2g). In contrast, the 64 protein-coding genes were not enriched for aging-associated protein-coding genes (Extended Data Fig. 4l). Surprisingly, both protein-coding genes and lncRNAs identified as CRISPR screening hits were evolutionarily conserved (Fig. 2g and Extended Data Fig. 4l). We annotated 13 of the 72 CRISPR screening hits as NFKBMARLs based on their evolutionary conservation, aging-associated expression and the presence of the RELA motif in the TSS region (Fig. 2h).

Upregulation of *NFKBMARL-1* by NFκB in aging, inflammation and senescence. Among the 13 NFKBMARLs, only LINC02085 is both responsive to TNF treatment in monocytes and displays senescence-related expression change. Thus, we selected LINC02085 for further downstream mechanistic study and renamed it *NFKBMARL-1* to denote its functional origin (Fig. 2h). From the available sequencing datasets, the transcript region of *NFKBMARL-1* was consistent with the reference genome annotation (Extended Data Fig. 5a). Based on the annotation and our FirstChoice RLM-RACE result for the 5′ and 3′ end of *NFKBMARL-1*, we identified its length as 1,633 nucleotides (Extended Data Fig. 5b). We further estimated the copy number of *NFKBMARL-1* in IMR90 as 32.4 (Extended Data Fig. 5c,d). *NFKBMARL-1* has a conserved lncRNA in the macaca genome and conserved transcript sequence in exon 4 (Extended Data Fig. 5e). Since the evolutionary age of *NFKBMARL-1* is 29 Ma (Supplementary Table 1), *NFKBMARL-1* belongs to a relatively young evolutionary age and is a primate-specific lncRNA, and is not only upregulated in primate tissue aging, but also during human cell inflammation and senescence (Fig. 2h), thus we studied the function of *NFKBMARL-1* in human cell inflammation and cellular senescence. *NFKBMARL-1*

Fig. 2 | Aging-associated lncRNAs regulate, and are regulated by, the NFκB pathway. a, Heat map showing the GSEA of normalized enrichment scores (NESs) from guilt-by-association analysis to annotate the function of brain aging-associated lncRNAs. Each column is a lncRNA and each row is a Gene Ontology (GO) term. The bottom two rows are expression correlations to age and expression levels of brain aging-associated lncRNAs. CS, chemical stimulus; PSP, peptidyl serine phosphorylation. **b**, TF motif enrichments in the TSS of brain aging up or down lncRNAs were determined by Homer compared to an equal number of randomly selected genomic regions as background. **c**, The GFP fluorescence intensity of NFκB pathway reporter cells after knocking out IKBKB by two pgRNAs in the TNF stimulation versus control (CT), as measured by fluorescence-activated cell sorting (FACS). The mean fluorescence intensity (MFI) was analyzed by FlowJo. Data are presented as the mean ± standard deviation (s.d.) from $n=3$ independent experiments. All *P* values were calculated using a two-sided unpaired Student's *t*-test. GRN, green. **d**, Workflow of CRISPR screen for NFκB pathway modulators. The HEK NFKB-GFP Cas9 cells were infected by lentivirus containing packaged pgRNAs and selected by stable expression of RFP in the same construct. The cell fractions with GFP upper quartile fluorescence (GFP high) and the GFP lower quartile fluorescence (GFP low) were isolated by FACS. Then genomic DNA was extracted from isolated cell populations to amplify the barcoded gRNA, followed by deep sequencing. **e**, With the canonical NFκB pathway genes spiked in as positive controls, GSEA identified NFκB pathway enrichment in sorted GFP low genes (CRISPR positive hits) of the CRISPR screen. **f**, The robust rank aggregation (RRA) scores of GFP significantly low lncRNAs and protein-coding genes ($P<0.05$), defined by the MAGeCK algorithm. **g**, The enrichment analysis of sorted GFP low lncRNAs (CRISPR hits) in aging-associated lncRNAs and different features of functional lncRNAs. TAP, topological anchor point. **h**, Functional annotation of CRISPR positive hits from CRISPR screening in NFκB signaling. Step 1: 764 highly expressed (reads per kilobase million (RPKM) > 0.5) lncRNAs (616 aging-associated and 148 non-aging-associated lncRNAs) were selected for CRISPR screening of NFκB pathway modulators. Step 2: 72 positive hits ($P<0.05$) were found by CRISPR screening (column 1), defined by MAGeCK. Among the 72 positive hits, we further filtered by evolution conservation (that is, at least conserved in primates; column 2), either brain or blood aging-related expression change (column 3), and containing the RELA motif in TSS ± 250-bp region (column 4). In total, 13 lncRNAs satisfied these criteria that we termed NFKBMARLs. Step 3: the 13 NFKBMARLs were further filtered by both TNF response (expression $\log_2$[fold change] ($|\log_2[FC]|>1$ in monocytes; GSE100382) and cellular senescence-related expression change ($|\log_2[FC]|>1$ in IMR90 senescence; GSE74328).

is upregulated during human blood aging based on our previously published RNA-seq data[28] (Spearman's rank correlation coefficient (RCC) = 0.332, $P = 3.161 \times 10^{-8}$, Fig. 3a and Extended Data Fig. 5f). Using published RNA-seq data of CD14^{+}CD16^{-} monocytes from 20 young and 20 older male participants[29], we observed significant upregulation of *NFKBMARL-1* in aging monocytes (Extended



Data Fig. 5g). Similarly, significant upregulation of *NFKBMARL-1* occurred in seven tissues in GTEx samples, such as lung and breast, with no significant changes in other tissues during aging (Extended Data Fig. 5h).

NFKBMARL-1 is about 100 kb upstream of *NFKBIZ*, which encodes IkBζ, an upstream regulator of the NFκB pathway. NFKBIZ is reported as an important activator for immune cell activation and senescence-associated secretory phenotype (SASP) release of the senescent cells[30–33]. Consistent with previous reports, NFKBIZ KD in 293T cells reduced the expression of *NFKBMARL-1*, CXCL1, interleukin (IL)-6 and IL-8 under TNF stimulation (Extended Data Fig. 5i). In contrast, NFKBIZ overexpression induced expression of CXCL1, IL-6 and IL-8, but not *NFKBMARL-1* (Extended Data Fig. 5j). We deleted NFKBIZ by using the CRISPR–Cas9 system with two pgRNAs and selected single clones (Extended Data Fig. 5k). In the NFKBIZ KO cells, the NFκB pathway was blocked under TNF stimulation, even simultaneous TNF, lipopolysaccharide (LPS) and interferon-α (IFNα) stimulation (Extended Data Fig. 5l,m).

NFKBMARL-1 and NFKBIZ were positively correlated in aging peripheral blood mononuclear cells (PBMCs; RCC = 0.286, $P = 2.307 \times 10^{-6}$; Fig. 3a). Therefore, we asked whether upregulation of *NFKBMARL-1* during blood aging is due to NFκB activation and increased production of cytokines during aging, such as TNF[34–36]. *NFKBMARL-1* and NFKBIZ were both upregulated in THP1 cells upon TNF, LPS and IFNα co-treatment, and positively correlated in a published RNA-seq dataset of $CD14^+$ monocytes[37] (Fig. 3b and Extended Data Fig. 5n). Moreover, *NFKBMARL-1* and NFKBIZ were concertedly upregulated in various cell lines, including 293T, IMR90 and neutrophils[38] during inflammation (Fig. 3c–e and Extended Data Fig. 5o).

Aging also causes cellular senescence and the release of inflammatory cytokines as part of the SASP. We observed increased expression of the senescent marker gene *CDKN2A* (p16-INK4A) during human blood and brain aging (Extended Data Fig. 6a). In IMR90 cells, expression of both *NFKBMARL-1* and NFKBIZ is activated by several types of senescence inducers: etoposide and bleomycin, which cause DNA damage-induced senescence, oncogene-induced senescence caused by RAS overexpression[39], and replicative senescence (both early and late senescence; Extended Data Fig. 6b–e). To verify these changes in other cell lines, we treated 293T cells with bleomycin, and found that *NFKBMARL-1* and NFKBIZ could also be activated (Extended Data Fig. 6b). In conclusion, *NFKBMARL-1* expression is activated during aging, senescence and inflammation, which are all regulated by the NFκB pathway.

RELA is a core subunit of NFκB, and binds the promoters of *NFKBMARL-1* and NFKBIZ upon TNF, LPS, Pam2CSK4, PolyIC or MtriDAP treatment of $CD14^+$ monocytes, adipocytes, IMR90 cells, human umbilical vein endothelial cells (HUVECs) and D562 cells as revealed by several published datasets[37,40–43] (Fig. 3f and Extended Data Fig. 7a–d). Other NFκB subunits, such as RelB, c-Rel, NFKB1 and NFKB2, simultaneously bind to the NFKBIZ and *NFKBMARL-1* promoters in GM12878 (Extended Data Fig. 7e). In senescent IMR90 cells, RELA KD globally inhibits senescence-activated genes including *NFKBMARL-1* and NFKBIZ, suggesting that RELA co-regulates NFKBIZ and *NFKBMARL-1* expression (Extended Data Fig. 6f).

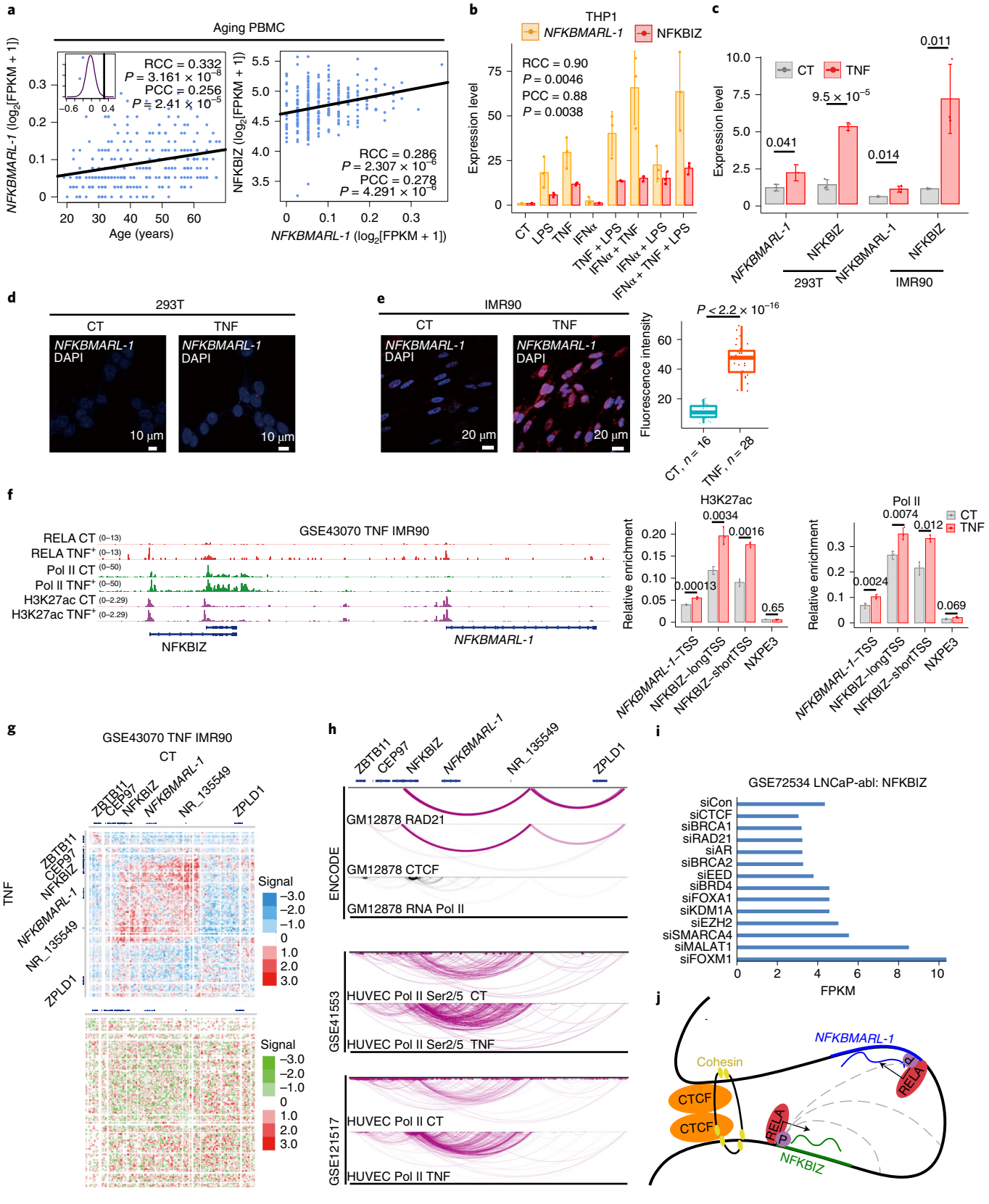
At the NFKBIZ and *NFKBMARL-1* promoters, greater chromatin accessibility after TNF (Extended Data Fig. 7b) treatment and higher histone acetylation were observed upon both inflammation and senescence (Fig. 3f and Extended Data Fig. 6g,h). The enhanced histone acetylation of the *NFKBMARL-1* and NFKBIZ regions generated a new super-enhancer in senescent cells, which is absent in proliferating cells (Extended Data Fig. 6i). Inhibition of the histone acetylation reader BRD4 by JQ1 or short hairpin RNA (shRNA) can repress global expression levels including *NFKBMARL-1* and NFKBIZ (Extended Data Fig. 6j)[44]. Although *NFKBMARL-1* and NFKBIZ are located in the same TAD, we did not observe significant changes of TAD structure and intensity after TNF treatment, which suggests there are preexisting DNA loops between the promoters of *NFKBMARL-1* and NFKBIZ before signal-dependent RELA binding to the promoters (Fig. 3g). A similar phenomenon was also observed in the senescent process (Extended Data Fig. 6k). Within this TAD, preexisting DNA loops between *NFKBMARL-1* and NFKBIZ could be captured by chromatin interaction analysis using paired-end tags (ChIA-PET) of Pol II, and the signal significantly increased under TNF treatment (Fig. 3h and Extended Data Fig. 6l). Breaking the anchor of TAD by CTCF or RAD21 KD downregulated expression of NFKBIZ, suggesting the integrity of the TAD is important in the expression of NFKBIZ (Fig. 3i,j).

To identify the subcellular location of *NFKBMARL-1*, we performed RNA single-molecule fluorescence in situ hybridization (smFISH). *NFKBMARL-1* was detected in one focal location in the nucleus of 293T cells under TNF treatment (Fig. 3d). In TNF-stimulated or bleomycin-induced senescent IMR90 cells, *NFKBMARL-1* had a much stronger signal in more pervasive foci within the nucleus, suggesting that *NFKBMARL-1* has a higher expression level in IMR90 cells and mainly localizes in the nucleus (Fig. 3e and Extended Data Fig. 6e). Quantitative PCR with reverse transcription (RT–qPCR) of subcellular fractions also showed *NFKBMARL-1* in the nucleus with the largest fraction associated with chromatin (Extended Data Fig. 7f).

***NFKBMARL-1* regulates NFKBIZ in *cis*.** The CRISPR screen identified *NFKBMARL-1* as an activator of NFκB during TNF treatment, so we asked whether *NFKBMARL-1* mediates NFκB activation directly by regulating NFKBIZ in *cis*. We deleted

Fig. 3 | *NFKBMARL-1* is activated during aging, inflammation and senescence by the NFκB pathway. a, Left: dot plot shows the expression of *NFKBMARL-1* in 280 aging PBMC samples. Right: dot plot shows the expression levels of *NFKBMARL-1* and NFKBIZ. The RCC and PCC between ages and expression of *NFKBMARL-1* or expression of NFKBIZ and *NFKBMARL-1* and their *P* values are shown. Top left: distribution of Spearman's RCC of all genes to ages in blood samples. The gray line is the correlation of *NFKBMARL-1* to ages. **b**, Expression of *NFKBMARL-1* and NFKBIZ after TNF (10 ng ml^{-1}), LPS (10 ng ml^{-1}) and IFNα (25 ng ml^{-1}) treatment for 3 h in THP1 cells, as revealed by RT–qPCR. **c**, Expression of *NFKBMARL-1* and NFKBIZ in TNF (10 ng ml^{-1}) treated in 293T and IMR90 cells. **d**, Distinct localization of *NFKBMARL-1* in TNF-stimulated 293T cells by smFISH. **e**, Distinct localization of *NFKBMARL-1* in TNF-stimulated IMR90 cells by smFISH (left: smFISH imaging; right: quantification of *NFKBMARL-1* fluorescence intensity in each cell). The center line indicates the median value; box limits are the interquartile range of 25th and 75th percentiles; whiskers each extend to 1.5 times the interquartile range of maxima and minima; dots represent fluorescence intensity. The results were processed by ImageJ. **f**, Left: genome browser view showing the binding profiles of RELA, Pol II and the epigenetic status of H3K27ac in TNF-stimulated IMR90 cells. Right: ChIP–qPCR shows increased H3K27ac and Pol II occupancy on the promoter region of *NFKBMARL-1*, NFKBIZ after TNF treatment in IMR90. NXPE3 is used as a negative control. **g**, Top: genomic contact matrices of control and TNF-stimulated IMR90 cells in chromosome 3 (chr3: 101625407–102547170). Bottom: subtraction contact matrices from TNF versus control. **h**, ChIA-PET of RAD21, CTCF and Pol II showing the PETs between NFKBIZ and *NFKBMARL-1* in GM12878. Pol II ChIA-PET showing the PETs between NFKBIZ and *NFKBMARL-1* under TNF treatment in HUVECs. **i**, The expression of NFKBIZ in gene regulators KD RNA-seq of LNCaP-abl cells. **j**, Schematic of the genomic structure between NFKBIZ and *NFKBMARL-1*. P, Pol II; CTCF, CCCTC-binding factor; dashed lines, ChIA-PET signals. Data in **b**,**c** and **f** are presented as the mean ± s.d. from $n = 3$ independent experiments. All *P* values in **b**,**c**,**e** and **f** were calculated using a two-sided unpaired Student's *t*-test.

NFKBMARL-1 with the CRISPR–Cas9 system and selected single clones with homozygous *NFKBMARL-1* KO (Extended Data Fig. 8a). Under TNF treatment, NFKBIZ was significantly repressed in *NFKBMARL-1* KO 293T cells (Fig. 4a), and the fluorescence intensity of the NFκB reporter was reduced (Fig. 4b). THP1 cells are used as a model to study chronic inflammation in aging[45,46], and the TNF-induced genes, including *NFKBIZ*, showed significant overlap between THP1 and CD14^{+} monocytes (Fisher's exact test



$P < 2.2 \times 10^{-16}$, Extended Data Fig. 5n), suggesting the NFκB pathway is activated in THP1 similar to the in vivo activation pattern. We further validated the regulatory function of *NFKBMARL-1* by generating heterozygous and homozygous KO THP1 clones (Extended Data Fig. 8b). The expression level of NFKBIZ was stepwise reduced from TNF-stimulated wild-type (WT) to heterozygous and to homozygous KO THP1 cells (Fig. 4c). The protein levels of NFKBIZ and IL-1β were also decreased in *NFKBMARL-1* KO compared to WT THP1 cells (Fig. 4d).

We next investigated whether the chromatin region of *NFKBMARL-1* functions as an enhancer. Using an enhancer reporter luciferase assay, we found that the genomic region corresponding to the promoter or gene body region of *NFKBMARL-1* failed to activate a luciferase reporter under TNF stimulation (Fig. 4e). We further knocked in the EF1 promoter into the TSS of *NFKBMARL-1* to overexpress *NFKBMARL-1* in *cis* (Extended Data Fig. 8c). In cells overexpressing *NFKBMARL-1* from its endogenous genomic locus, NFKBIZ was simultaneously induced (Fig. 4f). However, when we overexpressed an exogenous *NFKBMARL-1* in *trans*, NFKBIZ expression was not induced (Fig. 4g). These findings suggest that the transcript of *NFKBMARL-1* may function to activate NFKBIZ in *cis*.

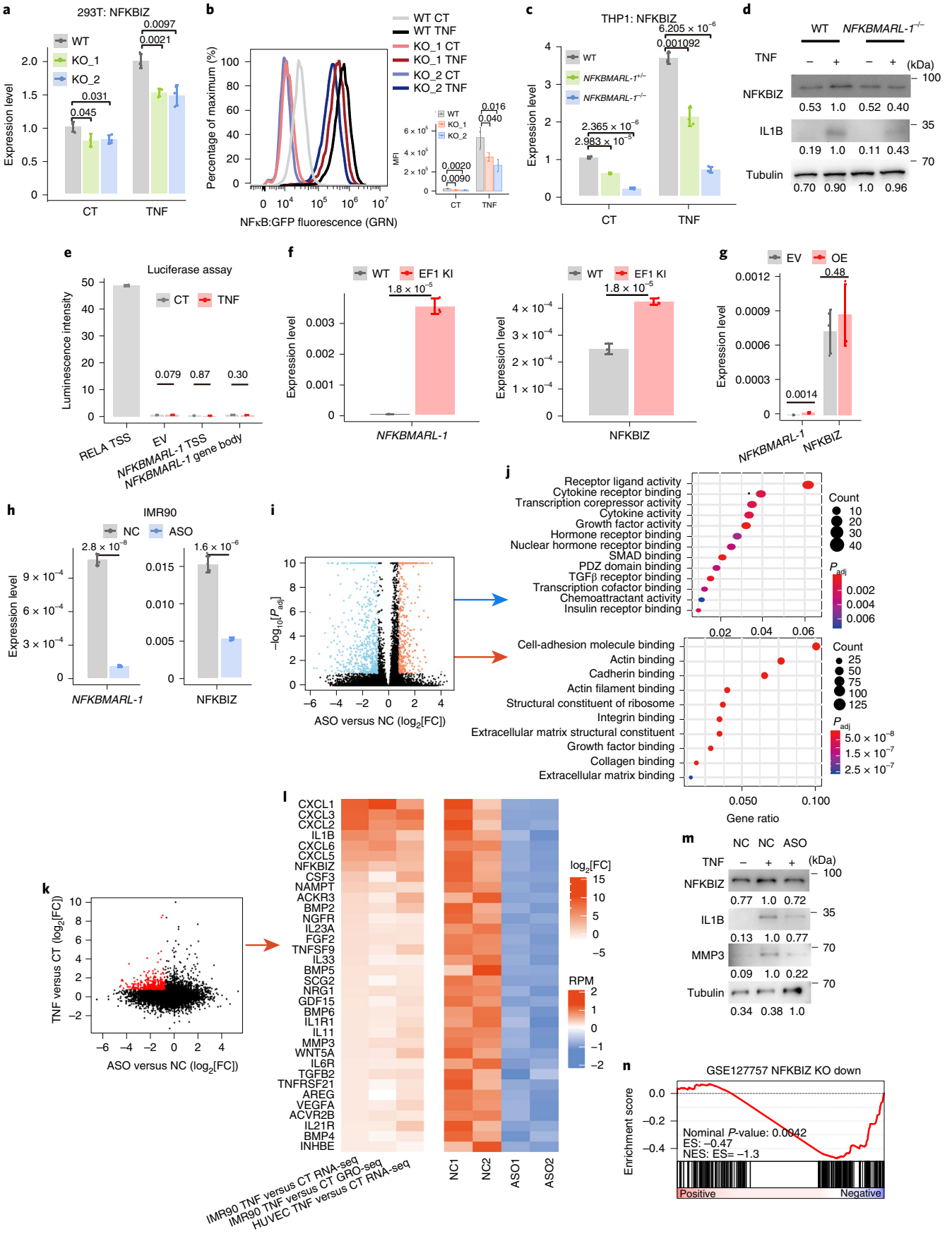
After *NFKBMARL-1* KD and TNF stimulation of IMR90 cells, NFKBIZ expression was downregulated, as shown by RT–qPCR, ribo-strand-specific RNA-seq, western blot and immunofluorescence (IF; Fig. 4h–m and Extended Data Fig. 8d). By RNA-seq, we observed that other downregulated genes were enriched for cytokine activity and upregulated genes were enriched for cell-adhesion-related genes (Fig. 4i–l). Then, we selected the intersection of genes significantly activated in TNF-stimulated IMR90 cells and repressed gene expression by *NFKBMARL-1* KD; this, again, enriched for cytokine genes, such as *IL1B* and *MMP3*, which were validated by western blot (Fig. 4k–m). NFKBIZ KO downregulated genes were highly similar to *NFKBMARL-1* KD downregulated genes, as determined by GSEA[47] (Fig. 4n). In *NFKBMARL-1* KD and bleomycin-induced senescent IMR90 cells, the expression levels of NFKBIZ and cytokine-related genes, including *CXCL1*, *IL6* and *IL8*, were also repressed. The repression of IL1B was validated by western blot and IF (Extended Data Fig. 8e–h). However, the senescence marker gene p16-INK4A did not change and neither did the senescence-associated heterochromatic foci observed by DAPI and 5-bromodeoxyuridine intensity (Extended Data Fig. 8f,g). Therefore, the senescent process is not altered by *NFKBMARL-1* KD.

In the monocyte/macrophage differentiation process, *NFKBMARL-1* and NFKBIZ are both activated (Extended Data Fig. 8i). To examine the effect of *NFKBMARL-1* on monocyte/macrophage differentiation, we differentiated *NFKBMARL-1* KO THP1 cells into M0 and M1 macrophages by 12-*O*-tetradecanoylphorbol-13-acetate (PMA) and LPS treatment, respectively. The expression levels of NFKBIZ, CXCL1, IL8 and CD14 were inhibited in *NFKBMARL-1* KO M1 macrophages (Extended Data Fig. 8i). The inhibition of IL1B was also validated by western blot (Extended Data Fig. 8j). Flow cytometry also validated the reduction of CD14-positive cells (Extended Data Fig. 8k). As M1 macrophages can release the IL-1β to repress tumor growth, we co-cultured *NFKBMARL-1* KO M1 macrophages with Hela cells and found that *NFKBMARL-1* KO cells allowed more tumor colonies to grow (Extended Data Fig. 8l). Therefore, *NFKBMARL-1* may regulate macrophage differentiation and cytokine activation.

***NFKBMARL-1* recruits RELA to the promoter of NFKBIZ.** To investigate the mechanism by which *NFKBMARL-1* regulates NFKBIZ transcription, we performed ChIP–qPCR of RELA bound to the NFKBIZ promoter in *NFKBMARL-1* KO THP1 and *NFKBMARL-1* KD IMR90 cells. In *NFKBMARL-1* KO or KD cells, less RELA bound the NFKBIZ promoter upon TNF stimulation (Fig. 5a,b). Biotin-RNA pull-down assay revealed that RELA, HNRNPA2B1 and HNRNPU can bind to *NFKBMARL-1* (Fig. 5c and Extended Data Fig. 9a). *NFKBMARL-1* binding to RELA was validated by tRSA-RNA pull-down assay (Extended Data Fig. 9a). In RNA immunoprecipitation (RIP) assays with anti-RELA antibody, *NFKBMARL-1* was more strongly enriched in TNF-stimulated or bleomycin-induced senescent IMR90 cells compared to untreated controls (Fig. 5d and Extended Data Fig. 9b). An electrophoretic mobility shift assay further revealed a direct interaction between purified RELA from *Escherichia coli* and *NFKBMARL-1* (Extended Data Fig. 9c). Next, we asked which domain of RELA can bind to *NFKBMARL-1*. RELA has a DNA-binding domain called the Rel homology domain (RHD) in the N terminus and activation domain (AD) in the C terminus. RELA-N containing RHD can bind to *NFKBMARL-1* directly (Extended Data Fig. 9c). We predicted the RNA-binding residues of RHD using Pprint and KYG software scores[48,49]. Two regions were identified as RNA-binding residues: one site from amino acid 30 to 45, and the other from amino acid 198 to 202 (Fig. 5e). Then we generated mutated recombinant RELA-N-R1 and RELA-N-R2 for the in vitro binding assay with *NFKBMARL-1*. Both of the RELA-N mutations weakened their binding affinity with *NFKBMARL-1* (Fig. 5f). Next, we artificially split *NFKBMARL-1* into four fragments for the RNA pull-down assay, and identified fragment 4 (F4) as having the strongest binding ability to RELA, also validated by electrophoretic mobility shift assay and in silico simulated docking (Fig. 5g and Extended Data Fig. 9d,e). Deletion of the genomic region of *NFKBMARL-1* F4 by the CRISPR–Cas9 system weakened the activation of NFKBIZ in TNF- or bleomycin-treated IMR90 cells (Fig. 5m and Extended Data Fig. 9f).

Fig. 4 | *NFKBMARL-1* regulates NFKBIZ expression under TNF stimulation in *cis*. **a**, Expression level of NFKBIZ in WT and *NFKBMARL-1* KO TNF-stimulated 293T cells, revealed by RT–qPCR. **b**, Fluorescence intensity of NFκB pathway reporter in WT and *NFKBMARL-1* KO TNF-stimulated 293T cells, revealed by FACS. The MFI was analyzed by FlowJo. **c**, Expression level of NFKBIZ in WT and *NFKBMARL-1* KO TNF-stimulated THP1 cells by RT–qPCR. **d**, Western blot of NFKBIZ and IL1B in WT and *NFKBMARL-1* KO THP1 cells. **e**, Luciferase assay revealed the enhancer activity of TSS and gene body region of *NFKBMARL-1*. EV, empty vector. **f**, Expression level of *NFKBMARL-1* and NFKBIZ in the EF1 knock-in (KI) 293T cells, revealed by RT–qPCR. **g**, Expression levels of *NFKBMARL-1* and NFKBIZ were quantified by RT–qPCR after in-*trans* overexpression (OE) of *NFKBMARL-1* in TNF-stimulated 293T cells. **h**, Expression levels of *NFKBMARL-1* and NFKBIZ in control and *NFKBMARL-1* KD TNF-stimulated IMR90 cells, revealed by RT–qPCR and using ACTIN as internal control. NC, negative control. **i**, The volcano plot of genes regulated by *NFKBMARL-1* KD in TNF-stimulated IMR90 cells, revealed in ribo-strand-specific RNA-seq and analyzed by HTSeq and DESeq. The blue dots are downregulated genes with a $\log_2$[FC] < −0.8 and an adjusted *P* value (P_{adj}) < 0.1, and the red dots are upregulated genes with a $\log_2$[FC] > 0.8 and P_{adj} < 0.1. **j**, GO annotations of significantly downregulated (top) or upregulated (bottom) genes in TNF-stimulated *NFKBMARL-1* KD IMR90 cells (P_{adj} < 0.1). Differentially expressed genes and their adjustment were defined by DESeq (**i** and **j**). **k**, Dot plot of $\log_2$[FC] in TNF-stimulated WT IMR90 cells and *NFKBMARL-1* KD IMR90 cells. The red dots annotated the gene with a $\log_2$[FC] > 0.8 in TNF-stimulated IMR90 cells and $\log_2$[FC] < −0.8 in *NFKBMARL-1* KD. **l**, Heat map of $\log_2$[FC] of cytokine-related genes by TNF stimulation in IMR90 cells measured by RNA-seq and global run-on sequencing and in HUVECs by RNA-seq, and their expression levels in control and *NFKBMARL-1* antisense oligonucleotides (ASO) TNF-stimulated IMR90 cells was measured by RNA-seq. **m**, Western blot for NFKBIZ, IL1B and MMP3 in *NFKBMARL-1* KD control and TNF-stimulated IMR90 cells. **n**, GSEA of NFKBIZ KO downregulated genes with signatures computed by the fold change of *NFKBMARL-1* KD. Normalized enrichment score was adjusted by GSEA. Data in **a**–**c** and **e**–**h** are presented as the mean ± s.d. from *n* = 3 independent experiments. All *P* values in **a**–**c** and **e**–**h** were calculated using a two-sided unpaired Student's *t*-test.

Chromatin isolation by RNA purification (ChIRP) showed 236 genomic DNA-binding peaks for *NFKBMARL-1* in TNF-stimulated IMR90 cells. Among them, 117 peaks were located in the intergenic region (Fig. 5h). About 74% of the intergenic ChIRP peaks were located in the Pol II ChIA-PET anchors (Fig. 5h). Hence, we speculated that *NFKBMARL-1* might regulate gene expression



through genomic interaction. The genome browser view shows that *NFKBMARL-1* binding to the 3′ end ChIRP peak of NFKBIZ was induced by TNF and senescence in IMR90 cells (Fig. 5i). As this ChIRP peak is a nucleosome-free region, we performed an RNA–DNA triplex-formation experiment and found a fragment in *NFKBMARL-1* (F2) that had the strongest binding to this peak (Fig. 5j), which was also predicted by LongTarget software[50] (Fig. 5k). Deleting the binding anchor of *NFKBMARL-1* on the NFKBIZ 3′ end ChIRP peak weakened the induction of NFKBIZ in TNF-stimulated or bleomycin-induced senescent IMR90 cells (Fig. 5l and Extended Data Fig. 9f). Chromosome conformation capture coupled with quantitative PCR (3C–qPCR) revealed a strong pre-existing enhancer–promoter interaction between the 3′ end ChIRP peak and the promoter of NFKBIZ in IMR90 cells that was unaffected by *NFKBMARL-1* KD (Fig. 5m and Extended Data Fig. 9g). These findings suggest that *NFKBMARL-1* F2 binds on the enhancer of NFKBIZ and F4 recruits RELA to the promoter of NFKBIZ to activate gene expression in inflammation and senescence (Fig. 6j).

***NFKBMARL-1* participates in RELA phase separation.** Activation of NFKBIZ by TNF or Kdo2-lipid A (KLA) treatment can be blocked by 1,6-hexanediol in MCF7 cells[51] (Fig. 6a). We also validated the inhibition of NFKBIZ by 1,6-hexanediol in TNF-stimulated IMR90 (Fig. 6b). This indicates that these TFs likely undergo phase separation on the promoter of NFKBIZ during NFKB activation and inflammation, and this may regulate activation of *NFKBIZ*. Consistent with this hypothesis, RELA contains intrinsically disordered regions (IDRs), with a high alanine and proline content (Extended Data Fig. 10a). RELA and MED1 IF indeed showed nuclear puncta in TNF-stimulated 293T cells (Extended Data Fig. 10b). Endogenous RELA tagged with EGFP in live 293T cells also revealed discrete nuclear puncta (Extended Data Fig. 10c). We next investigated whether RELA foci possess the features of liquid-like condensates. Internal dynamical reorganization and rapid exchange kinetics in liquid-like condensates can be detected by measuring the rate of fluorescence recovery after photobleaching (FRAP). To study the dynamics of RELA foci in live 293T cells, we performed FRAP in cells with tagged EGFP-RELA. After photobleaching, EGFP-RELA foci recovered after 150 s (Extended Data Fig. 10d). To investigate whether RELA can form droplets in vitro, we purified recombinant RELA-EGFP fusion proteins and MED1-RFP as a positive control. RELA-EGFP and MED1-RFP solutions formed GFP-positive micron-sized spherical droplets (Fig. 6c). To examine the dynamic reorganization of these droplets, we photobleached them and observed recovery after 200 s (Fig. 6d). As TFs can form phase-separated condensates with mediators[52], such as MED1, we mixed RELA-EGFP or EGFP with MED1-RFP, which showed that RELA-EGFP incorporated into MED1-RFP droplets and enlarged the foci size (Extended Data Fig. 10e).

To detect whether RELA participates in phase separation on the chromatin region of NFKBIZ, we performed DNA FISH of NFKBIZ with IF for RELA and found significant colocalization of RELA foci on the chromatin region of NFKBIZ in TNF-stimulated IMR90 cells (Fig. 6e). More importantly, we observed the *NFKBMARL-1* transcripts by smFISH signal colocalizing with RELA or MED1 foci (Extended Data Fig. 10f). We mixed Cy3-labeled *NFKBMARL-1* and RELA-EGFP, and found that Cy3-*NFKBMARL-1* was packaged into the droplet of RELA-EGFP and enlarged the size of the droplets (Fig. 6f). RELA-EGFP droplets are salt sensitive and were reduced in size by higher NaCl concentrations (Fig. 6g). Next, we performed mass spectrometry to investigate whether the protein families pulled down by *NFKBMARL-1* have IDRs that participate in phase separation (Extended Data Fig. 10g). Proteins pulled down by *NFKBMARL-1* were enriched with IDRs (Fig. 6h), including members of the heterogeneous nuclear ribonucleoproteins (HNRNPs) and ribosomal protein family (RPL; Fig. 6i). We validated HNRNPU and HNRNPA2B1 from the HNRNP family by RNA pull-down assay and RIP experiments in IMR90 cells (Fig. 5c and Extended Data Fig. 10h). At least some of these HNRNPs and other RBPs were also required for the full activation of NFKBIZ expression and NFκB pathway response (Extended Data Fig. 10i–n). Based on this evidence, we constructed a regulatory model in which *NFKBMARL-1* binds to the enhancer region of NFKBIZ by F2 and recruits RELA by F4 to the promoters of NFKBIZ through genomic interactions during inflammation and participates in the phase separation of RELA (Fig. 6j).

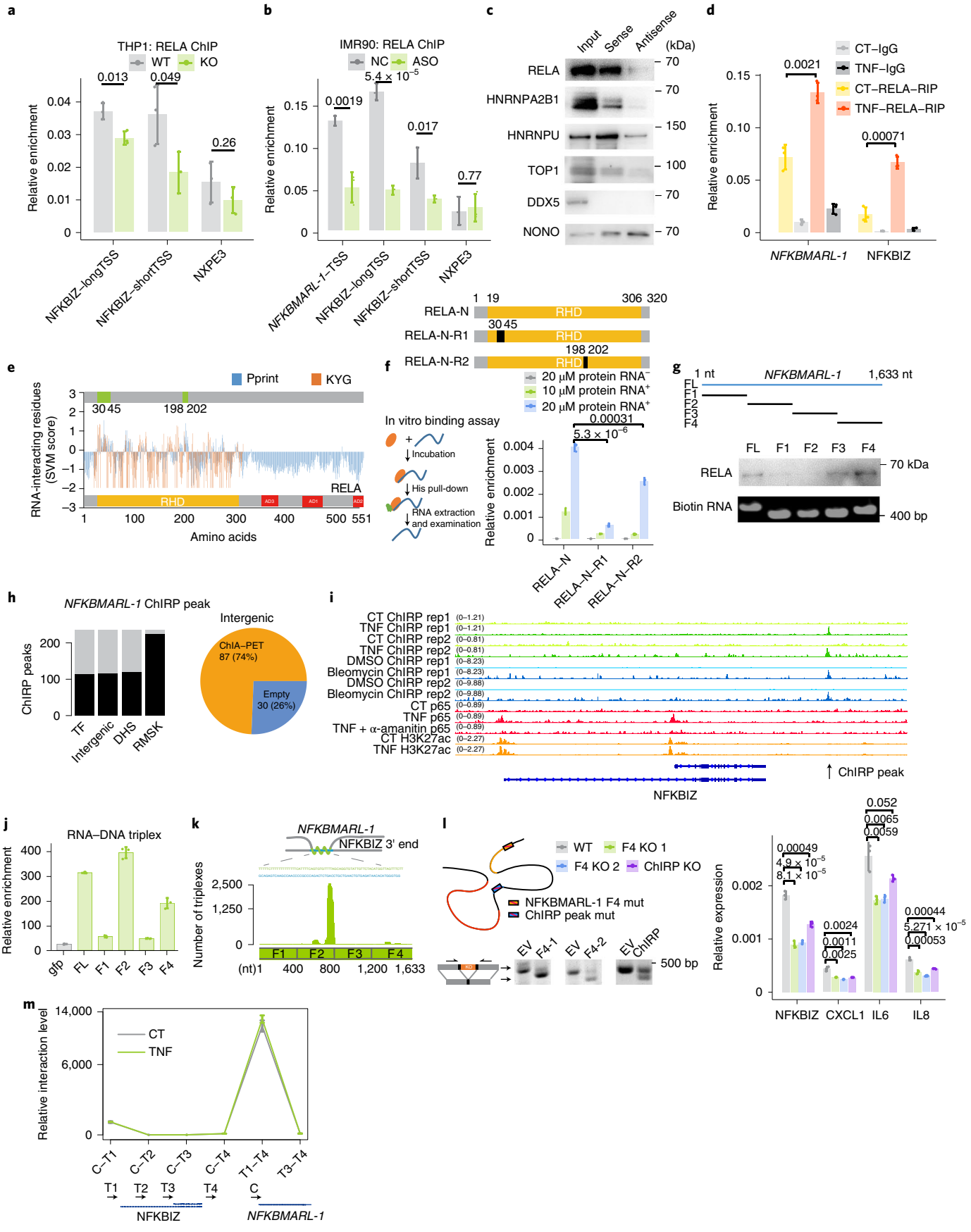
Discussion

Overall, our study provides a comprehensive analysis of aging-associated lncRNAs and their evolutionary, regulatory and functional properties. We identify 13 lncRNAs as NFKBMARLs, and demonstrate that *NFKBMARL-1* is both activated by NFκB and controls subsequent NFκB-driven gene expression by regulating the transcription of NFKBIZ.

Protein-coding genes that regulate important biological processes are strongly constrained evolutionarily[1]. lncRNAs that are dynamically regulated during development are also reported

Fig. 5 | NFKBMARL-1 recruits RELA to NFKBIZ by binding on the 3′ end enhancer of NFKBIZ. a, Enrichment of RELA binding on the promoter of NFKBIZ in WT and *NFKBMARL-1* KO TNF-stimulated THP1 cells, as shown by ChIP–qPCR. NXPE3 acted as the negative control. **b**, Enrichment of RELA binding on the promoter region of NFKBIZ in control and *NFKBMARL-1* KD TNF-stimulated IMR90 cells, as shown by ChIP–qPCR. NXPE3 acted as the negative control. **c**, The *NFKBMARL-1* RNA pull-down assay for RELA, HNRNPU, HNRNPA2B1, TOP1, DDX5 and ILF3, as revealed by western blot. The positive control was input of total cell lysis and the negative control was the antisense RNA. **d**, Enrichment of *NFKBMARL-1* and NFKBIZ in native anti-p65 or IgG RIP assays in TNF-stimulated IMR90 cells, quantified by RT–qPCR. **e**, Score of RNA-binding residue in RELA predicted by Pprint and KYG software using support vector machine (SVM) with position-specific scoring matrix profiles. **f**, Top: schematic of RELA-N and RELA-N truncation lacking predicted RNA-binding region R1 or R2. Bottom: *NFKBMARL-1* enrichment level by in vitro binding assay, revealed by RT–qPCR. **g**, RNA pull-down assay by four biotin-labeled RNA fragments of *NFKBMARL-1* in TNF-stimulated IMR90 followed by western blot. **h**, Distribution of chromosome regions isolated by biotin-labeled antisense probes targeting *NFKBMARL-1* in TNF-stimulated IMR90 cells. Left: pie chart showing the genomic distribution of 236 ChIRP peaks in TF, intergenic, DNase I hypersensitive site (DHS) and RepeatMasker (RMSK file). Right, distribution of 117 intergenic ChIRP peaks in Pol II ChIA-PET of TNF-stimulated HUVECs. **i**, Genome browser view of ChIRP-seq binding profile of *NFKBMARL-1* in the 3′ end ChIRP peak of NFKBIZ in TNF- and bleomycin-induced IMR90 cells, and RELA binding profile and histone modification of H3K27ac in TNF-stimulated IMR90. The arrow highlights the 3′ end ChIRP peak of NFKBIZ. **j**, In vitro triplex-formation assay for four fragments of *NFKBMARL-1* with the NFKBIZ 3′ end ChIRP peak, revealed by qPCR. **k**, Distribution of triplex formation between ChIRP peaks and RNA sequence of *NFKBMARL-1* predicted by LongTarget, based on Hoogsteen base-pairing rules to align the triplex-formation oligonucleotides (TFO) and triplex target sites (TTS). Top: triplex-formation sequence of *NFKBMARL-1* (green) and NFKBIZ 3′ end ChIRR peak (blue). **l**, Left: schematic of mutated *NFKBMARL-1* fragment 4 and ChIRP peak region. PCR-validated mutated region. Right: after mutated F4 or ChIRP peak region, IMR90 cells were activated by TNF and quantified for expression of NFKBIZ, CXCL1, IL6 and IL8 for RT–qPCR. **m**, Genomic interaction from *NFKBMARL-1* and ChIRP binding regions to the long and short promoter regions of NFKBIZ in TNF-stimulated IMR90 cells, revealed by 3C–qPCR. The upper arrows denote the locus of the primers designed for the 3C experiment. Data in **a**,**b**,**d**,**f**,**j**,**l** and **m** are presented as means ± s.d. from $n = 3$ independent experiments. All P values in **a**,**b**,**d**,**f**,**j**,**l** and **m** were calculated using two-sided, unpaired Student's t-tests.

to have strong conservation[2]. There is some evidence that life span regulation is due to relaxed selection[53]. Here we found that aging-associated lncRNAs are more conserved and have older evolutionary ages than non-aging-associated lncRNAs, and are more likely to be functional, as judged by enrichment for cancer survival and disease GWAS loci. Aging-associated lncRNAs



accumulate more regulatory binding sites for TFs and RBPs, and are closer to protein-coding genes within the same TAD than non-aging-associated lncRNAs.

Intriguingly, we found that lncRNAs upregulated during aging are under NFκB pathway regulation, whereas CRISPR screening for NFκB pathway regulators were significantly enriched

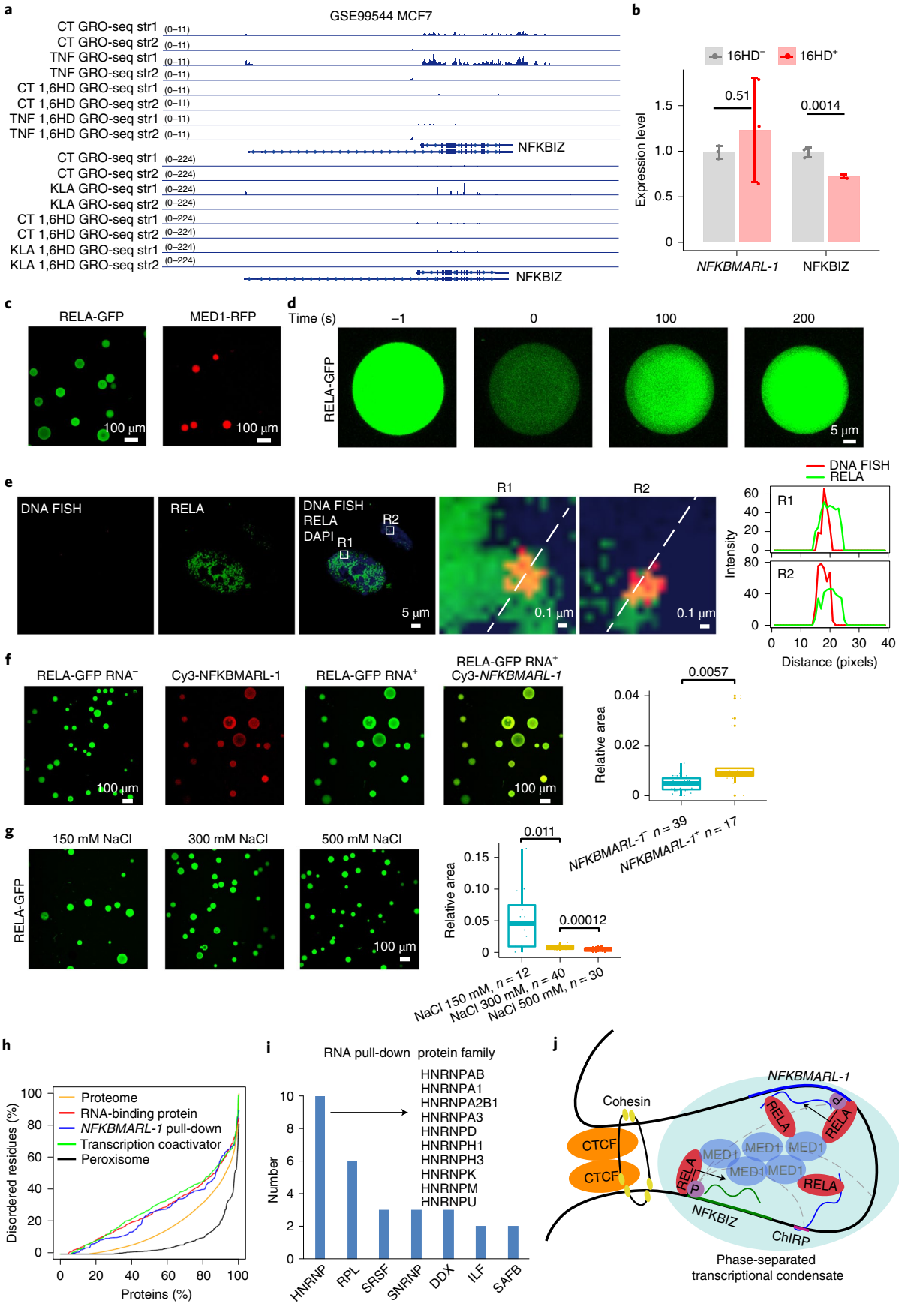


Fig. 6 | *NFKBMARL-1* regulates phase separation of the NFKBIZ chromatin region. a, Genome browser view showing the expression profile of NFKBIZ in 8% 1,6-hexanediol (1,6HD; 5 min) follow by 100 nM TNF or KLA (30 min) treatment in MCF7 cells. **b**, Expression of *NFKBMARL-1* and NFKBIZ with or without 8% 1,6-hexanediol (5 min) treatment follow by 10 ng ml^{-1} TNF (30 min) treatment in IMR90 cells, quantified by RT–qPCR. **c**, Droplet foci formation in 10 μM purified recombinant RELA-GFP and MED1-RFP. **d**, Images of FRAP experiments were performed on RELA-GFP. **e**, Left: image of DNA FISH of the chromatin in NFKBIZ and IF of RELA in TNF-stimulated IMR90 cells. Middle: two enlarged selected regions of DNA FISH into R1 and R2. Right: quantification of the fluorescence signal represented by the dashed lines in the images. **f**, Left: images of droplet foci formation for RELA-GFP with cy3-labeled *NFKBMARL-1*. Right: quantification of the relative area of RELA-GFP. **g**, Left: the images of droplet foci formation of RELA-GFP in 150, 300 and 500 mM NaCl droplet formation buffer. Right: the relative area of RELA-GFP. The center line indicates the median value; box limits are the interquartile range of 25th and 75th percentiles; whiskers each extend to 1.5 times the interquartile range of maxima and minima; and dots represent the relative area (**f** and **g**). **h**, The percentage of disordered residues of each protein (*y* axis) against the cumulative fraction of a different subset of proteins (*x* axis). **i**, Statistics of protein families pulled down by *NFKBMARL-1*. **j**, Model of *NFKBMARL-1* regulates NFKBIZ in *cis*. P, Pol II; dashed lines, ChIA-PET signals. All image results in **e**–**g** were processed by ImageJ. Data in **b** are presented as the mean ± s.d. from $n = 3$ independent experiments. All *P* values in **b**,**f** and **g** were calculated using two-sided, unpaired Student's *t*-tests.

for aging-associated lncRNAs, suggestive of a positive-feedback relationship between NFκB and lncRNAs in the processes of inflammation, aging and senescence. These CRISPR hits also show an increased fraction in older evolutionary ages. In fact, in all the species analyzed here that have evolved the NFκB component RELA (from fly and above), aging-associated lncRNAs are highly enriched for the RELA targeting motif (Fig. 1f), suggesting NFκB may be a key regulator responsible for the lncRNA expression changes observed during aging. Likewise, given our CRISPR screening results, we envision many aging-associated lncRNAs in human and other species may also be involved in the positive-feedback regulation of NFκB signaling. This is exemplified by the molecular mechanism of *NFKBMARL-1*, which we found to regulate NFKBIZ in *cis* by binding to the 3′ end of NFKBIZ and directly bind and recruit RELA to chromatin. Blocking *NFKBMARL-1* expression blocks the toxic SASPs without affecting the senescence progression, representing an appealing mild alternative to senolytics. We expect many more aging-associated lncRNAs may be uncovered as regulators of inflammation, aging and senescence in the future, highlighting them as a previously overlooked key player driving the vicious cycle of 'inflammaging' and underlying the etiology of many aging-associated diseases.

Methods

Homologous lncRNA identification through one-to-one orthology assignment. To identify the homologous lncRNAs across different species, we applied two criteria: syntenic genomic mapping and transcript sequence alignment. Identification of the syntenic region between each species was performed accordingly[2]. For the syntenic genomic mapping, the pairwise genome alignments between 11 species, a total of C_{11}^2 combinations, were aligned to each other and were used to construct a projector. Specifically, we split genomic sequences soft-masked with RepeatMasker by chromosomes and aligned each chromosome between two species using BLASTZ with the parameters --notransition --step = 20 --nogapped --format = maf --chain and constructed the chain of the genomic coordinates between the two species. For each lncRNA, we projected the genomic coordinate from one species to the other species using the liftOver tool via the chain file with a minimal ratio of bases that must remap parameter set to 0.1 and allowing multiple results as the output. Then we applied intersectBed to obtain the intersection of lncRNAs around 10 kb. After obtaining the syntenic lncRNA, we further applied criterion 2 to obtain transcript sequences with conserved lncRNA. For each intersected lncRNA, we obtained the transcript sequence by twoBitToFa from FASTA files, and aligned each pair by BLASTZ with default parameters to obtain the alignment scores. Protein-coding, small nucleolar RNAs and micro RNA genes were constrained to search within the same category; the remaining lncRNAs were mapped to each other and the lncRNA with the highest alignment score was defined as the evolutionarily conserved lncRNA.

De novo assembly of transcribed regions and identification of lncRNAs. We collected transcriptomic datasets for the aging process in brain, blood and liver samples across 11 species (human, macaca, mouse, rat, pig, wolf, killer fish, zebrafish, green monkey, *Caenorhabditis elegans* and *Drosophila*), containing ribo-strand-specific and poly(A)-selected RNA-seq libraries. Reads were mapped to the genome by STAR (020201) and assembled using Cufflinks (v2.2.1). The qualified transcripts were further filtered for multi-exonic or single-exon transcripts with fragments per kilobase of transcript per million mapped reads (FPKM) > 5 and length > 200 bp. The transcripts from same species were merged using the cuffmerge function from the Cufflinks package. We further detected the constructed transcripts' protein-coding potential by CPAT (v1.2.2) and Pfamscan (v3.0). CPAT calculates the protein-coding score through an alignment-independent logistic regression model trained from the protein-coding genes and lncRNAs in the same species. Then the noncoding transcripts with a coding score < 0.44 were kept. To increase the power to remove the protein-coding transcripts, we further applied Pfamscan to align the annotated coding sequence from the PFAM database to the assembled transcripts. The transcripts identified with putative coding peptides were further removed from the lncRNA list.

Gene expression quantification and aging-associated lncRNA identification. Gene expression levels were quantified by Cufflinks (v2.2.1) as FPKM values. The two-sample dataset with young and old ages was used with a two-sided *t*-test to obtain significant aging-associated genes for worm, zebrafish, rat and pig. The age distribution for these four species was around 0.5–0.6 and the fold changes for age distribution were scaled. We calculated the Spearman's RCC for time-series aging samples of wolf, green monkey, fly, macaca, killifish, human and mouse to obtain significant aging-associated genes. The lncRNAs with a cutoff of $P < 0.05$ were defined as aging-associated lncRNAs. At a type I error rate < 0.05 ($P < 0.05$), the powers of all aging datasets analyzed here were larger than or close to 0.5, corresponding to a relatively low type II error. For example, a *P* value < 0.05 for the human brain dataset ($n = 8$) corresponds to an RCC > 0.7 and power > 0.6, while $P < 0.05$ for the human blood dataset ($n = 280$) corresponds to an RCC > 0.11 and power > 0.5.

The TF, RBP and repeat distribution analysis on lncRNAs. For the TF enrichment analysis in a gene set, the promoter region was defined as 250 bp around the TSS. We obtained the TF binding sites from the REMAP database with 131 nonredundant TFs from thousands of ChIP–seq experiments and mapped the peak regions. We overlapped the TF peaks with genes and quantified enrichment using two-sided Fisher's exact test. The RBP binding sites were retrieved from the POSTAR2 database with 171 human RBPs based on cross-linked immunoprecipitation followed by sequencing. Then, the complexity of RNA processing was defined as the percentage of the whole-gene interval overlapping with RBP binding signals. Repeat coverage was similarly defined as the percentage of the gene interval overlapping with the eight types of DNA sequence for interspersed repeat and low-complexity DNA sequences annotated by the RepeatMasker database (enrichment score = binding sites/length × 1,000,000).

Coexpression analysis for aging-associated lncRNAs in TAD regions. The raw Hi-C reads were processed by Homer (v4.11.1) to call TAD regions from the ENCODE Hi-C dataset (GSE87112). The closest protein-coding genes to human brain aging-associated lncRNAs were identified in the same TAD regions. PCC was calculated for the gene pairs within the same TAD. For control gene pairs, we selected a set of 10,000 randomly assigned lncRNA and protein-coding gene pairs. The biological functions of protein-coding genes immediately upstream and downstream of the brain aging-associated lncRNAs in same TAD were annotated by DAVID for GO annotation.

Tissue specificity score calculation. Jensen–Shannon specificity (JSS) and divergence (JSD) were used to measure the similarity and divergence, respectively, between two probability distributions. To measure the specific score for the gene in

a tissue, a binary list **b**, with the tissue assigned as 1 and the remaining samples as 0. The expression of gene list **g**, is defined by $G=g_i/||G||$, i = 1, 2… n, where n is the number of samples.

$$\mathrm{JSS}(G||B) = 1 - \mathrm{JSD}(G||B)$$

$$\mathrm{JSD}(G||B) = \frac{1}{2}\left(D(G||M) + D(B||M)\right), \text{where } M = \frac{1}{2}(G+B)$$

$$D(G||M) = \sum_{i=1}^{n} g_i \log\frac{g_i}{m_i}, D(B||M) = \sum_{i=1}^{n} b_i \log\frac{b_i}{m_i},$$
$$\text{where } m_i = \frac{1}{2}(g_i + b_i)$$

Motif enrichment. The 'findmotif' function in Homer software was used to analyze the enrichment of TF motifs in the promoter regions of selected lncRNAs (250 bp surrounding the TSS). The known motifs were annotated using the Homer and JASPAR databases.

Construction of CRISPR screen system. *Reporter cell line.* The 293T cells were infected by lentivirus packaged Cas9 and selected by puromycin. Reporter Cas9 293T cells were generated by lentivirus delivery of reporter plasmid in Cas9 293T cells and followed by sorting out single-cell clones stably expressing GFP. The reporter cell line was validated by quantifying the NFκB-GFP response element under serial concentrations of TNF to activate the NFκB pathway (Extended Data Fig. 4d). Then we validated the NFκB reporter cells by using shRNAs to knock down NFKB1 and RELA or pgRNAs to knock out IKBKB. From the FACS results, NFKB1 and RELA KD weakened GFP strength after induction by TNF (Extended Data Fig. 4e). The pgRNAs knock out of IKBKB completely blocked TNF activation of the NFκB pathway (Fig. 2c).

Library design. For targeting lncRNAs, approximately 616 aging-associated and 148 non-aging-associated lncRNAs were selected to design up to ten pgRNAs for each targeted lncRNA. For targeting protein-coding genes, 242 aging-associated and 82 non-aging-associated protein-coding genes were selected to design pgRNAs. Positive-control genes were 20 genes of the NFκB signaling pathway, including *CHUK*, *IKBKB*, *TNF*, *NFKBIA*, *NFKBIB*, *IKBKG*, *TRAF1*, *TRAF2*, *TRAF3*, *CD40*, *TOLLIP*, *IL1R1*, *TLR4*, *TLR2*, *TNFRSF1A*, *TNFRSF1B*, *TRAC*, *BCR*, *CBP* and *EP300*. Among the 20 positive selection genes, 60 pgRNAs target promoters (the distance between two pgRNAs is between 500 bp and 5 kb), and the 140 pgRNAs were designed to target the promoter and gene body with the same orientations with gene, as this has a better knockout effect than the opposite direction of gRNAs. For negative controls, we designed 100 nontarget genomic pgRNAs and 100 pgRNAs targeting the *AAVS1* region, which is a nonessential region in the genome and frequently used in CRISPR studies for efficiency tests.

pgRNA design. For all the targeted lncRNAs and positive-control genes, we selected the targeted regions across the promoter and gene body region. For the promoter region, we chose the region upstream of each TSS as the targeted region. After selecting the region, we identified all the potential single-guide RNAs (sgRNAs) by scanning the PAM motif in both sense and antisense genome sequences and the spacing between sgRNAs of at least 20 nucleotides. Then, we removed all multiple mapping and selected the sgRNA with at least two mismatches with the other region by using BWA to map the 20-nucleotide spacer to the reference genome (hg38). After selecting sgRNAs with the least probability of causing off-target effects, we calculated the efficiency score. The sgRNA GC content was between 20% and 80%. We also removed the sgRNAs that had more than four continuous bases or a BsmBI restriction enzyme cutting site. For designing pgRNAs, we first selected the sgRNAs with the highest efficiency and enumerated the paired sgRNAs, with one ahead of the TSS and the other one after the TSS in the same orientation. When designing pgRNAs for lncRNAs, we removed the sgRNA overlap with exon and 5 kb ahead of the TSS of protein-coding genes. For selection of essential genes and designing a pgRNA efficiency score predictor, we first generated the read counts for all samples and merged them into the matrix from the public pgRNA screening dataset[11]. Then, we normalized the data by the 'median ratio normalization' function and estimated significant pgRNAs by the negative binomial distribution. For estimating the essential lncRNAs, MAGeCK was applied using the RRA algorithm for 20 pgRNAs to target each lncRNA to calculate the RRA score. Then, we selected the significant lncRNAs as essential lncRNAs. After identifying the 121 essential lncRNAs in both Huh75 and Hela cell lines targeted by 2,420 pgRNAs, we next filtered out the efficient pgRNAs and inefficient pgRNAs. In both Huh75 and Hela cells, we set the threshold of log-ratio to less than −0.5 for identification of efficient pgRNAs and more than 0.5 for identification of inefficient pgRNAs. Next, we extracted the sequence feature of these pgRNAs spacers. To measure the sequence preference, we calculated the log odds ratio of nucleotide frequency between efficient pgRNAs and inefficient pgRNAs. Then, we indexed the nucleotide frequency and used it as the gold standard for machine learning. We trained the pgRNA-predictor using the support vector machine in linear kernel and used it to predict the efficiency of pgRNA. The highly efficient pgRNAs were selected into the library.

pgRNA library construction. Oligonucleotide pools were designed with flanking cloning and PCR sites, synthesized by CustomArray, and the oligonucleotides amplified into double-strand DNA with overlapping regions to the plasmid. The amplification product was cloned into the plasmid by Gibson assembly. Ligation reactions were pooled by ten independent replicates (pool 1–10). Then, we amplified the plasmid by transformation into highly efficient competent *E. coli* cells according to the manufacturer's protocol. Following incubation overnight at 37 °C, we counted the single clone after 16 h and ensured the library had high complexity with individual pgRNAs and more than 100× coverage. The plasmid DNAs were collected by using two replicates with QIAGEN Midi Prep Kit. The assembled plasmids were digested by endonuclease again. The ligations used digested plasmid and the linker contained tracrRNA-U6 with ten independent replicates by using T4 DNA ligase (NEB). The ligated plasmid was transformed into the competent cell and collected with at least 1,000× coverage for individual pgRNAs. The CRISPR library was validated by sequencing the PCR product of the barcode gRNA. From the sequencing results, the sequencing depth was 24 million reads, and the mapping ratio was 84.6%. Around 99.9% of pgRNAs were successfully constructed into the library.

Package virus and FACS selection. sgRNA libraries were packaged into lentiviral particles using HEK293T cells. For each pooled package, 5 million cells were transfected 24 h after plating 100 µl Fugene HD reagent (Promega) diluted in 6 ml of OPTI-MEM combined with 20 µg sgRNA library, 15 µg Δ8.9 package plasmid and 9 µg VSVG envelope plasmid. Then the medium was replaced by 1% BSA DMEM. At 48 and 72 h after transfection, the lentivirus was collected and filtered through a 0.45-µm membrane frozen at −80 °C. The lentivirus was tested for the multiplicity of infection (MOI) in 293T cells to ensure it was less than 0.5 before the screening. Experiments were performed in biological triplicates from the infection step. The pgRNA library was delivered into 1.2×10^7 reporter cells by lentivirus infection with 10 µg ml^{-1} polybrene. Infected cells were collected by FACS for red fluorescence 3 d after infection.

FACS-based screening. The pgRNA library was transduced at an MOI of 0.5, aiming to cover an average of 1,000 cells per pgRNA reagent. MOI was measured by the infection rate of RFP-positive cells in 12-point dose-response ranging from 0 to 400 µl of viral supernatants in 10 µg ml^{-1} polybrene. For the genome-wide screen, 60 million cells were seeded at 20 million cells per 15-cm dish (Corning) in the triplicate experiment. At 24 h after plating, the culture medium was replaced with fresh medium containing 0.5 MOI of lentivirus and 10 µg ml^{-1} polybrene. At 24 h infection, the culture was replaced with fresh medium. After 48 h, the infected cells were selected with RFP-positive cells by FACS. Cells were maintained in culture and split as needed to ensure the confluence did not exceed 90%. The pgRNA libraries were screened 7 d after infection. After stimulation with 10 ng ml^{-1} TNF, the cells were selected as GFP high and GFP low by FACS, for which cells were collected, resuspended at 10 million cells per ml, and RFP-positive cells were sorted (BD FACSAria II) for lower GFP and higher GFP. For the genome-wide screen, 50 million cells were collected for sorting out GFP high and GFP low. Approximately 50 million cells were collected as the input sample.

Illumina library construction. The genomic DNA of every replicate from 4×10^6 live cells was isolated using QIAamp DNA mini kit (QIAGEN). Then, 0.6 µg of genome DNA was used for amplification. The genomic DNA was subjected to two-step PCR. For the first step, the amount of input genomic DNA was more than 200× the coverage of the designed gRNA library (assuming 10^6 cells had 6.6 µg of genomic DNA), which was calculated as 20 µg for genomic DNA PCR. For each sample, at least five independent PCR reactions were done, after which all the amplicons were pooled together. gRNAs were amplified by PCR (NEB Q5 hot-start DNA polymerase) with 28 cycles of reaction using primers targeting the U6 promoter and linker between two guide RNAs of each pair. Then the PCR products were second amplified with the Illumina adaptor and index sequenced. The PCR products were purified with QIAquick PCR purification kit (QIAGEN), followed by Illumina deep sequencing. The RRA algorithm was used to identify the statistically significant pgRNA abundance changes using the negative binomial model, and the genes targeted by these pgRNAs were identified by the RRA algorithm.

Cell culture. IMR90 cells were obtained from the American Type Culture Collection (CCL-186), and 293T and THP1 cells were from Cell Bank, Chinese Academy of Sciences. The 293T cells were maintained in DMEM supplemented with 10% FBS and 1% penicillin–streptomycin. THP1 cells were maintained in RPMI1640 supplemented with 10% FBS, 1× nonessential amino acids solution and 1% penicillin–streptomycin. IMR90 cells were maintained in MEM supplemented with 10% FBS and 1% penicillin–streptomycin. Cell culture reagents were obtained from Gibco. All cells were kept in a humidified incubator at 37 °C and 5% CO_2 and were tested routinely for mycoplasma contamination.

Western blot and immunofluorescence. Cells were collected and washed in ice-cold PBS, and cell pellets were resuspended using RIPA buffer (50 mM Tris, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% NP40 and Roche Complete Protease inhibitor) and incubated in ice for 10 min. The cell lysate was sonicated using a Bioruptor2000 (30 s on, 30 s off; 3 cycles), then centrifuged at high speed for 10 min, and the supernatant was collected and boiled for 5 min with loading buffer. The samples were separated by SDS–PAGE and transferred into PVDF membrane and blocked by 5% non-fat milk and incubated with antibodies. Signals were detected by Tanon 5200 imaging systems. The uncropped images of immunoblot are available as source data. The antibodies used in western blotting were NFKBIZ (CST, 9244S), IL1B (Proteintech, 16806), tubulin (Proteintech, 11224-1-AP), RELA (CST, 8242 S), MMP3 (BOSTER, PB0265), HNRNPA2B1 (Proteintech, 14813-1-AP), HNRNPU (Proteintech, 14599-1-AP), TOP1 (Proteintech, 20705-1-AP), DDX5 (Proteintech, 10804-1-AP), NONO (Proteintech, 11058-1-AP), p16-INK4A (Proteintech, 10883-1-AP) and HRAS (Proteintech, 18295-1-AP).

Cells were grown on glass coverslips overnight, then the medium was removed from cells. Wash buffer at room temperature (1× PBS, Hyclone, SH30265.01B) was added to wash cells three times, 5 min each time. Cells were incubated in 500 µl fixative solution (4% paraformaldehyde in PBS; Affymetrix, 4210324) at 20–25 °C for 20 min. After removing the fixative solution, cells were washed three times with 1 ml of PBS for 2–5 min, three times, followed by incubation in permeabilization solution (0.5% Triton X-100 in 1× PBS) for 15 min at 20–25 °C. After removing permeabilization solution, cells were washed three times with 1 ml of PBS for 2–5 min, and cells were incubated overnight at 4°C. After washing three times with 1 ml of PBS for 10 min, diluted secondary antibody (250 µl) was added and cells were incubated in the dark for 60 min at 20–25 °C. After washing three times with 1 ml of PBS for 5 min, pictures were taken with an inverted confocal Zeiss LSM880 fluorescence microscope.

Knockout and knock-in by CRISPR–Cas9, knockdown by ASO or shRNA, and overexpression to perturb gene expression. HEK Cas9 NFκB-GFP cells were used to generate *NFKBMARL-1* and NFKBIZ KO single clones. A pair of sgRNAs targeting the genomic loci of *NFKBMARL-1* and NFKBIZ were constructed into pLL3.7 plasmids. The plasmids were transfected into HEK Cas9 NFκB-GFP cells using Fugene HD according to the manufacturer's manual. After culturing for 7 d, the cells were sorted into 96-well plates by FACS. The KO clones were verified by PCR after cell growth. The EF1 promoter was knocked in to the TSS of *NFKBMARL-1*. Two sgRNAs were constructed in the pX330 plasmids and the EF1 promoter with 820-bp homology arms flanking the TSS of *NFKBMARL-1* was constructed into T vector pMD20. The plasmids were transfected using Fugene HD according to the manufacturer's manual. Then, the single clones were picked as described above.

To knock down NFKBIZ, HNRNPU and HNRNPA2B1, shRNAs were constructed into the plko.1 vector for targeted and scrambled RNAs. The virus was packaged according to the CRISPR screening steps, and the virus was concentrated by about 100-fold using a Lenti-X Concentrator (TAKARA) and resuspended in PBS and stored at −80 °C until use. To knock down *NFKBMARL-1*, nontarget control and *NFKBMARL-1* ASOs were electroporated into IMR90 cells with the 4D-nucleofector system (LONZA) according to the manufacturer's instructions.

To overexpress *NFKBMARL-1* and NFKBIZ, the cDNA sequence was cloned into pCDH plasmid. The virus was packaged and concentrated as previously described.

Disordered residue enrichment analysis. Protein disorder regions were evaluated using the D^2P^2 database[54] with at least 75% of the confidential score of all algorithms for each amino acid. The reference human proteome as the background was obtained from Uniprot (UP000005640), and the proteins were selected according to the GO enrichment annotation for peroxisome (GO:0005777) and transcription coactivator (GO:0003717). The nonredundant RBPs were obtained from RBPDB and the list was filtered through the human proteome down to 895 proteins.

Triplex model. To evaluate triplex formation, the NFKBIZ enhancer ChIRP peak genomic DNA sequence and *NFKBMARL-1* were input into the software LongTarget to search all base-pairing rules with parameters -1000 TT penalty and 0 CC penalty, and output the triplex model with >60 identity, >50 nucleotides and >1 mean stability, with the 100× shuffled RNA sequence as the background.

Three-dimensional modeling of RELA and *NFKBMARL-1* interaction. The MPRDock server[55] was used to model the docking between RELA and *NFKBMARL-1*. The protein structure of RELA (1NFI chain A) was obtained from the Protein Data Bank. MPRDock uses a freestyle docking strategy and calculates the docking energy score DITScorePR and the ligand root mean squared deviation (Å).

Statistics and reproducibility. Statistical significance was performed by Student's *t*-test, Fisher's exact test and Wilcoxon rank-sum test as described in figure legends with R (v3.3.2). The statistical results of source data are available in Supplementary Table 7. The sample sizes are described in the figure legends for all experiments with at least three biological replicates generated per treatment. No statistical methods were used to predetermine sample sizes, but our samples are similar to those reported in a previous publication[24]. Data were not excluded from analyses. The experiments were independently repeated at least twice and led to the same conclusions, including experiments of IF and western blot. The cell lines were obtained from public sources and authenticated; no preestablished selection criteria for the cells were used. The cells were assigned randomly to culture plates for passage or different treatments (drug treatment or genetic perturbation). When the NFκB reporter cells were treated by different concentrations of TNF, the GFP intensity changed dramatically to reflect the treatment. Tthe senescent cell induction can also be visualized by the reduction of cell number and phenotype compared to proliferating cells. Blinding was not always possible during experimental setup; however, when performing co-culture experiments for M1 macrophages and Hela cells, the investigators were blinded to quantify the tumor colonies.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The transcriptome, histone modification, transcription factor, genomic structure data from the Gene Expression Omnibus (GEO) and NODE database were used for analyzing evolutionary conservation, gene expression or epigenetic profiles for *NFKBMARL-1* or NFKBIZ under different treatments. All the publicly available datasets used in this study are labeled in the figures. The datasets used for evolutionary conservation analysis (Supplementary Fig. 1a with 11 species transcriptomes) included GSE80440, GSE101964, GSE108282, GSE66362, GSE107049, GSE85377, GSE52462, GSE66715, GSE123590, GSE106670, OEP001041, GSE75192 and GSE106670. The gene expression and epigenetic profiles of *NFKBMARL-1* or NFKBIZ under different treatments were from GSE43070, GSE72534, GSM996197, GSE109700, GSE74328, GSE130306, GSE99074, GSE100382, GSE64233, GSE91020, GSE121522, GSE55105 and GSE99544. The public databases applied in this study were REMAP (http://remap.univ-amu.fr/), DAVID (https://david.ncifcrf.gov/) and POSTAR2 (http://lulab.life.tsinghua.edu.cn/postar/).
The sequencing data of CRISPR screening sequencing, ribo-strand-specific RNA-seq and ChIRP-seq have been deposited in the GEO under accession number GSE167511. Statistics for corresponding figures are available as Supplementary Information. Any other relevant data or information about this study are available from the corresponding author upon reasonable request.

Code availability

Software and parameters used to perform the analyses are described in Supplementary Software.

Received: 1 September 2020; Accepted: 10 March 2021;
Published online: 6 May 2021

References

1. Jordan, I. K., Rogozin, I. B., Wolf, Y. I. & Koonin, E. V. Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. *Genome Res.* **12**, 962–968 (2002).
2. Sarropoulos, I., Marin, R., Cardoso-Moreira, M. & Kaessmann, H. Developmental dynamics of lncRNAs across mammalian organs and species. *Nature* **571**, 510–541 (2019).
3. Iyer, M. K. et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* **47**, 199–208 (2015).
4. Necsulea, A. et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* **505**, 635–640 (2014).
5. Li, X. et al. GRID-seq reveals the global RNA-chromatin interactome. *Nat. Biotechnol.* **35**, 940–950 (2017).
6. Kopp, F. & Mendell, J. T. Functional classification and experimental dissection of long noncoding RNAs. *Cell* **172**, 393–407 (2018).
7. Mele, M. & Rinn, J. L. 'Cat's cradling' the 3D genome by the act of lncRNA transcription. *Mol. Cell* **62**, 657–664 (2016).
8. Ulitsky, I. Evolution to the rescue: using comparative genomics to understand long noncoding RNAs. *Nat. Rev. Genet.* **17**, 601–614 (2016).
9. Liu, S. J. et al. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* **355**, aah7111 (2017).
10. Joung, J. et al. Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* **548**, 343–346 (2017).
11. Zhu, S. Y. et al. Genome-scale deletion screening of human long noncoding RNAs using a paired-guide RNA CRISPR–Cas9 library. *Nat. Biotechnol.* **34**, 1279–1286 (2016).
12. Cheng, H. et al. Repression of human and mouse brain inflammaging transcriptome by broad gene-body histone hyperacetylation. *Proc. Natl Acad. Sci. USA* **115**, 7611–7616 (2018).

13. Glass, C. K., Saijo, K., Winner, B., Marchetto, M. C. & Gage, F. H. Mechanisms underlying inflammation in neurodegeneration. *Cell* **140**, 918–934 (2010).
14. Kawahara, T. L. A. et al. SIRT6 links histone H3 lysine 9 deacetylation to NFκB-dependent gene expression and organismal life span. *Cell* **136**, 62–74 (2009).
15. Zhang, G. et al. Hypothalamic programming of systemic ageing involving IKK-β, NFκB and GnRH. *Nature* **497**, 211–216 (2013).
16. Gorgoulis, V. et al. Cellular senescence: defining a path forward. *Cell* **179**, 813–827 (2019).
17. Guttman, M. & Rinn, J. L. Modular regulatory principles of large non-coding RNAs. *Nature* **482**, 339–346 (2012).
18. Bao, Z. Y. et al. LncRNADisease 2.0: an updated database of long noncoding RNA-associated diseases. *Nucleic Acids Res.* **47**, D1034–D1037 (2019).
19. Landrum, M. J. et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* **42**, D980–D985 (2014).
20. Chen, W. et al. CRISPRlnc: a manually curated database of validated sgRNAs for lncRNAs. *Nucleic Acids Res.* **47**, D63–D68 (2019).
21. Amaral, P. P. et al. Genomic positional conservation identifies topological anchor point RNAs linked to developmental loci. *Genome Biol.* **19**, 32 (2018).
22. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* **19**, 327–341 (2018).
23. Green, C. D. et al. Impact of dietary interventions on noncoding RNA networks and mRNAs encoding chromatin-related factors. *Cell Rep.* **18**, 2957–2968 (2017).
24. De Cecco, M. et al. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature* **566**, 73–78 (2019).
25. Fanucchi, S. et al. Immune genes are primed for robust transcription by proximal long noncoding RNAs located in nuclear compartments. *Nat. Genet.* **51**, 138–150 (2019).
26. Beerman, I. et al. Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc. Natl Acad. Sci. USA* **107**, 5465–5470 (2010).
27. Imamura, K. et al. Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Mol. Cell* **53**, 393–406 (2014).
28. Xia, X. et al. Three-dimensional facial-image analysis to predict heterogeneity of the human ageing rate and the impact of lifestyle. *Nat. Metab.* **2**, 946–957 (2020).
29. Shchukina, I. et al. Enhanced epigenetic profiling of classical human monocytes reveals a specific signature of healthy aging in the DNA methylome. *Nat. Aging* **1**, 124–141 (2020).
30. Miyake, T. et al. IκBζ is essential for natural killer cell activation in response to IL-12 and IL-18. *Proc. Natl Acad. Sci. USA* **107**, 1 7680–17685 (2010).
31. Okamoto, K. et al. IκBζ regulates T_H17 development by cooperating with ROR nuclear receptors. *Nature* **464**, 1381–U13 (2010).
32. Muller, A. et al. IκBζ is a key transcriptional regulator of IL-36-driven psoriasis-related gene expression in keratinocytes. *Proc. Natl Acad. Sci. USA* **115**, 10088–10093 (2018).
33. Alexander, E. et al. IκBζ is a regulator of the senescence-associated secretory phenotype in DNA damage- and oncogene-induced senescence. *J. Cell Sci.* **126**, 3738–3745 (2013).
34. Koelman, L., Pivovarova-Ramich, O., Pfeiffer, A. F. H., Grune, T. & Aleksandrova, K. Cytokines for evaluation of chronic inflammatory status in ageing research: reliability and phenotypic characterisation. *Immun. Ageing* **16**, 11 (2019).
35. de Gonzalo-Calvo, D. et al. Differential inflammatory responses in aging and disease: TNF-α and IL-6 as possible biomarkers. *Free Radic. Biol. Med.* **49**, 733–737 (2010).
36. Rea, I. M. et al. Age and age-related diseases: role of inflammation triggers and cytokines. *Front. Immunol.* **9**, 586 (2018).
37. Park, S. H. et al. Type I interferons and the cytokine TNF cooperatively reprogram the macrophage epigenome to promote inflammatory activation. *Nat. Immunol.* **18**, 1104–1116 (2017).
38. Wright, H. L., Thomas, H. B., Moots, R. J. & Edwards, S. W. RNA-seq reveals activation of both common and cytokine-specific pathways following neutrophil priming. *PLoS ONE* **8**, e58598 (2013).
39. Young, A. R. J. et al. Autophagy mediates the mitotic senescence transition. *Genes Dev.* **23**, 798–803 (2009).
40. Jin, F. L. et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **503**, 290–294 (2013).
41. Schmidt, S. F. et al. Acute TNF-induced repression of cell identity genes is mediated by NFκB-directed redistribution of cofactors from super-enhancers. *Genome Res.* **25**, 1281–1294 (2015).
42. Higashijima, Y. et al. Coordinated demethylation of H3K9 and H3K27 is required for rapid inflammatory responses of endothelial cells. *EMBO J.* **39**, e103949 (2020).
43. Borghini, L., Lu, J., Hibberd, M. & Davila, S. Variation in genome-wide NFκB RELA binding sites upon microbial stimuli and identification of a virus response profile. *J. Immunol.* **201**, 1295–1305 (2018).
44. Tasdemir, N. et al. BRD4 connects enhancer remodeling to senescence immune surveillance. *Cancer Discov.* **6**, 612–629 (2016).
45. Furman, D. et al. Expression of specific inflammasome gene modules stratifies older individuals into two extreme clinical and immunological states. *Nat. Med.* **23**, 174–184 (2017).
46. Zannas, A. S. et al. Epigenetic upregulation of FKBP5 by aging and stress contributes to NFκB-driven inflammation and cardiovascular risk. *Proc. Natl Acad. Sci. USA* **116**, 11370–11379 (2019).
47. Kakiuchi, N. et al. Frequent mutations that converge on the NFKBIZ pathway in ulcerative colitis. *Nature* **577**, 260–265 (2020).
48. Kumar, M., Gromiha, M. M. & Raghava, G. P. Prediction of RNA binding sites in a protein using SVM and PSSM profile. *Proteins* **71**, 189–194 (2008).
49. Kim, O. T., Yura, K. & Go, N. Amino acid residue doublet propensity in the protein–RNA interface and its application to RNA interface prediction. *Nucleic Acids Res.* **34**, 6450–6460 (2006).
50. He, S., Zhang, H., Liu, H. & Zhu, H. LongTarget: a tool to predict lncRNA DNA-binding motifs and binding sites via Hoogsteen base-pairing analysis. *Bioinformatics* **31**, 178–186 (2015).
51. Nair, S. J. et al. Phase separation of ligand-activated enhancers licenses cooperative chromosomal enhancer assembly. *Nat. Struct. Mol. Biol.* **26**, 193–203 (2019).
52. Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855 (2018).
53. Cui, R. F. et al. Relaxed selection limits life span by increasing mutation load. *Cell* **178**, 385–399 (2019).
54. Oates, M. E. et al. D^2P^2: database of disordered protein predictions. *Nucleic Acids Res.* **41**, D508–D516 (2013).
55. He, J., Tao, H. & Huang, S. Y. Protein-ensemble-RNA docking by efficient consideration of protein flexibility through homology models. *Bioinformatics* **35**, 4994–5002 (2019).

Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (91749205, 92049302, 32088101), China Ministry of Science and Technology (2020YFA0804000, 2016YFE0108700) and Shanghai Municipal Science and Technology Major Project (2017SHZDZX01) to J.-D.J.H. We thank J. Sedivy at Brown University, Y. Sun at Shanghai Institute of Nutrition and Health and F. Lan at Fudan University for invaluable suggestions and discussions, and the analysis and technology platform at Shanghai Institute of Nutrition and Health for FACS and confocal microscope support.

Author contributions

J.-D.J.H. and D.C. conceived and designed the study, analyses and experiments, and wrote the manuscript. D.C. performed the analyses and experiments.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s43587-021-00056-0.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43587-021-00056-0.

Correspondence and requests for materials should be addressed to J.-D.J.H.

Peer review information *Nature Aging* thanks Shinichi Nakagawa, Yousin Suh, Gian Gaetano Tartaglia and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

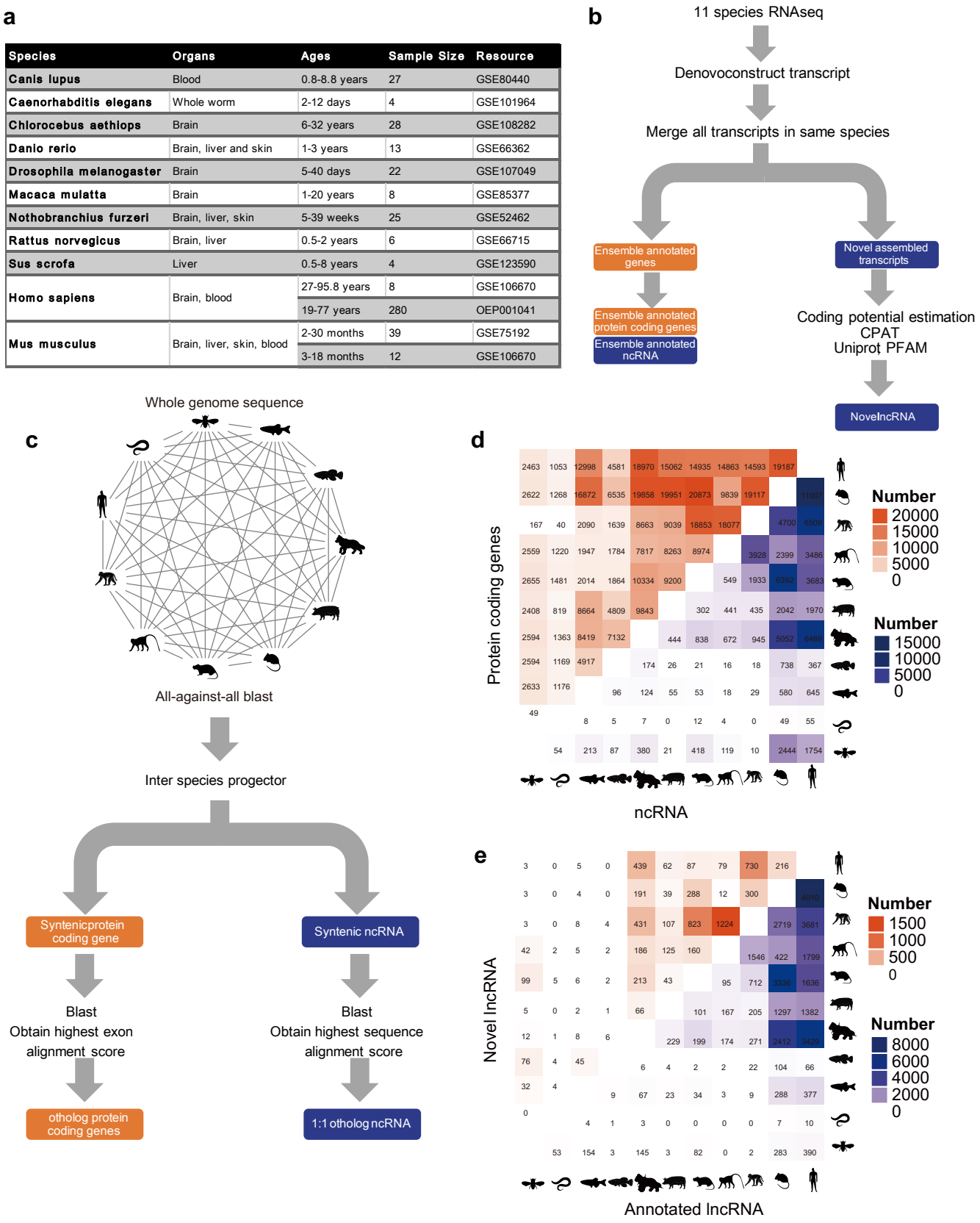
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

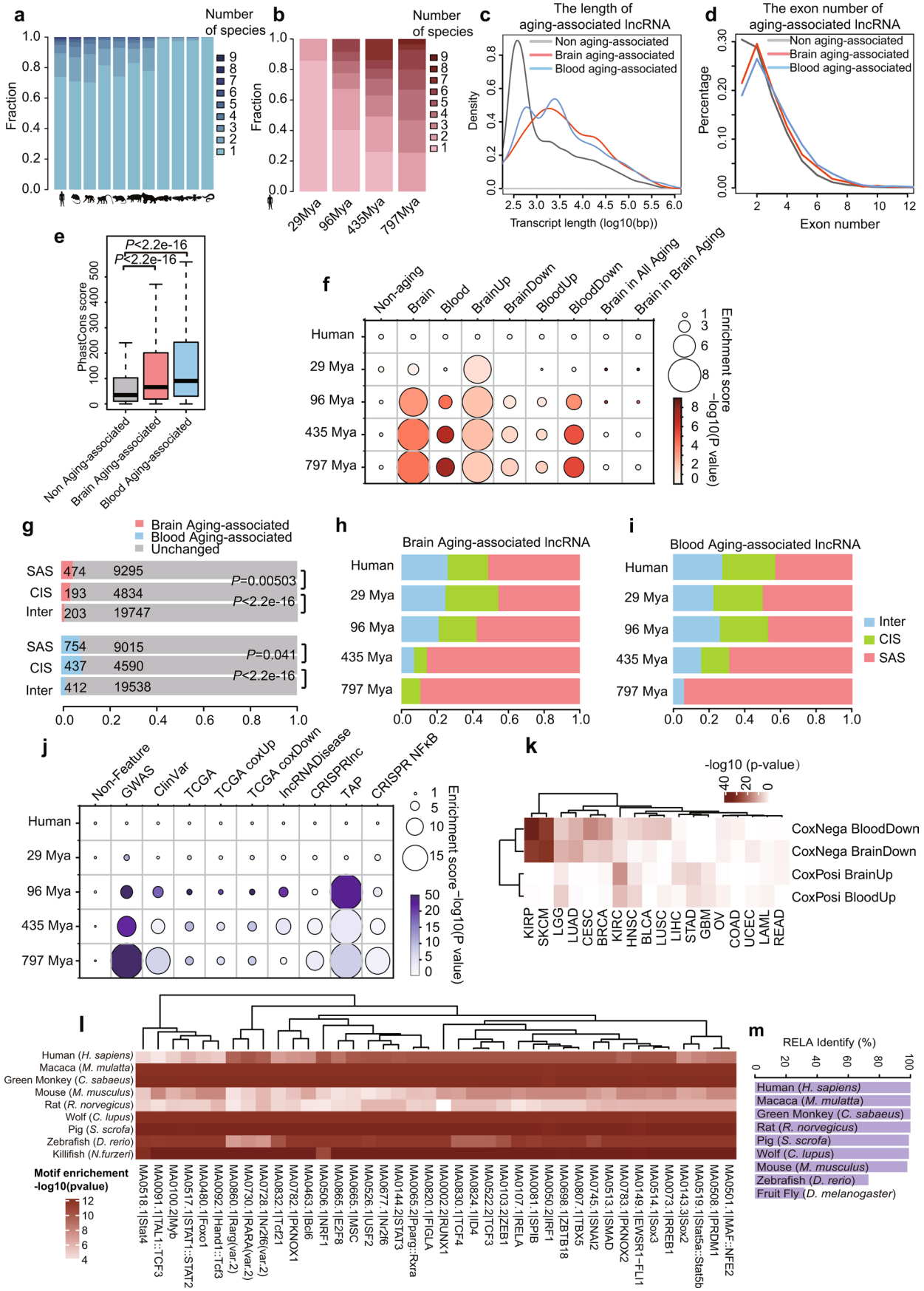
© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021

a

Species	Organs	Ages	Sample Size	Resource
Canis lupus	Blood	0.8-8.8 years	27	GSE80440
Caenorhabditis elegans	Whole worm	2-12 days	4	GSE101964
Chlorocebus aethiops	Brain	6-32 years	28	GSE108282
Danio rerio	Brain, liver and skin	1-3 years	13	GSE66362
Drosophila melanogaster	Brain	5-40 days	22	GSE107049
Macaca mulatta	Brain	1-20 years	8	GSE85377
Nothobranchius furzeri	Brain, liver, skin	5-39 weeks	25	GSE52462
Rattus norvegicus	Brain, liver	0.5-2 years	6	GSE66715
Sus scrofa	Liver	0.5-8 years	4	GSE123590
Homo sapiens	Brain, blood	27-95.8 years	8	GSE106670
		19-77 years	280	OEP001041
Mus musculus	Brain, liver, skin, blood	2-30 months	39	GSE75192
		3-18 months	12	GSE106670

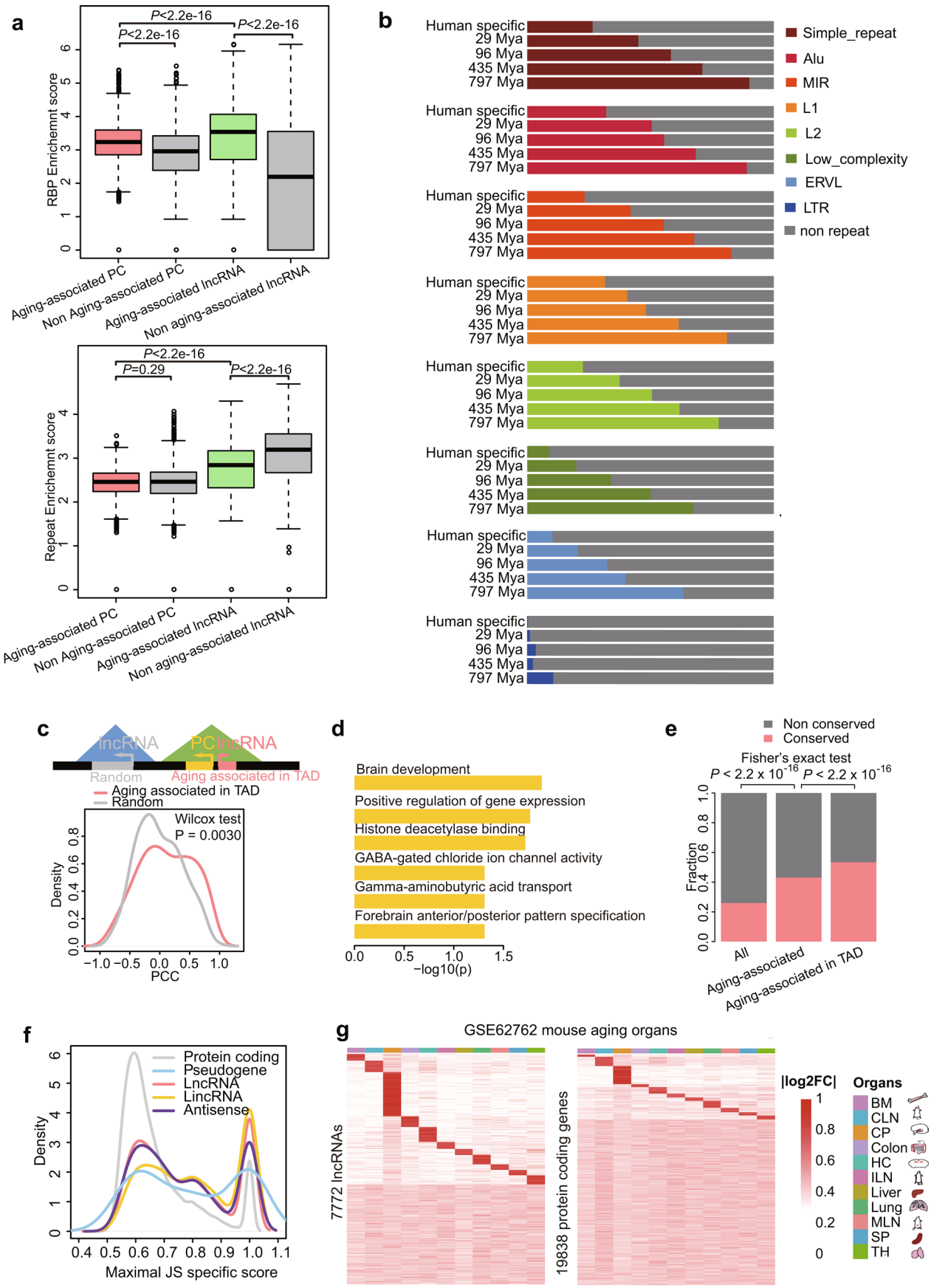


Extended Data Fig. 1 | Identification of novel lncRNAs and evolutionarily conserved lncRNAs. a, The public RNA-seq data of 11 species for evolutionary associated lncRNAs analysis. **b**, Pipeline for de novo assembly of transcripts and annotation of lncRNAs. **c**, Schematic presentation of pipeline for detecting one-to-one orthologous lncRNA families. **d**, Distribution of one-to-one orthologous protein-coding genes (red) and ncRNAs (blue) in 11 species. **e**, Distribution of one-to-one ortholog annotated novel lncRNAs (red) and Ensembl annotated lncRNAs (blue) in 11 species.



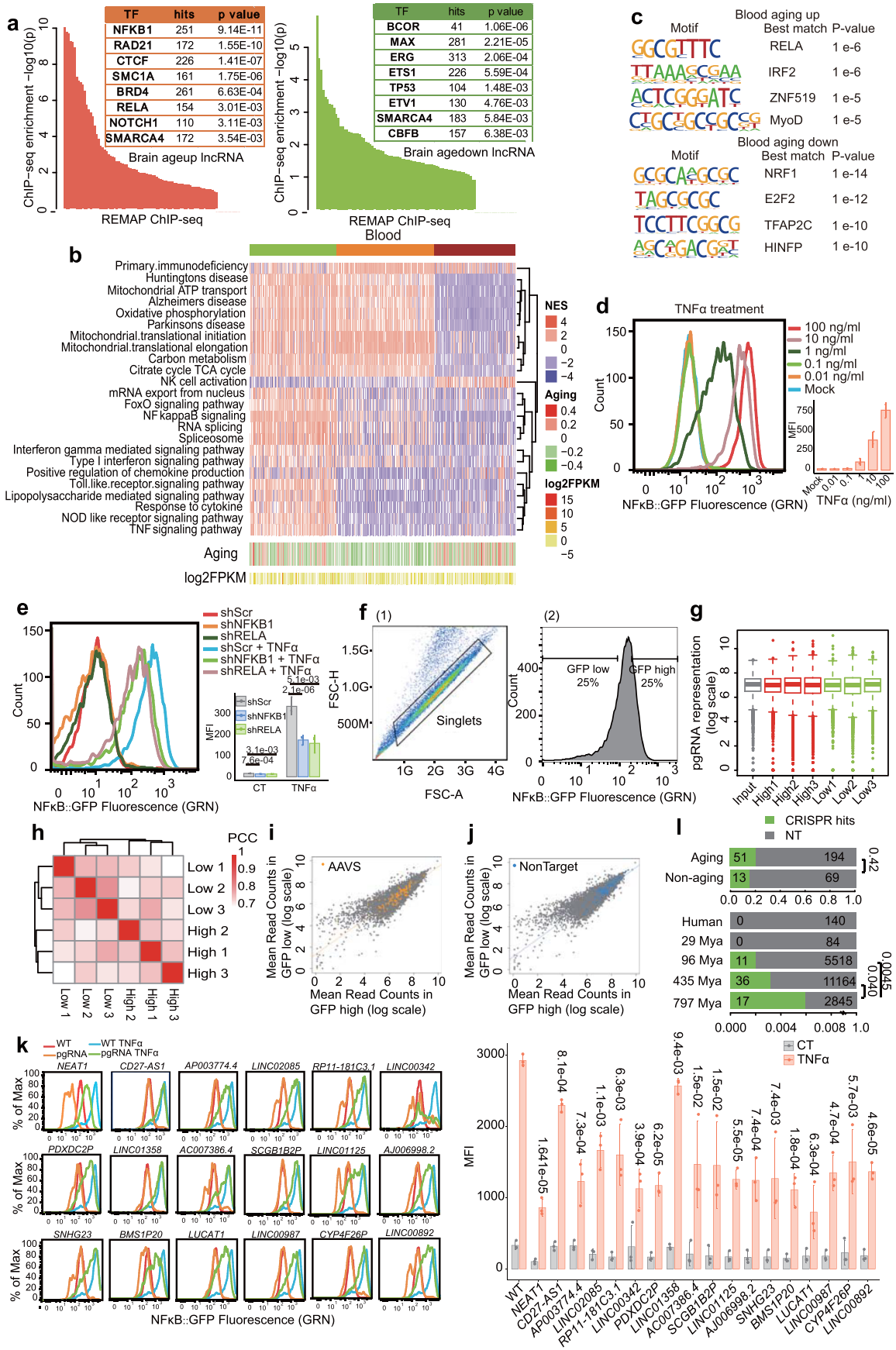
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Features of aging-associated lncRNAs in evolutionary ages. a, Fraction of the lncRNA orthologous conserved to other species. **b**, Fraction of human-associated lncRNAs conserved to other species in different evolutionary ages. **c**, Density plot of transcripts length of human brain and blood aging-associated lncRNAs and non-aging-associated lncRNAs. **d**, Fraction of transcripts exon numbers of human brain and blood aging-associated lncRNAs and non-aging-associated lncRNAs. **e**, PhasCons20 scores in aging-associated and non-aging-associated lncRNAs. Wilcoxon rank-sum test. Centerline indicates median value; box limits are the interquartile range of 25th and 75th percentiles; whiskers each extend 1.5 times of the interquartile range of maxima and minima. P values were calculated by two-sided Wilcoxon rank-sum test with correction. **f**, Enrichment analysis of aging-associated protein-coding genes in different evolutionary ages by two-sided Fisher's exact test. 2000 random non-aging-associated protein-coding genes serve as control. **g**, Distribution of physical proximity with protein-coding genes for aging-associated lncRNAs. SAS is defined by antisense transcript overlapping with protein-coding genes. CIS is defined by the distance between lncRNA and protein-coding gene <2 kb. Inter is defined by the distance to protein-coding genes >2 kb. **h**, Fraction of brain aging-associated lncRNAs with different physical proximity to protein-coding genes are conserved to different evolutionary ages. **i**, Fraction of blood aging-associated lncRNAs with different physical proximity to protein-coding genes are conserved to different evolutionary ages. **j**, Enrichment analysis of functional lncRNAs in different evolutionary ages by two-sided Fisher's exact test. 2000 random non-feature lncRNAs serve as control. **k**, Survival associated and aging-associated lncRNAs enriched in different cancer types in TCGA dataset. **l**, Motif enrichment heatmap in TSS regions of aging-associated lncRNAs by Homer. The heatmap shows the motifs significantly enrich in all species ($p < 0.05$). **m**, Identity percentage of RELA protein sequence from humans to other species by BLASTP.



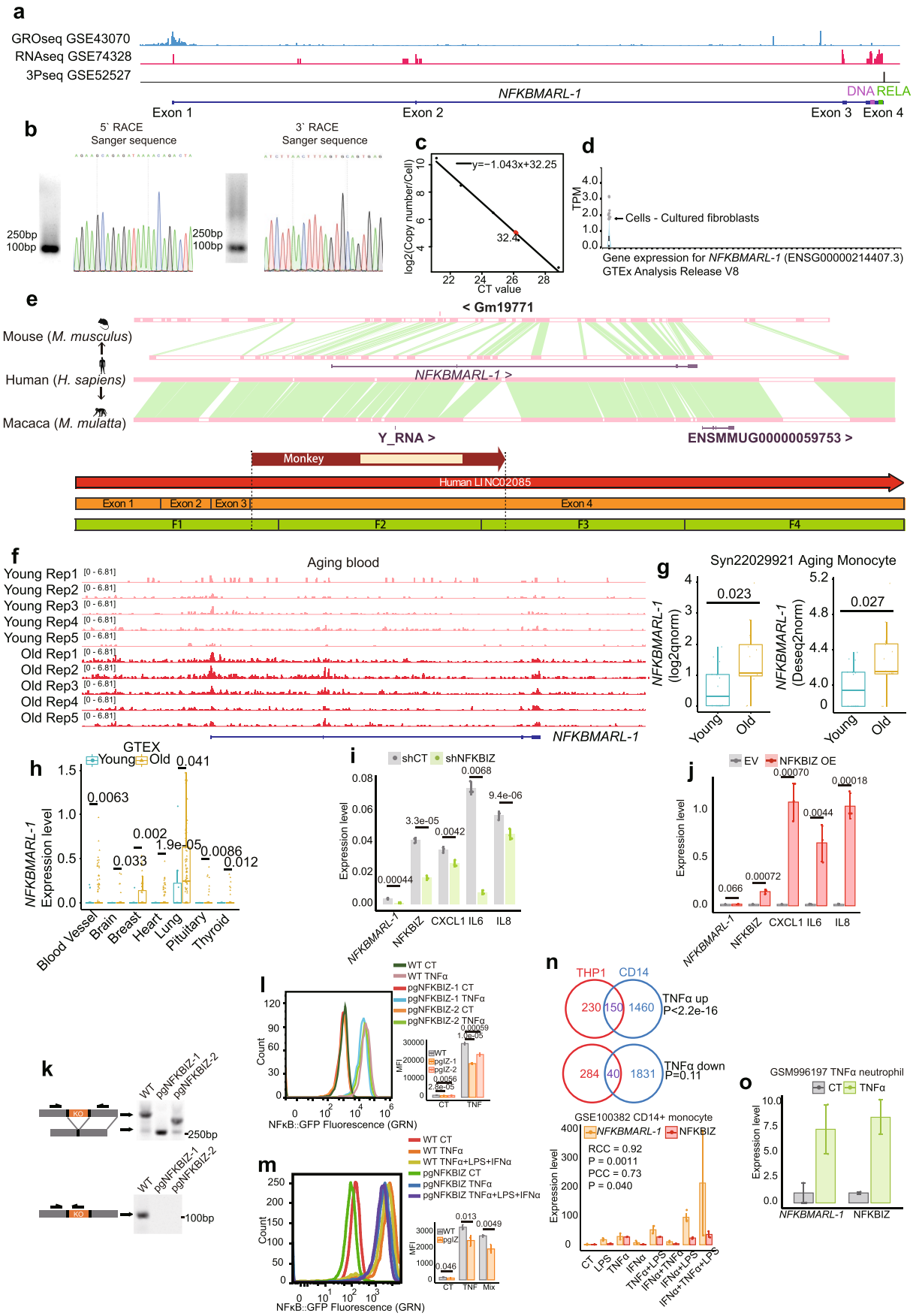
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Regulatory elements in aging-associated lncRNAs. a, Binding sites enrichment of RNA binding proteins (up) and repeats (down) distribute on the protein-coding genes, aging-associated lncRNAs and non-aging-associated lncRNAs. P values were calculated by two-sided Wilcoxon rank-sum test with correction. Centerline indicates median value; box limits are the interquartile range of 25th and 75th percentiles; whiskers each extend 1.5 times of the interquartile range of maxima and minima; dots are outliers. **b**, Fraction of different types of transposon elements in lncRNAs conserved to different evolutionary ages. **c**, Density distribution of the Pearson correlation coefficients between closest protein-coding genes to aging-associated lncRNAs in same TAD region. TAD regions are annotated from Hi-C dataset of dorsolateral prefrontal cortex (GSE87112). P values were calculated by two-sided Wilcoxon rank-sum test with correction. **d**, Functional enrichment of closest protein-coding genes in the same TAD region with aging-associated lncRNAs. **e**, Fraction of conserved lncRNAs in aging-associated lncRNAs within TAD. P values were calculated by two-sided Fisher's exact test. **f**, Density distribution of maximal Jensen Shannon specific score for fold change of each protein-coding genes and lncRNA including lincRNA and antisense. **g**, Heatmap of the |log2FC| of protein-coding genes and lncRNAs in 11 tissues. The abbreviations are as follows: BM for bone marrow; CLN for cervical lymph nodes; CP for choroid plexus; HC for hippocampus; ILN for inguinal lymph nodes; MLN for mesenteric lymph nodes; SP for spleen; TH for thymus.



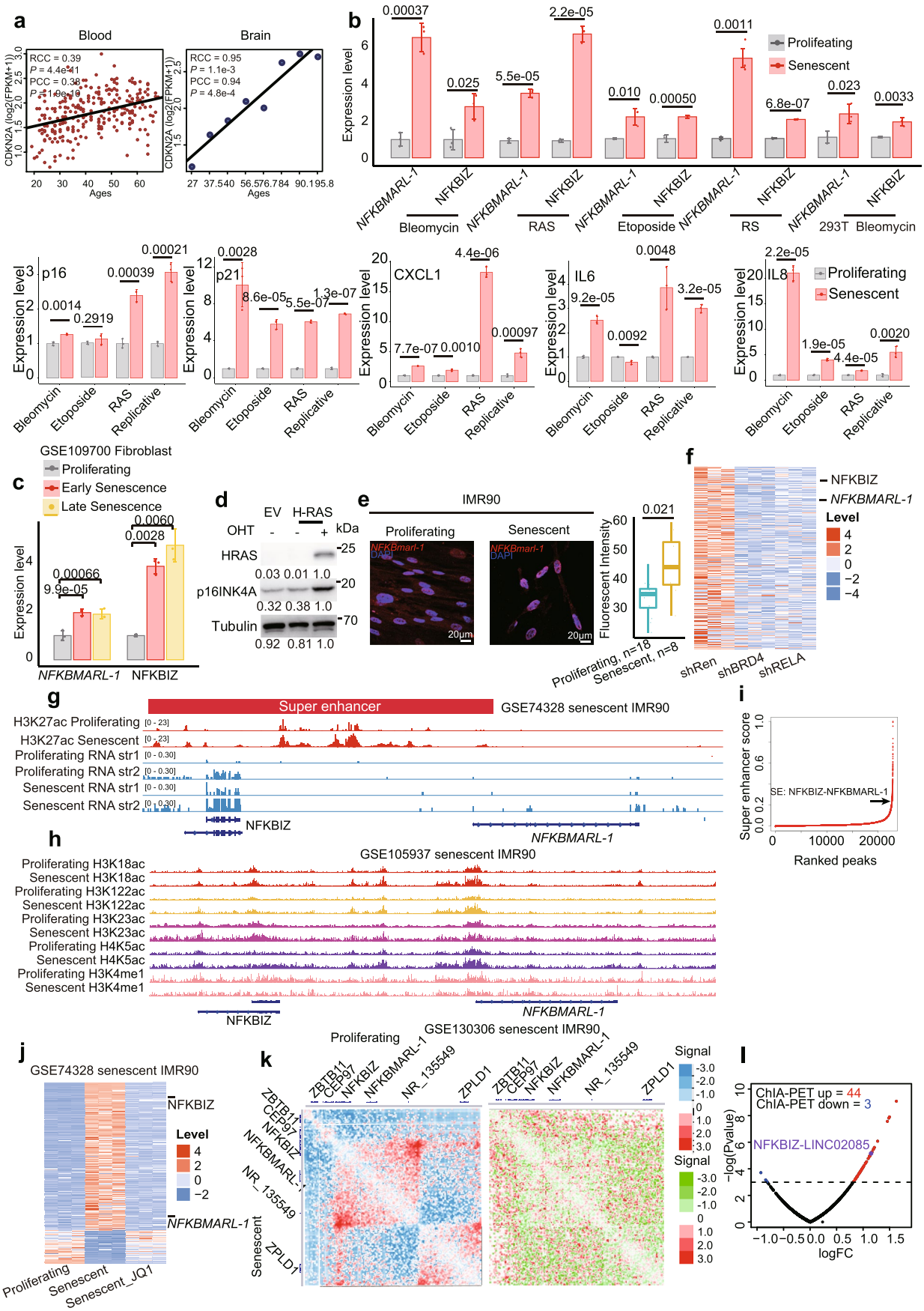
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Enriched functions of blood aging-associated lncRNAs and the general information of CRISPR screen. a, TF enrichment in brain aging up or down lncRNAs in REMAP ChIP-seq dataset. **b**, Heatmap showing NES from guilt-by-association to annotate the function of blood aging-associated lncRNAs. Each column is lncRNA and each row is GO term. The other 2 columns are age correlations and expression levels of blood aging-associated lncRNAs **c**, Motif enrichment of promoter regions of blood aging up lncRNAs by Homer. Equal number of random background regions serve as control. **d**, GFP fluorescent intensity of serial concentrations of TNFα stimulated NFκB pathway reporter cells, revealed by FACS. **e**, GFP fluorescent intensity of TNFα stimulated NFκB pathway reporter cells after knocking down NFKB1 or RELA by shRNAs. **f**, Gating strategy: (1) singlets were selected by FSC-H against FSC-A; (2) singlets were gated for GFP high/low populations. **g**, Boxplot showing the pgRNAs distribution after selected GFP high and low under TNFα stimulation from n = 3 replicates. Centerline indicates median value; box limits are the interquartile range of 25th and 75th percentiles; whiskers each extend 1.5 times of the interquartile range of maxima and minima; dots are outliers. **h**, The Pearson correlation coefficient of CRISPR sequence normalized reads between 3 independent replicates of GFP high and GFP low sorted cells. **i**, Mean read counts of pgRNAs targeting the AAVS1 loci. **j**, Mean read counts of pgRNAs targeting the non-target controls. **k**, GFP fluorescent intensity of TNFα stimulated NFκB pathway reporter cells after knocking out 18 GFP low selected lncRNAs by each pgRNAs. All the Fluorescent intensity measurements including WT were performed in the same batch. The MFI in (**d, e, k**) were analyzed by FlowJo. Data in (**d, e, k**) are presented as mean ± s.d. in n = 3 independent experiments. All p values in (**d, e, k**) were calculated using two-sided unpaired Student's t-test.



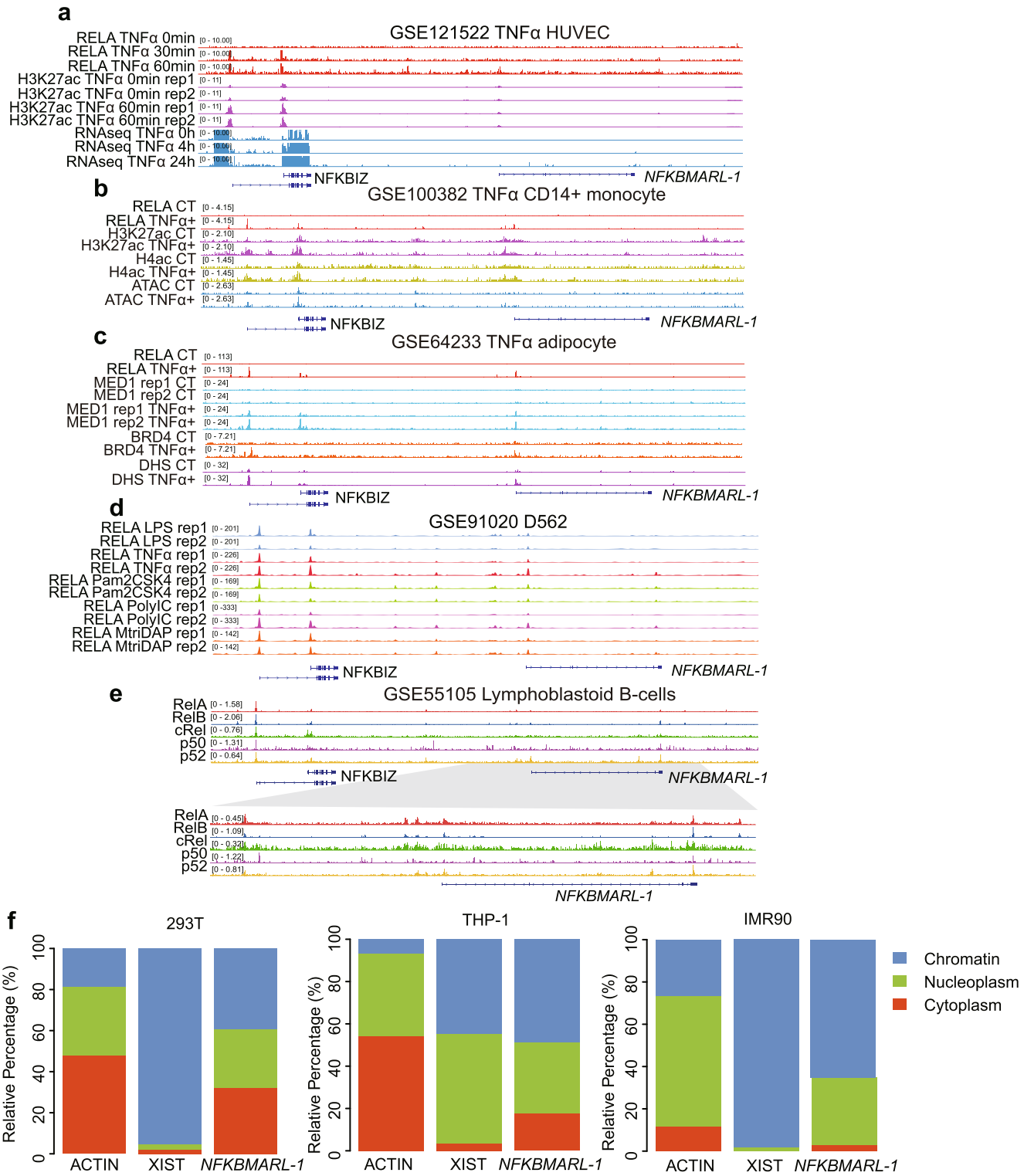
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | General feature and gene expression of *NFKBMARL-1* in aging and inflammation. a, The genome browser overview of GRO-seq, RNA-seq and polyA-seq in *NFKBMARL-1*. Purple mark is the DNA binding region, green mark is RELA binding region. **b**, Gel imaging of PCR product and Sanger sequencing of 5′ end and 3′ end RACE. **c**, Copy number of *NFKBMARL-1* in IMR90, revealed by RT-qPCR. **d**, Expression level of *NFKBMARL-1* in different tissue and cell from GTEx analysis release V8 database. **e**, Top, syntenic region comparison from human to mouse and macaca in *NFKBMARL-1* genomic region from Ensembl database. Bottom, alignment of transcripts from monkey to human. **f**, Genome browser view of RNA-seq profile of *NFKBMARL-1* in aging PBMC. **g**, The expression level of *NFKBMARL-1* in aging monocytes with two normalization methods, log2qnorm and Deseq2norm. **h**, The expression level of *NFKBMARL-1* in GTEX database for 7 organs of blood vessel, brain, breast, heart, lung, pituitary, thyroid. All tissues have no significant changes and are thus not shown. The young samples are ages <30, and old samples are ages >60. Centerline indicates median value; box limits are the interquartile range of 25th and 75th percentiles; whiskers each extend 1.5 times of the interquartile range of maxima and minima; dots are expression value (**g, h**). **i**, Expression of *NFKBMARL-1*, CXCL1, IL6 and IL8 in NFKBIZ shRNA knockdown and TNFα stimulated 293 T, revealed by RT-qPCR. **j**, Expression of *NFKBMARL-1*, CXCL1, IL6 and IL8 in control and NFKBIZ overexpressed TNFα stimulated 293 T cells, revealed by RT-qPCR. **k**, Gel image of PCR validating the deletion of NFKBIZ by 2 pgRNAs. The MFI were analyzed by FlowJo. **l**, Fluorescent intensity of NFκB pathway reporter cell, after knocking out NFKBIZ and TNFα stimulation, revealed by FACS. The MFI were analyzed by FlowJo. **m**, Fluorescent intensity of NFκB pathway reporter cell, after knocking out NFKBIZ and TNFα, LPS and IFNα co-treatment for FACS. The MFI were analyzed by FlowJo. **n**, Up: The overlap of TNFα up- and down-regulated genes between THP1 and monocytes, shown by Venn diagram. The significance of overlap is determined by Fisher's exact test. Down: Bar plot showing the expression of *NFKBMARL-1* after TNFα (10 ng/ml), LPS (10 ng/ml) and IFNα (25 ng/ml) treatment for 3hrs in CD14 + monocyte, revealed by RT-qPCR. **o**, Expression of *NFKBMARL-1* and NFKBIZ in TNFα (10 ng/ml) treated neutrophils. Data in (**i, j, l, m**) are presented as mean ± s.d. in n = 3 independent experiments. All p values in (**g-j, l-n**) were calculated using two-sided unpaired Student's t-test.

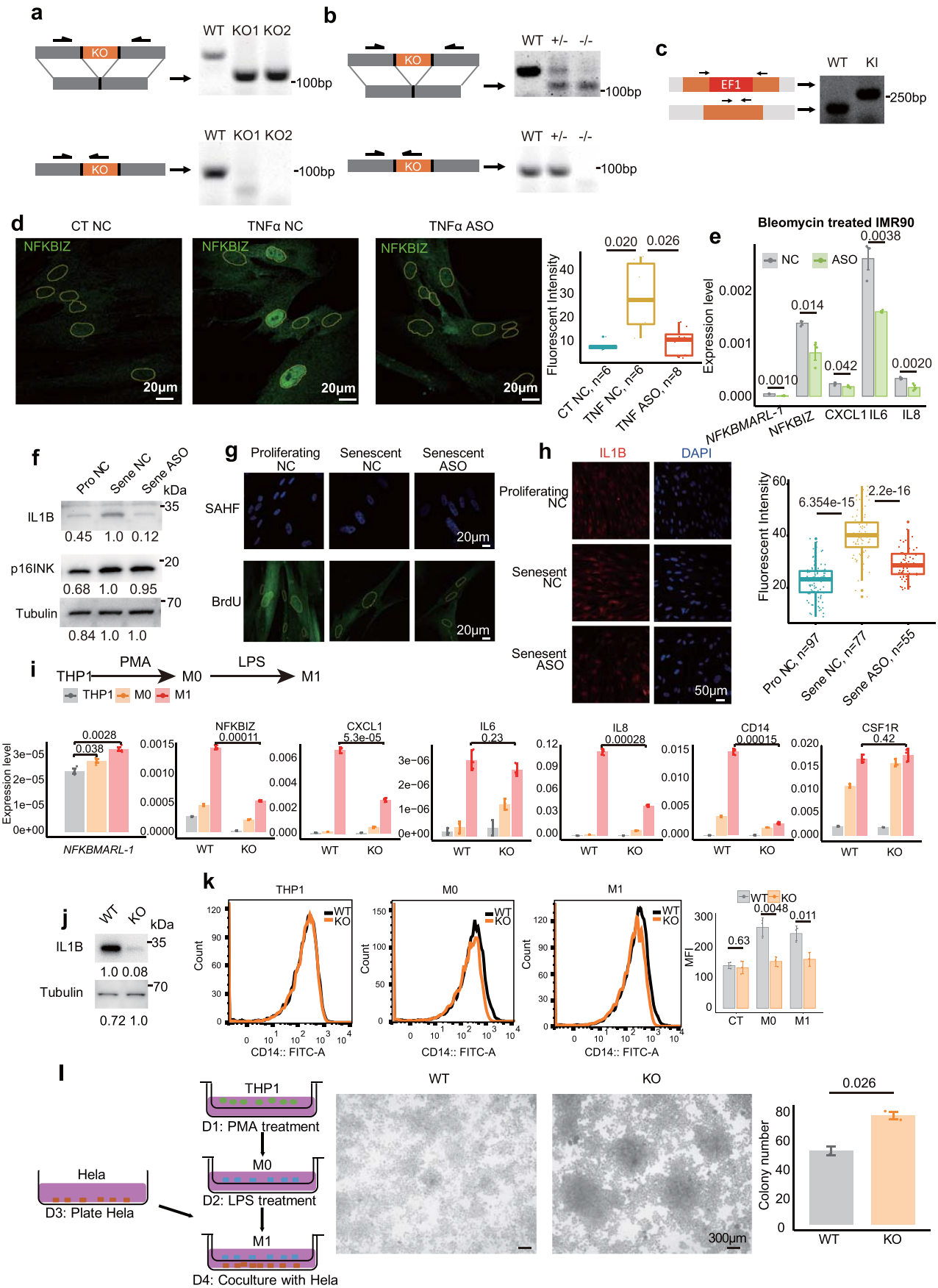


Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | *NFKBMARL-1* is up-regulated during cellular senescence and upon NFκB pathway activation. a, The expression of senescent marker gene CDKN2A (p16INK4A) in aging brain and blood. The RCC and PCC between ages and expression of CDKN2A and their p values are shown. **b**, Up: Expression level of *NFKBMARL-1* and NFKBIZ after treated by bleomycin (10ug/ml), etoposide (10uM), 4OHT (100 nM) induced HRAS overexpress and replicative senescence in IMR90 or 293 T, revealed by RT-qPCR. Down: The expression levels of senescent markers including p16 and p21 and SASP factors including CXCL1, IL6 and IL8 in 4 different senescence models. **c**, Expression of NFKBIZ and *NFKBMARL-1* in early and late senescent IMR90 cells revealed by RNA-seq. **d**, 4OHT (100 nM) induced HRAS overexpressing IMR90 cells with HRAS and p16-INK4A protein level measured by Western blot. **e**, smFISH for *NFKBMARL-1* in bleomycin-induced senescent IMR90 (left). Statistics of fluorescent intensity of *NFKBMARL-1* in senescent IMR90 (right). Centerline indicates median value; box limits are the interquartile range of 25th and 75th percentiles; whiskers each extend 1.5 times of the interquartile range of maxima and minima; dots are each fluorescent intensity. The results were analyzed by ImageJ. **f**, Global expression profile of RNA-seq for RELA knockdown in oncogene-induced senescent IMR90. The arrows point out NFKBIZ and *NFKBMARL-1*. **g**, Genome browser view for the H3K27ac modification and expression level of *NFKBMARL-1* and NFKBIZ under cellular senescence process. **h**, The genome browser view of histone modification H3K18ac, H3K122ac, H3K23ac, H4K5ac and H3K4me1 in senescent IMR90. **i**, Dot plot of super-enhancer scores in senescent H3K27ac identified peak. The score in the super-enhancer covered NFKBIZ and *NFKBMARL-1* is 0.25 (Chromosome region is chr3:101809216-101964911). **j**, Global expression profile of RNA-seq for oncogene-induced senescent IMR90 and treated by JQ1. The arrows point out NFKBIZ and *NFKBMARL-1*. **k**, Left, genomic contact matrices of proliferating and senescent IMR90 in chromosome 3. Right, subtraction heatmap of senescent vs proliferating. **l**, The TNFα induced change of Pol II ChIA-PET signal in HUVEC within each of the 2407 TADs obtained from the HUVEC HiC dataset (GSE63525). The dotted line indicates $p = 0.05$. The number of TADs with significantly increased and decreased Pol II ChIA-PET signals are indicated by the red and blue labels. Under TNFα treatment, 41 and 3 of the annotated TADs show significantly increased and decreased Pol II in HUVEC cells, respectively. Importantly, the ChIA-PET signals in the TAD enclosing NFKBIZ and *NFKBMARL-1* is strongly increased after TNFα treatment, and the Pol II PET signals ranks at the 10th of the ChIA-PET signal up-regulated TAD among a total 2407 TADs. P values were defined and adjusted by edgeR. Data in (**b, c**) are presented as mean ± s.d. in $n = 3$ independent experiments. All p values in (**b, c**) were calculated using two-sided unpaired Student's t-test.

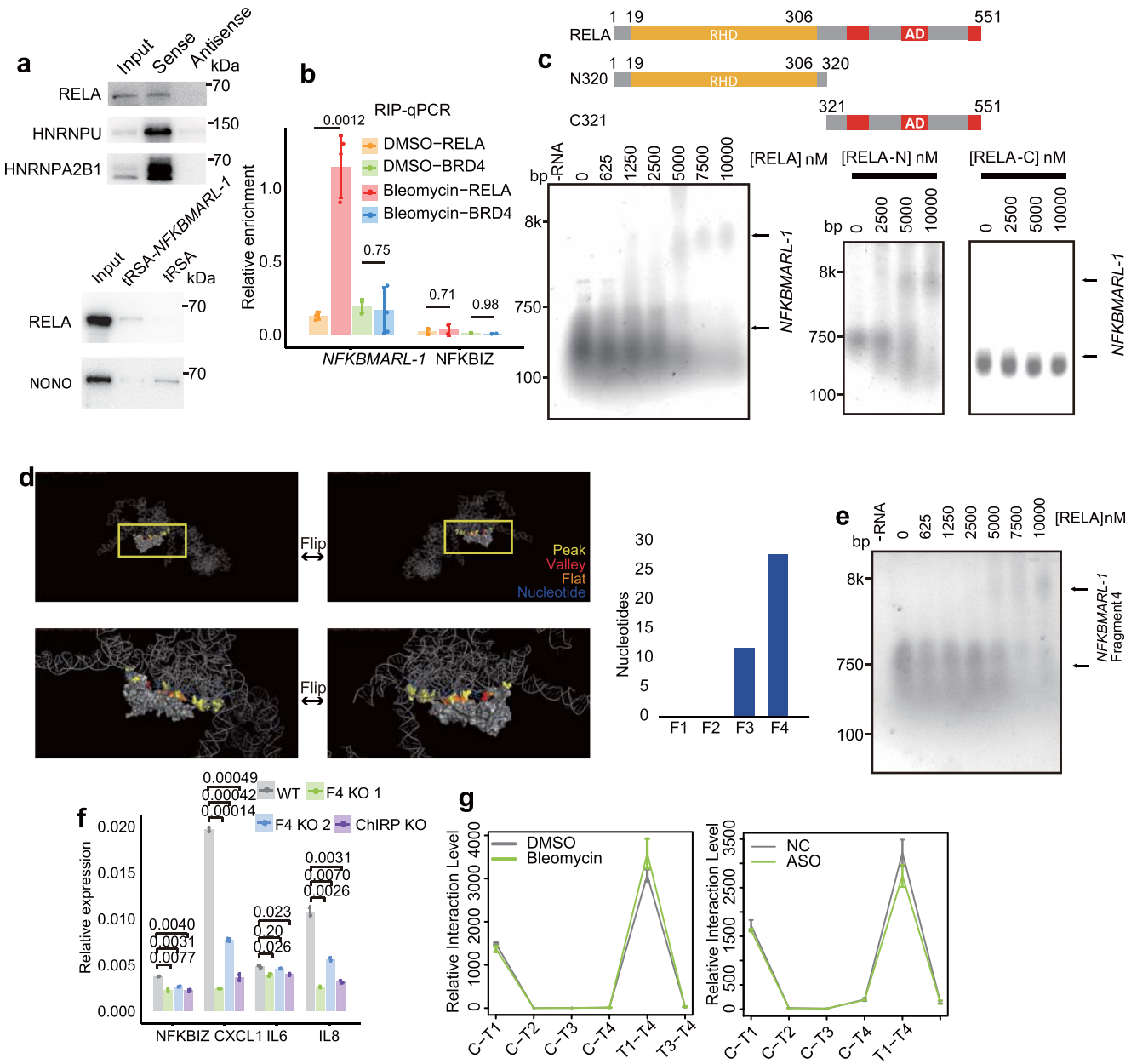


Extended Data Fig. 7 | The chromatin binding profile of *NFKBMARL-1*. **a**, Genome browser view showing the transcription factor binding profile of RELA and the epigenetic status of H3K27ac after in TNFα stimulated HUVEC. **b**, Genome browser view showing the transcription factor binding profile of RELA and the epigenetic status of H3K27ac, H4ac and ATAC after TNFα treatment in CD14 + monocyte. **c**, Genome browser view showing the transcription factor binding profile of RELA, MED1 and BRD4 and the epigenetic status of H3K27ac after TNFα treatment in adipocyte. **d**, Genome browser view showing RELA binding profile on NFKBIZ and *NFKBMARL-1* under LPS, TNFα, Pam2CSK4, PolyIC and MtriDAP treatment in D562. **e**, Genome browser view showing 5 NFκB pathway family members binding on NFKBIZ and *NFKBMARL-1* in lymphoblastoid B cells. **f**, Cellular fraction distribution of *NFKBMARL-1* in TNFα (10 ng/ml) induced 293 T, THP1 and IMR90, revealed by RT-qPCR.

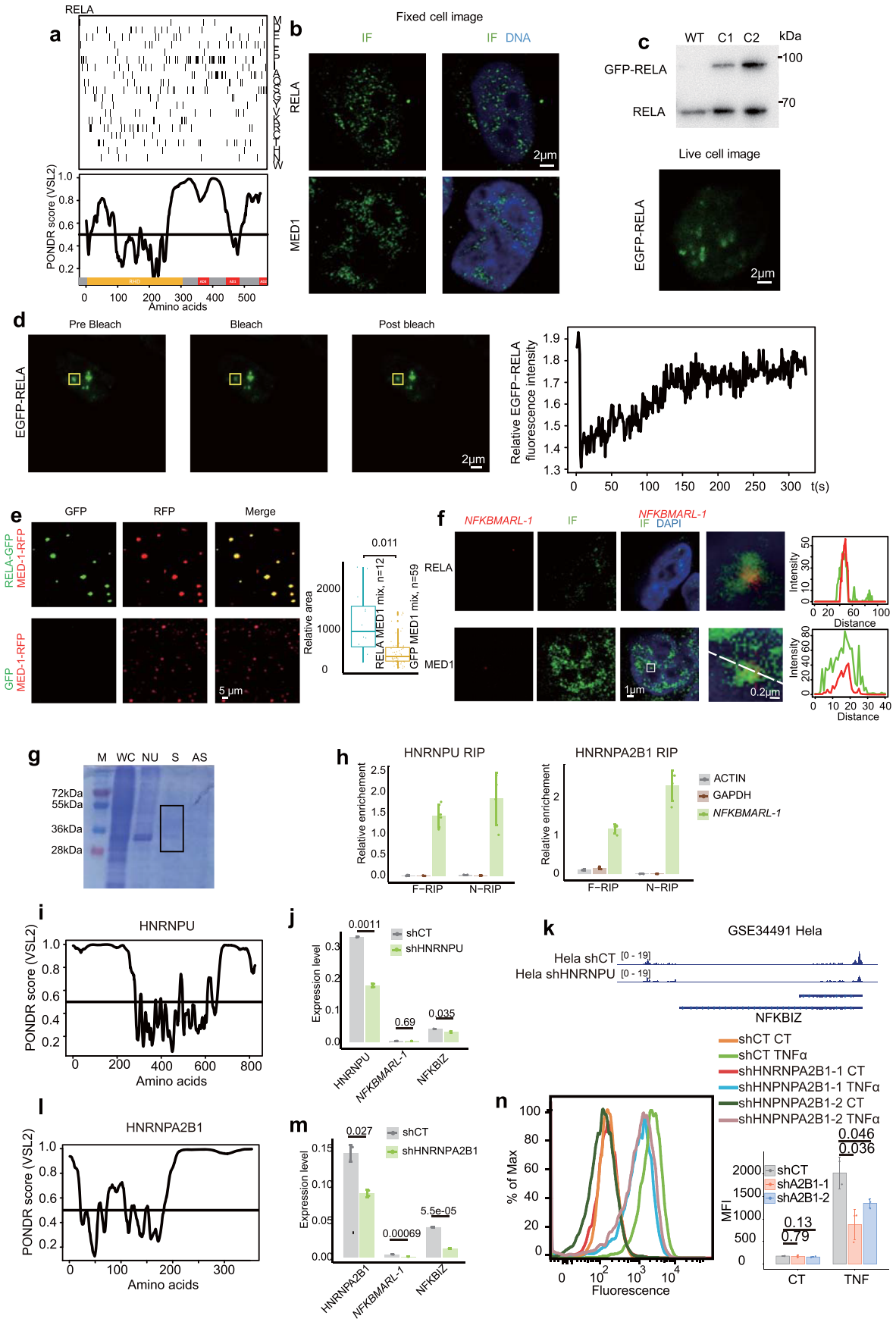


Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | *NFKBMARL-1* regulates NFKBIZ expression in cis. a, Gel image of PCR products for validating the genomic deletion of *NFKBMARL-1* in 293 T cells. **b**, Gel image of PCR validates the genomic deletion of *NFKBMARL-1* in THP1 cells. **c**, Gel image of PCR validated insertion of EF1 promoter in 293 T cells. **d**, Immunofluorescence (IF) of NFKBIZ in control and *NFKBMARL-1* KD TNFα stimulated IMR90 cells (yellow circles are identified by DAPI for nucleus). The right panel is the statistics of fluorescent intensity of NFKBIZ within the yellow circle by ImageJ. **e**, Expression level of NFKBIZ, CXCL1, IL6 and IL8 in control and *NFKBMARL-1* KD bleomycin-induced senescent IMR90 cells, revealed by RT-qPCR. **f**, Western blot of IL1B and p16-INK4A in *NFKBMARL-1* KD and bleomycin-induced senescent IMR90 cells. **g**, Image of SAHF dyed by DAPI and BrdU in control and *NFKBMARL-1* KD bleomycin-induced senescent IMR90 cells. **h**, IF images of IL1B in control and *NFKBMARL-1* KD bleomycin-induced senescent IMR90 cells. The left panel of statistics of fluorescent intensity of IL1B in IMR90. Centerline indicates median value; box limits are the interquartile range of 25th and 75th percentiles; whiskers each extend 1.5 times of the interquartile range of maxima and minima; dots are each fluorescent intensity (**d, h**). The results were analyzed by ImageJ. **i**, WT and *NFKBMARL-1* KO THP1 cell line was induced into M0 stage by PMA (100 ng/ml) and induced into M1 stage by LPS (15 ng/ml), the expression level of *NFKBMARL-1*, NFKBIZ, CXCL1, IL6, IL8, CD14 and CSF1R were revealed by RT-qPCR. **j**, Western blot of IL1B in WT and *NFKBMARL-1* KO THP1 cells. **k**, FACS analysis of CD14-FITC in WT and *NFKBMARL-1* KO THP1 cells induced into M0 and M1 macrophage. The MFI were analyzed by FlowJo. **l**, Left, schematic of coculture of M1 stage macrophage from WT and *NFKBMARL-1* KO THP1 with Hela cells. Middle, image of 0.5% crystal violet dyeing the Hela cells. Right, quantification of the colony number of Hela cells. Data in (**e, i, k**) are presented as mean ± s.d. in n = 3 independent experiments. All p values in (**d, e, h, i, k, l**) were calculated using two-sided unpaired Student's t-test.



Extended Data Fig. 9 | The mechanism of *NFκBMARL-1* regulates NFKBIZ. a, Top: *NFκBMARL-1* RNA pull-down assay for RELA, HNRNPU and HNRNPA2B1 in THP1, revealed by western blot. Bottom: RELA binding in tRSA-RNA pull-down for *NFκBMARL-1*, quantified by Western blot. **b**, Native RIP assay by anti-p65 and anti-BRD4 in Bleomycin-induced senescent IMR90, followed by RT-qPCR for *NFκBMARL-1* and NFKBIZ. **c**, The EMSA of 1ug *NFκBMARL-1* binds to serial concentrations of his tag RELA, RELA-N or RELA-C. **d**, Left, 3D modeling the interaction structure of RELA (1NFI) and *NFκBMARL-1*. The docking 3D structure was visualized by Pymol with the interaction region between RNA and protein within 3 Å colored. The docking nucleotides from RNA are blue. The docking amino acid from RELA is shown with the yellow peak region, the red valley region and orange flat region. Right, the interacting nucleotides in each fragment of *NFκBMARL-1*. **e**, 1μg *NFκBMARL-1* F4 binds to serial concentrations of his tag RELA, revealed by EMSA. **f**, Expression quantification of NFKBIZ, CXCL1, IL6 and IL8 by RT-qPCR after mutating F4 or ChIRP peak region, induced senescence by bleomycin. **g**, The genomic interaction from *NFκBMARL-1* and ChIRP binding region to the long and short promoter region of NFKBIZ in bleomycin treated or *NFκBMARL-1* KD IMR90, revealed by 3C-qPCR. The Data in (**b, f**) are presented as mean ± s.d. in n = 3 independent experiments. All p values in (**b, f**) were calculated using two-sided unpaired Student's t-test.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | The evidences of RELA forming droplet foci. a, Top, amino acid composition of RELA. Each row is a kind of amino acid in protein sequence. Bottom, intrinsic disorder score of RELA (PONDR VSL2). The Rel homology domain (RHD) and activation domain (AD) are indicated in the bottom of intrinsic disorder score. **b**, Immunofluorescence (IF) imaging of RELA and MED1 in TNFα stimulated 293 T cells. Fluorescence signal is shown left and merged with DAPI stain shown right. **c**, Live imaging of TNF-α induced endogenous tagged EGFP-p65 293 T cells. **d**, Images of FRAP experiments performed on the TNFα stimulated endogenous tagged EGFP-p65 293 T cell line. Right, quantification of fluorescent intensity of EGFP-p65 after photobleaching. **e**, In vitro droplet formation of MED1-RFP mixed with GFP or RELA-GFP in droplet formation buffer. The statistics of relative area of RELA-GFP. Centerline indicates median value; box limits are the interquartile range of 25th and 75th percentiles; whiskers each extend 1.5 times of the interquartile range of maxima and minima; dots are each relative area. The results were analyzed by ImageJ. **f**, Images of smFISH for *NFKBMARL-1* and IF for RELA and MED1 in TNFα stimulated 293 T. Middle, enlarged selected region of smFISH. Right, quantification of fluorescent signal of the dash line. **g**, Coomassie blue dyed SDS-PAGE gel of proteins from RNA pull-down with biotin labeled *NFKBMARL-1* in IMR90. M represents marker, WC represents whole cell, NU represents nuclear fraction, S represents sense strand pull-down fraction, AS represents antisense strand pull-down fraction. The piece of gel in sense strand pull-down proteins was cut for MS. **h**, RNA IP by anti-HNRNPA2B1 and anti-HNRNPU and reveal the enrichment level of *NFKBMARL-1*, ACTIN and GAPDH by RT-qPCR. **i**, Graph showing intrinsic disorder score of HNRNPU (PONDR VSL2). **j**, Expression of *NFKBMARL-1* and NFKBIZ by RT-qPCR after knocking down HNRNPU by shRNA in TNFα stimulated IMR90. **k**, Genome browser view of NFKBIZ expression after knocking down HNRNPU in Hela RNA-seq. **l**, Graph showing intrinsic disorder score of HNRNPA2B1 (PONDR VSL2). **m**, Expression of *NFKBMARL-1* and NFKBIZ by RT-qPCR after knocking down HNRNPA2B1 by shRNA in TNFα treated IMR90. **n**, Fluorescent intensity of GFP by FACS after knocking down HNRNPA2B1 by shRNA in NFκB reporter 293 T cells. The MFI was analyzed by Flowjo. Data in (**h, j, m, n**) are presented as mean ± s.d. in n = 3 independent experiments. All p values in (**e, h, j, m, n**) were calculated using two-sided unpaired Student's t-test.

nature research

Corresponding author(s): Jing-Dong Jackie Han

Last updated by author(s): Feb 27, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Collection of qPCR data: MxPro by Agilent, version 3000P. Microscope imaging: ZEN Black Edition SP1 by Carl Zeiss, version 8.1. Immunoblot Imaging: Tanon imaging systems, version 5200. Luciferase reporter assays: Synergy Microplate Reader, version H1.
Data analysis	R version 3.3.2 was used to perform general statistical analysis (t-tests, Fisher exact test,etc). Homologous lncRNA identification: BLASTZ (1.04.00); liftOver (LINUX 64-bit); intersectBed (version 2.17.0); twoBitToFa (LINUX 64-bit). For RNAseq data analysis, the following software were used: STAR (version 020201); HTSeq (version 0.11.2); DESeq (version 1.38.0); cufflink (version 2.2.1); CPAT (version 1.2.2); Pfamscan (version 3.0); GSEA (version 2.07). For ChIPseq and ChIRPseq data analysis, the following software were used: STAR(version 020201); Homer (version 4.11.1). Disordered residue enrichment analysis: D2P2 (web server). Modeling: LongTarget (AGPLv3); MPRDock (Web server). For CRISPRseq data analysis, the following software were used: python (version 3.6.0); MAGeCK (version 0.5.5). ImageJ (version 2.0.0) software was used to analyze immunoblotting and image. FlowJo (version 7.6) software was used to analyze the FACS data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The transcriptome, histone modification, transcription factor, genomic structure data from GEO and NODE database are used for analyzing evolutionary conservation, gene expression or epigenetic profile for NFKBMARL-1 or NFKBIZ under different treatments. All the public datasets used in this study are labeled in the figure. The datasets used for evolutionary conservation analysis (Supplementary Figure 1a with 11 species transcriptomes) include GSE80440, GSE101964, GSE108282, GSE66362, GSE107049, GSE85377, GSE52462, GSE66715, GSE123590, GSE106670, OEP001041, GSE75192, GSE106670. The gene expression and epigenetic profiles of NFKBMARL-1 or NFKBIZ under different treatments are from GSE43070, GSE72534, GSE100382, GSM996197, GSE109700, GSE74328, GSE130306, GSE99074, GSE100382, GSE64233, GSE91020, GSE121522, GSE55105, GSE99544, GSE99544, GSE99544. The public database applied in this study are REMAP (http://remap.univ-amu.fr/); DAVID (https://david.ncifcrf.gov/); POSTAR2 (http://lulab.life.tsinghua.edu.cn/postar/).
The sequencing data of CRISPR screening sequencing, ribo- strand-specific RNA-seq and ChIRP-seq have been deposited in the Gene Expression Omnibus (GEO) with accession number GSE167511. Other source data and statistics for corresponding figures are in the supporting information accompany the paper. Any other relevant data or information about this study are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size of all samples have been described in figure legends for all experiments with at least 3 biological replicates were generated per treatment. No statistical methods were used to pre-determine sample sizes but our samples are similar to those reported in previous publication (Marco De cecco et al. 2019. Nature).
Data exclusions	Data was not excluded from analysis.
Replication	The experiments were independently repeated at least twice and leading to the same conclusions.
Randomization	The cell lines were obtained from public source and authenticated, no pre-established selection criteria for the cell were used. When passaging cell or cell were assigned randomly to different well for different treatment (drug-treatment, genetic perturbation).
Blinding	When the NFκB reporter cells were treated by different concentrations of TNFα, the GFP intensity change dramatically to reflect the treatment. And also the senescent cell induction can also be visualized by the reduction of cell number and phenotype compared to proliferating cell. So blinding was not always possible during experimental setup. However, when performing co-culture experiment for M1 macrophage and Hela, the investigators were blinded to quantify the tumor colonies.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

NFKBIZ, Immunoblotting and immunofluorescence, CST, 9244S, dilution: 1:1000.
IL1B, Immunoblotting and immunofluorescence, Proteintech, 16806, dilution: 1:1000.
Tubulin, Immunoblotting, Proteintech, 11224-1-AP, dilution: 1:5000.
RELA, immunofluorescence and Chromatin Immunoprecipitation, CST, 8242S, dilution: 1:100.
MMP3, Immunoblotting, BOSTER, PB0265 , dilution: 1:500.
HNRNPA2B1, Immunoblotting, Proteintech, 14813-1-AP, dilution: 1:1000.
HNRNPU, Immunoblotting, Proteintech, 14599-1-AP, dilution: 1:1000.
TOP1, Immunoblotting, Proteintech, 20705-1-AP, dilution: 1:1000.
DDX5, Immunoblotting, Proteintech, 10804-1-AP, dilution: 1:1000.
NONO, Immunoblotting, Proteintech, 11058-1-AP, dilution: 1:1000.
p16-INK4A, Immunoblotting, Proteintech, 10883-1-AP, dilution: 1:1000.
HRAS, Immunoblotting, Proteintech, 18295-1-AP, dilution: 1:1000.
MED1, immunofluorescence, Bethyl, A300-793A, dilution: 1:200.

Validation

All antibodies are from commercially available sources and have been validated from the manufacturer with supporting publications found on manufacturer's websites.
NFKBIZ, Immunoblotting and immunofluorescence, CST, 9244S, (https://www.cellsignal.com/products/primary-antibodies/ikb-z-antibody/9244?site-search-type=Products&N=4294956287&Ntt=9244s&fromPage=plp&_requestid=415959).
IL1B, Immunoblotting and immunofluorescence, Proteintech, 16806, (https://www.ptglab.com/products/IL1B-Antibody-16806-1-AP.htm).
Tubulin, Immunoblotting, Proteintech, 11224-1-AP, (https://www.ptglab.com/products/TUBA1B-Antibody-11224-1-AP.htm).
RELA, immunofluorescence and Chromatin Immunoprecipitation, CST, 8242S, (https://www.cellsignal.com/products/primary-antibodies/nf-kb-p65-d14e12-xp-rabbit-mab/8242?site-search-type=Products&N=4294956287&Ntt=8242s&fromPage=plp&_requestid=416274).
MMP3, Immunoblotting, BOSTER, PB0265 , (http://www.boster.com.cn/product/anti-mmp3-antibody_pb0265.html).
HNRNPA2B1, Immunoblotting, Proteintech, 14813-1-AP, (https://www.ptglab.com/products/HNRNPA2B1-Antibody-14813-1-AP.htm).
HNRNPU, Immunoblotting, Proteintech, 14599-1-AP, (https://www.ptglab.com/products/HNRNPU-Antibody-14599-1-AP.htm).
TOP1, Immunoblotting, Proteintech, 20705-1-AP, (https://www.ptglab.com/products/TOP1-Antibody-20705-1-AP.htm).
DDX5, Immunoblotting, Proteintech, 10804-1-AP, (https://www.ptglab.com/products/DDX5,p68-Antibody-10804-1-AP.htm).
NONO, Immunoblotting, Proteintech, 11058-1-AP, (https://www.ptglab.com/products/NONO-Antibody-11058-1-AP.htm)
p16-INK4A, Immunoblotting, Proteintech, 10883-1-AP, (https://www.ptglab.com/products/P16,P19-Antibody-10883-1-AP.htm).
HRAS, Immunoblotting, Proteintech, 18295-1-AP, (https://www.ptglab.com/products/HRAS-Specific-Antibody-18295-1-AP.htm).
MED1, immunofluorescence, Bethyl, A300-793A, (https://www.bethyl.com/product/A300-793A/MED1+Antibody).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

IMR90 cells were obtained from ATCC (CCL-186). 293T and THP1 cells were from Cell Bank, Chinese academy of sciences.

Authentication

IMR90, 293T and THP1 cell line were obtained from the sources above and used at low passage. They were not further authenticated.

Mycoplasma contamination

All cells were kept in a humidified incubator at 37°C and 5% CO2 and test negative of cell line routinely for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

None of these cell lines are listed in the ICLAC database.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Cells were maintained in culture and split as needed to ensure the confluence did not exceed 90%. The pgRNA libraries were screened 7 days post-infection. After stimulated by 10ng/ml TNF, the cells were selected for GFP high and GFP low by FACS, for which cells were harvested, resuspend at 10 million cells/ml, and RFP-positive cells were sorted (BD ARIA II) for lower GFP and upper GFP.

Instrument	BD ARIA II
Software	Flow cytometry data was analyzed using FlowJo (v7.6)
Cell population abundance	The cell population was identified by the GFP signal as GFP high or GFP low for the top or bottom 25%.
Gating strategy	FSC-H vs FSC-A was used in the gate to determine the singlets. The count plot of GFP signal was used to isolate the cell population. The gating strategy was presented in Supplementary Fig. 4.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.