nature communications

Article

https://doi.org/10.1038/s41467-024-49106-2

Progressive senescence programs induce intrinsic vulnerability to aging-related female breast cancer

Received: 16 June 2023

Accepted: 24 May 2024

Published online: 17 June 2024

Check for updates

Huiru Bai[1,2,3,6], Xiaoqin Liu[1,2,3,6], Meizhen Lin[1,2,3], Yuan Meng[1,2,3], Ruolan Tang[1,2,3], Yajing Guo[1,2,3], Nan Li[4], Michael F. Clarke[5] & Shang Cai[1,2,3]

Cancer incidence escalates exponentially with advancing age; however, the underlying mechanism remains unclear. In this study, we build a chronological molecular clock at single-cell transcription level with a mammary stem cell-enriched population to depict physiological aging dynamics in female mice. We find that the mammary aging process is asynchronous and progressive, initiated by an early senescence program, succeeded by an entropic late senescence program with elevated cancer associated pathways, vulnerable to cancer predisposition. The transition towards senescence program is governed by a stem cell factor *Bcl11b*, loss of which accelerates mammary ageing with enhanced DMBA-induced tumor formation. We have identified a drug TPCA-1 that can rejuvenate mammary cells and significantly reduce aging-related cancer incidence. Our findings establish a molecular portrait of progressive mammary cell aging and elucidate the transcriptional regulatory network bridging mammary aging and cancer predisposition, which has potential implications for the management of cancer prevalence in the aged.

Aging has been recognized as the greatest risk factor for the vast majority of cancer types. As a significant extension of global lifespan, the burden of cancer incidence and cancer mortality have been rapidly increasing as major challenges to human health worldwide[1,2]. Despite extensive advances in aging studies at the molecular, cellular and organismal levels, the exponential association between cancer occurrence and age[3–5] has persisted for years, and the underlying biology of this etiological phenotype remains largely unclear. Understanding how physiological aging programs impact carcinogenesis is therefore of particular interest and an imperative research direction for preventing cancer prevalence.

In contrast to the uncontrolled proliferation of cancer cells, in aging tissue, a senescence program is frequently activated, which suppresses cell proliferation, through the p53-p21 and p16^{Ink4a}-Rb pathways[6,7] and is accompanied by activated β-galactosidase[8], global alteration of H3K9me3 abundance[9,10], induction of senescence-associated secretory phenotype (SASP) factor activity[11] and disturbed immune ecosystem[12]. In general, these pathways impair long-term stem cell self-renewal ability[13], induce chronic inflammation[14] and interfere with tissue homeostasis[15], ultimately leading to the acquisition of degenerative phenotypes. The prevailing idea to explain aging-caused cancer involves accumulation of mutations during aging, which promotes the transformation of cells into cancer cells[16–19]; however, the degree to which this transformation drives cancer is not clear[20], largely because concordance between mutation rates and cancer profiles for aged individuals is lacking and because increased longevity is associated with lower cancer incidence[3,21]. Alternatively, aging tissue may secrete a plethora of SASP factors, forming a fertile environment

[1]Westlake Disease Modeling lab, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China. [2]School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China. [3]Key Laboratory of Growth Regulation and Translational Research of Zhejiang Province, Westlake University, Hangzhou, Zhejiang, China. [4]Westlake University High-Performance Computing Center, Westlake University, Hangzhou, Zhejiang, China. [5]Institute of Stem Cell and Regenerative Medicine, School of Medicine, Stanford University, Palo Alto, CA, USA. [6]These authors contributed equally: Huiru Bai, Xiaoqin Liu.
e-mail: caishang@westlake.edu.cn

for neighboring cells to promote cancer initiation[22]. Intriguingly, recent discoveries have indicated that the stemness program can be triggered by senescence during embryonic development[23,24], wound healing[25–29], and drug treatment in cancer[30]. Whether this reprogramming underpins the physiological aging process and whether it contributes to cancer initiation remain unclear. Understanding chronological aging dynamics under physiological conditions and the underlying driving force mediating aging are crucial in establishing a biological connection between aging and cancer.

In this study, we address the potential connection between aging and cancer by building a chronological transcriptome map at the single-cell level with a stem cell-enriched mammary population in mice of various ages (from 2 to 29 months). We identify heterogeneous cell states in mice at each individual age with distinct senescence programs (early or late) vulnerable to breast cancer predisposition. In addition, we identify a master transcription factor, *Bcl11b*, that comprehensively suppresses both early and late senescence programs and find that loss of *Bcl11b* expression dramatically accelerates aging and tumor formation. Reversing the senescence program by TPCA-1 treatment efficiently reduces the cancer incidence and extends the cancer latency time. Our study establishes a molecular aging trajectory for mouse mammary cells and reveals an intrinsic molecular link between aging and cancer, which may shed light on preventive strategies against breast cancer occurrence in the future.

Results

A chronological single-cell transcriptome analysis reveals asynchronous dynamics of a mammary stem cell-enriched population during aging

Recently, single-cell studies have revealed hallmarks of aging in mammary gland cells at discrete ages, defined as young and old[31–35]; however, the chronological aging process of mammary gland cells has not been well documented. Because initial oncogenic mutations occur in long-lived stem cells/progenitors[36,37], we were particularly interested in the molecular progression of aging in a mammary stem cell-enriched population. We therefore sorted mammary stem cell-enriched populations ($CD49f^{high}EpCAM^{low}$)[38,39], which included all the mammary stem cells identified to date[40–43], from mice of various biological ages spanning from 2 months to 29 months and performed single-cell RNA sequencing (RNA-seq) with 3′ UTR Smartseq[44–46] (Fig. 1a; Supplementary Fig. 1a; Supplementary Data 1). We extracted transcriptome information of 1981 cells from 25 mice (1–4 mice at each age) and detected 20629 genes. A subsequent t-distributed stochastic neighbor embedding (t-SNE) analysis revealed that the vast majority of the mammary cells obtained from mice of various ages were thoroughly intermingled (Fig. 1b). These cells shared a similar transcriptome and uniformly expressed *Keratin14* and *Keratin5* across different age groups (Fig. 1c, Supplementary Fig. 1h), suggesting that they share the same fate.

To build a molecular clock and thus gauge dynamic transcriptomic changes with age, we performed a trajectory analysis with Monocle 2 and reconstructed a linear pseudotime ordering of mammary cells at different mouse ages. Remarkably, the mammary cells at different mouse ages clearly followed a chronological order, with the cells isolated from younger mice aligning with the early pseudotime stage and the cells isolated from older mice aligning with the later pseudotime stage (Fig. 1d). This finding indicates that an age-related transcriptome program defines the intrinsic cell state. Indeed, when we clustered the differentially expressed genes on the basis of the pseudotime, the signature genes in the mammary cells were classified into four different states with distinct gene expression patterns (Fig. 1e, f and Supplementary Fig. 1c). Interestingly, the mammary cells of each individual mouse comprised all four-state cells, with their relative abundance being the only difference (Fig. 1g, h and Supplementary Fig. 1b). We then quantified the relative abundance of the cells in each cell state throughout the aging process and found that the number of State 1 cells kept decreasing with age and that of the State 4 cells showing an increasing trend (State 1: $p = 0.0052$; State 2: $p = 0.73$; State 3: $p = 0.72$; State 4: $p = 0.05$). The number of state 2 cells remained relatively constant, while that of State 3 cells temporarily increased and then decreased (Fig. 1i and Supplementary Fig. 1b, d–g). Even in the individual mice of same age, we observed variation of relative abundance of the four cell states, suggesting that the ageing of mice was not synchronous. But when they were analyzed as age groups, the trend was clear. These data suggest that the mammary $CD49^{hi}EpCAM^{low}$ cells are heterogeneous with distinct cell states and that the aging phenotype may be manifested by the relative abundance of various cell states.

Distinct senescent mammary cell states induce progressive aging dynamics

To further characterize the biological identities of each individual cell state, we performed pathway analysis. We plotted the activity of each signaling pathway over pseudotime to visualize the chronological dynamics, and we identified six distinct dynamic patterns (Fig. 2a). Pattern 1 pathways exhibited the highest activity in State 1, gradually declined throughout the entire time course to the last state. These pathways included ‘DNA replication’, ‘mismatch repair’, ‘oxidative phosphorylation’, ‘beta-alanine metabolism’ and ‘valine, leucine and isoleucine degradation’. The decreased activity of ‘DNA replication’ and ‘mismatch repair’ with increased pseudotime aligned with the notion that DNA mutations accumulate during aging[47,48]. In addition, this finding indicated that State 1 cells are younger cells with higher DNA repair ability and metabolic activity. Consistent with Pattern 1, Pattern 2 pathways showed a transient increase in activity during the State1-2 transition, followed by a rapid decline. Pattern 2 exhibited only one activated pathway, ‘mitochondria ribosome’. The low pathway enrichment level may suggest low transcriptional activity and a relatively quiescent state. Indeed, when we plotted metabolism-related pathways, we found that State 2 cells exhibited relatively low metabolic lipid, amino acid and carbon activity but higher TGFbeta signaling pathway activity (Fig. 2b) than State 1 cells. In addition, a cell cycle analysis[49] suggested that State 2 cells contain a relatively high portion of cells in the G0 phase compared to that of the other cells (Fig. 2c, d). Given the low metabolism in State 2 cells, we labeled these cells quiescent young (q-Young) cells, and we labeled State 1 cells active young (a-Young) cells.

Patterns 3–5 shared a similar increase in activity with nuanced differences. The activity of Pattern 3 pathways, including the ‘NF-kappaB signaling pathway’, ‘p53 signaling pathway’, ‘HIF1 signaling pathway’ and ‘ferroptosis’, peaked in State 3 and then quickly decreased in State 4 (Fig. 2a). Notably, activation of the NF-kB and p53 pathways correlated with aging phenotypes[50,51]. Pattern 4 pathways were activated in State 3 and then reached a plateau level that persisted into State 4. This pattern included ‘cellular senescence’, ‘cytokine–cytokine receptor interaction’ and ‘Toll-like receptor signaling pathway’ (Fig. 2a). This finding suggests that both State 3 and State 4 cells had an activated senescence program. In accordance with this supposition, the expression of SASP program genes, a hallmark of senescent cells, was elevated in both State 3 and State 4 cells (Fig. 2e–f, Supplementary Fig. 3a-b). Pattern 5 included pathways continuously increased activity in both States 3 and 4. These pathways included ‘IL-17 signaling’, ‘JAK-STAT signaling’, ‘mTOR signaling’, ‘PI3K-Akt signaling’, ‘MAPK signaling’, ‘Ras signaling’, ‘TGFbeta signaling’, ‘breast cancer’ and, most notably, ‘signaling pathways regulating pluripotency of stem cells’ (Fig. 2a). The activation of these gene pathways suggests that State 4 cells, despite expressing the senescence program, had acquired stem cell traits and cell growth/survival programs, which may have predisposed them to a precancerous phenotype. We therefore named State 3 cells early senescence (e-Sen) cells, while State 4 cells

https://doi.org/10.1038/s41467-024-49106-2

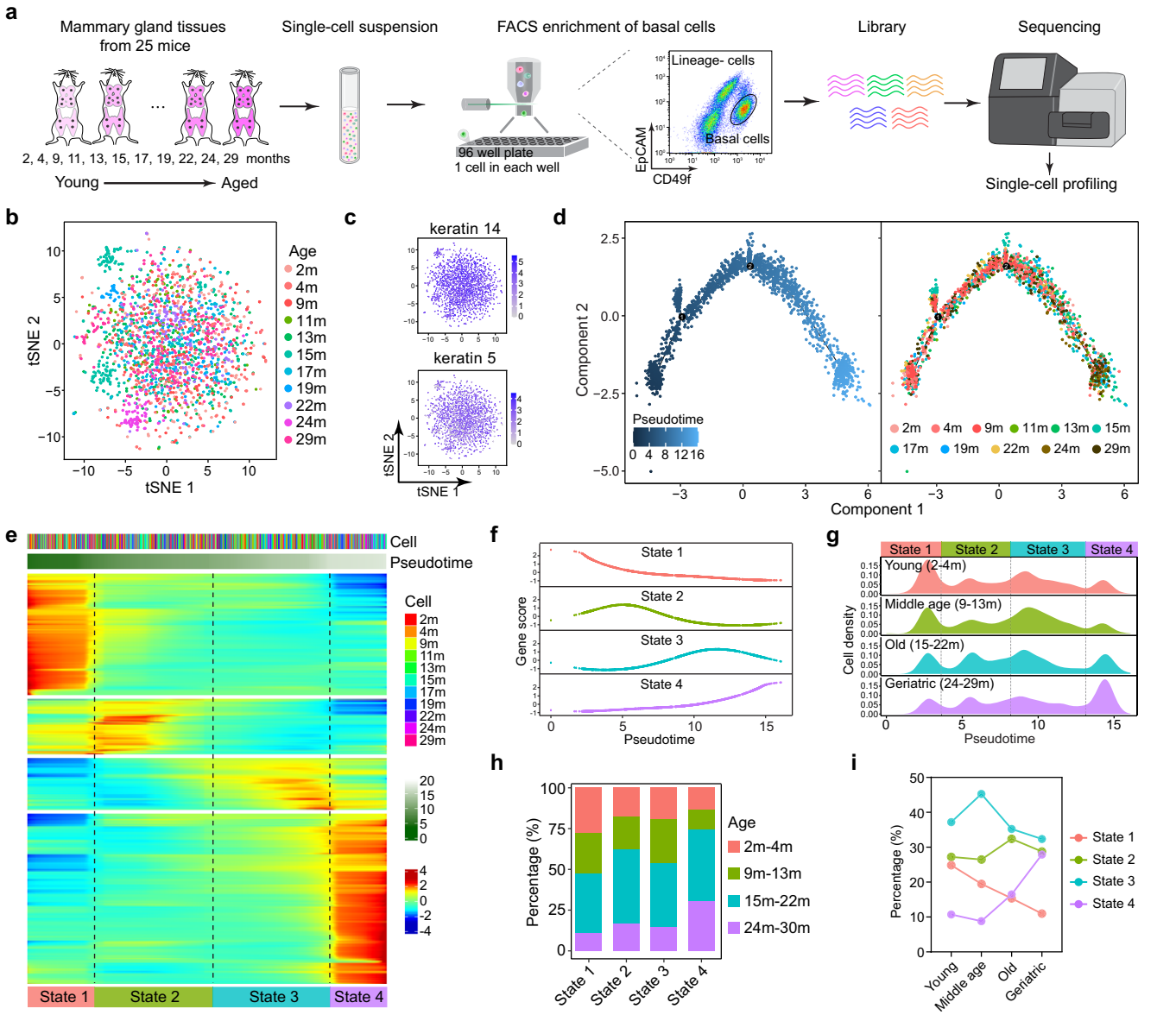

Fig. 1 | scRNAseq profiling of mammary stem cell enriched population at various chronological ages of mice. a Schematic diagram showing the pipeline of scRNA-seq of mouse mammary cells. Dissociated mammary cells from 2 month to 29 month old mice were sorted for CD49f^{high}EpCAM^{low} mammary stem cell enriched population, followed by 3′UTR SMARTseq for library construction and subsequent sequencing. **b** *t*-SNE plot showing the clustering of mammary cells originating from 2 m ($n = 2$), 4 m ($n = 3$), 9 m ($n = 2$), 11 m ($n = 2$), 13 m ($n = 2$), 15 m ($n = 4$), 17 m ($n = 2$), 19 m ($n = 2$), 22 m ($n = 2$), 24 m ($n = 1$) and 29 m ($n = 3$) mice. **c** Basal cell specific genes (*Keratin 5* and *Keratin 14*) on *t*-SNEs showing a uniform expression pattern. **d** Pseudotemporal ordering analysis of single cell transcriptomes by Monocle 2 inferring the mammary ageing trajectory. Cell origins are labeled by distinct colors. **e** Heatmap visualization of the dynamic gene expression over the pseudotime. Cells were divided into four states based on the differentially expressed gene clusters. **f** Expression pattern of the signature gene clusters for each cell state along the pseudotime. **g** Cell density map showing the distribution of mammary cells from young (2 m-4 m, $n = 5$), middle age (9 m–13 m, $n = 6$), old (15 m–22 m, $n = 10$) and geriatric (24 m-29 m, $n = 4$) mice along with pseudotime. **h** Relative cell proportion of each mouse age in each cell state. **i** Relative cell abundance of the four mammary cell states in each age group.

were called late senescence (l-Sen) cells. Pattern 6 contained pathways activated later, between State 3 and State 4, indicating that they might be functionally involved in driving the transition from the e-Sen phenotype to the l-Sen phenotype (Fig. 2g). These pathways included 'Hedgehog signaling', 'Notch signaling' and 'Wnt signaling' (Fig. 2a). All of the six pathway patterns dynamics remained the same in different age groups (Supplementary Fig. 2), suggesting that the dynamics were intrinsic to the cell state irrespective of their biological age. Some of the signaling pathways were confirmed by immunohistochemical (IHC) staining (Fig. 2h–k and Supplementary Fig. 3c-f).

These intriguing distinct senescence programs suggest that senescent cells are heterogeneous and that physiological aging progresses in a sophisticated manner with altered homeostasis among four different cellular states, not through a young-old binary switch. This unique physiological aging process is consistent with the in vitro senescence dynamics induced by oncogenes[52,53], as well as the aberrant activation of senescence and stem cell programs during embryogenesis[23,24], wound healing[26,28] and cancer drug treatment[30], indicating a pervasive underlying mechanism.

Breast cancer initiation is associated with l-Sen program

l-Sen cells exhibited aberrantly activated cancer- and stem cell-related programs, and have reduced P53 activity and enhanced PI3K-Akt activity. Considering that P53 and PIK3CA are the two most prominent

https://doi.org/10.1038/s41467-024-49106-2

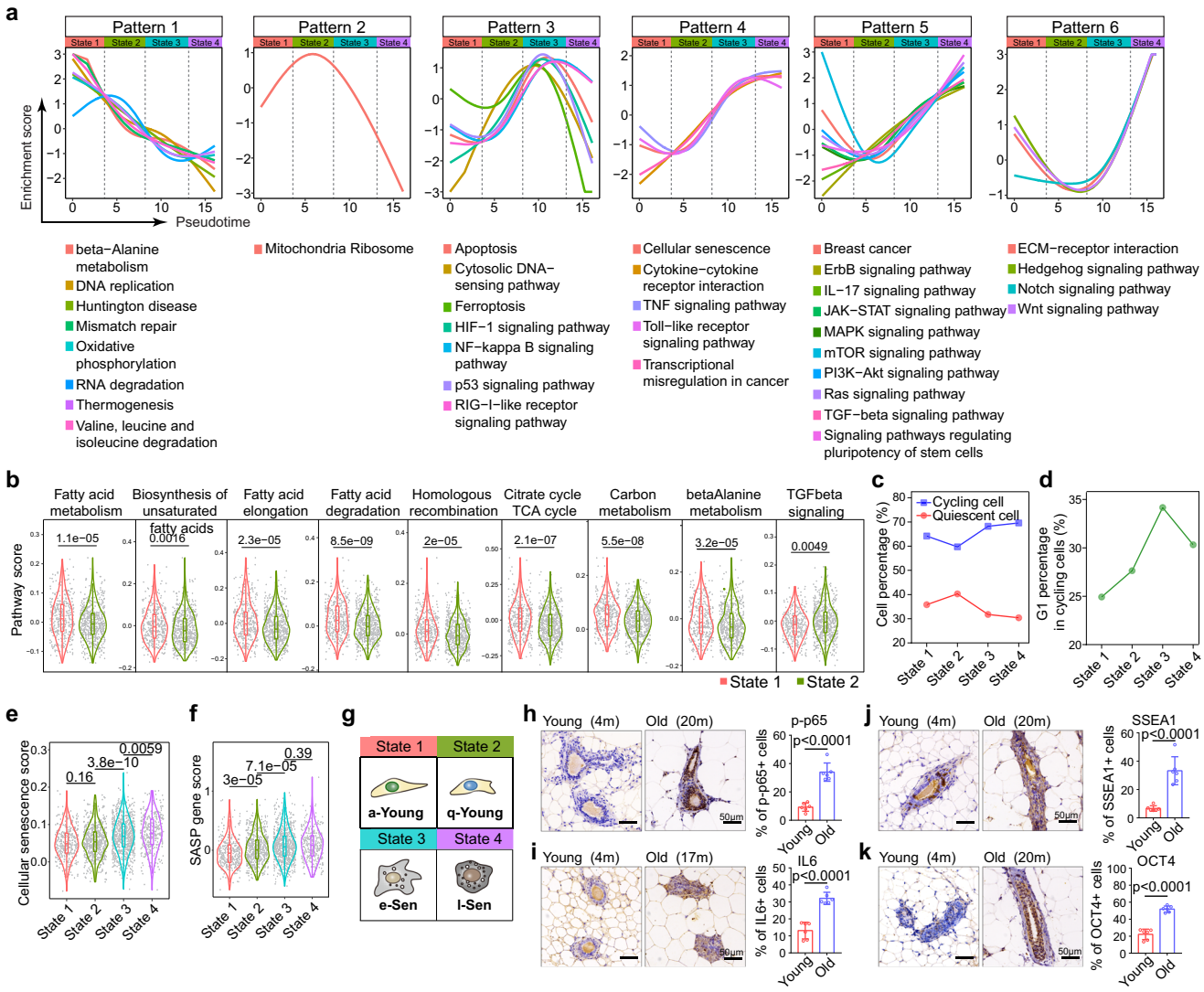

Fig. 2 | Characterization of distinct mammary cell states along the ageing trajectory. a Activity of each signaling pathway presented by enrichment score over the pseudotime. Note: six distinct dynamic patterns were identified indicating the identity of each cell state. **b** Differential pathway activities between state 1 ($n = 344$ cells) and state 2 ($n = 588$ cells) cells. Statistical analysis was performed using two-sided wilcoxon test. **c** Cell cycle analysis showing the relative fractions of cycling cells (blue) and quiescent cells (red) in each cell state. **d** Relative proportion of G1-phase cells in each cell state. **e**, **f** Senescence related pathway analysis showing the cellular senescence level (**e**) and SASP expression level (**f**) in each cell state. There are 344, 588, 740 and 309 cells in State1, State2, State3 and State4 separately. Statistical analysis was performed using two-sided wilcoxon test. **g** Schematic diagram of the four cell states designated as active young state (a-Young), quiescent young state (q-Young), early senescence (e-Sen) and late senescence (l-Sen). **h**–**k** Representative images of immunohistochemistry staining of p-p65 (**h**), IL6 (**i**), SSEA1 (**j**), OCT4 (**k**) in young and old mammary tissues. Values are means ± SD ($n = 6$ biological replicates); p, two-tailed unpaired t-test.

mutation genes in breast cancer[54], we speculate that l-Sen cells have increased their vulnerability toward cancer transformation. This prompted us to ask, do these programs predispose cells to a precancer state? We therefore analyzed the paired human breast samples (tumor and tumor adjacent normal tissue) in TCGA database for pathway activity and transcription factor activity (Fig. 3a). Interestingly, compared to the adjacent normal tissue, breast tumors were significantly elevated for various senescence related pathways, and the l-Sen program related pathways, including Notch signaling, Wnt signaling (TCF7L1, LEF1), Hedgehog signaling (GLI1, GLI2), and pluripotency related factors (MYC, SOX2 and KLF4), while e-Sen specific NFkB pathway was tuned down (Fig. 3b and Supplementary Fig. 4a-c). This suggests that breast tumor tissues are closely associated with l-Sen signature.

To ask whether the l-Sen program is turned on at the initiation stage of breast tumor, we used a dimethylbenz(a)anthracene (DMBA)-mediated breast cancer development model[55,56], which was shown to predominantly trigger breast tumors[57], and analyzed tumor initiation foci at very early stage (Fig. 3c). Consistent with human breast tumors, the senescence related signals and pluripotency related signals were all upregulated at the onset of tumor formation (Fig. 3d–h). Meanwhile, we found that DMBA, besides its function causing genetic mutations, also triggered mammary cell senescence with a prominent l-Sen cell expansion and senescence related pathway activation, suggesting the tumor initiation process is accompanied with l-Sen program activation (Fig. 3i–n). Consistent with this idea, we found that mammary cells colony formation ability significantly decreased from young to old mice, but recovered in the geriatric mice where they have significant expansion of l-Sen population (Fig. 3o, p).

To further characterize the kinetics of the physiological aging process, we employed a single-cell signaling entropy algorithm[58] to profile the dynamics of cellular entropy (Fig. 3q, r). The cellular entropy of a-Young, q-Young and e-Sen cells remained at a relatively low level, with a slight decrease from the a-Young to the q-Young

https://doi.org/10.1038/s41467-024-49106-2

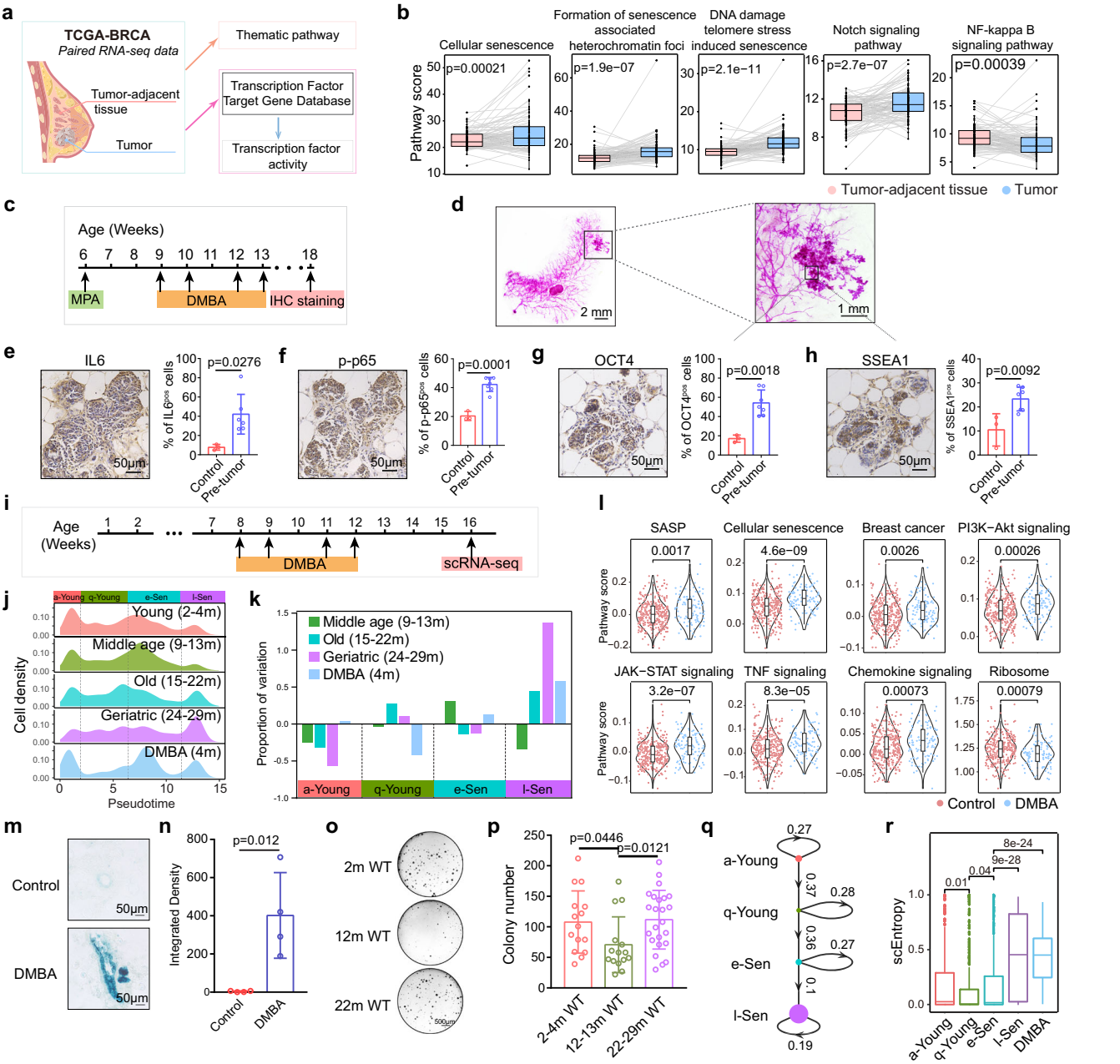

Fig. 3 | Senescent cells are vulnerable for cancer predisposition. a-**b** Diagram showing the workflow of the analysis for thematic pathway score and transcription factor activity score in human mammary gland tumor and tumor-adjacent tissue from TCGA database. $n = 112$ patients; p, two-tailed paired t-tests. **c** Schematic diagram showing the DMBA induced cancer assay for 6 week-old WT mice. **d** Representative wholemount staining images for DMBA-induced tumors in mammary gland. **e**–**h** Representative IL6 (**e**), p-p65 (**f**), OCT4 (**g**), SSEA1 (**h**) IHC staining images for DMBA-induced tumors in mammary gland. Values are means ± SD. Control: $n = 3$; Pre-tumor: $n = 6$ for IL6; $n = 7$ for p-p65, OCT4, SSEA1; n represents biological replicates; p, two-tailed unpaired t-test. **i** Diagram showing that 8 week-old mice were treated with DMBA (200 μL, 5 mg/mL) and analyzed by scRNA-seq. **j** Density map of mammary cell states in each age group and DMBA treated group. **k** Changes of the relative abundance of cell states in different age groups compared to the Young (2–4 m) group. **l** Senescence related pathway score in CD49f^{high}EpCAM^{low} cells of DMBA treated mammary gland. 236 cells from 3 control mice, 109 cells from 2 DMBA treated mice. Two-sided wilcoxon-test was used. **m**, **n**, Representative β-gal staining images (**m**) and quantification (**n**) of mammary gland from control (4 month, $n = 4$ mice) and DMBA treated mice (4 month, $n = 4$ mice) showing senescent cells. Statistical analysis was performed using two-tailed unpaired t-tests; data are presented as mean ± SD. **o**-**p** Representative images (**o**) and quantification (**p**) of colony formation ability of CD49f^{high}EpCAM^{low} mammary cells in 2–4 m WT ($n = 5$ mice), 12-13 m WT ($n = 5$ mice), 22-29 m WT ($n = 8$ mice) mammary gland. Statistical analysis was performed using two-tailed unpaired t-tests; 3 technical replicates for each mice; data are presented as mean ± SD. **q** A linear lineage trajectory shows the transition probabilities for the four cell states with the node size corresponding to the signaling entropy. **r** scEntropy analysis of the four cell states. Cell number: 344 (a-Young), 588 (q-Young), 740 (e-Sen), 309 (-Sen), 109 (DMBA group). Statistical analysis was performed using two-sided wilcoxon test and BH adjusted p-value was used.

cells ($p < 0.01$) and an increase from the q-Young to the e-Sen cells ($p < 0.05$). Remarkably, the entropy of the l-Sen cells was strikingly elevated compared with that of the e-Sen cells ($p < 0.001$), indicating a drastic systemic disorder and a potential chromatin reorganization[59]. The increase in entropy suggests that e-Sen cells transition to the l-Sen state in a passive spontaneous manner in the absence of extrinsic inputs. Therefore, we speculated that mammary aging might be initiated and determined during the early q-Young-to-e-Sen transition, which is crucial for subsequent l-Sen state commitment.

The q-Young-to-e-Sen cell transition is mediated by the mammary stem cell factor *Bcl11b*

To determine the factors that drive the transition from the q-Young to the e-Sen state, we constructed a limited gene regulation network using the chromatin immunoprecipitation followed by sequencing (ChIP-seq) ENCODE and ChEA databases, along with the text curated database TRRUST of transcriptional regulatory networks, and we superimposed our network onto the expression matrix of the four aging-cell states to infer the key transcription factors in each cell state (Fig. 4a; Supplementary Data 2). The highest fidelity factors supported by multiple databases were selected (Fig. 4b).

With a transcriptional inference algorithm, we pinpointed 38 state-specific fate determinants in total (Fig. 4b), of which the majority were found in l-Sen state cells (Supplementary Fig. 4d). Interestingly, the factors in young state cells, such as *Hcfc1* and *Srf*, have been previously associated with longevity[60,61]. In the l-Sen state cells, we found that factors such as *Cebpb*, *Cebpd* and *Cebpz* were associated with SASP secretion[52,62]; *Stat3*, which is the downstream effector in the Jak-stat pathway; *Tcf3* and *Tcf7l2*, which are key effectors in the Wnt signaling pathway; and *Kdm5b*, which is a histone demethylase that contributes to Rb-mediated cellular senescence[63]. Most interestingly, we observed *Nanog* and *Pou5f1* which were associated with pluripotent stem cells. These factors were consistent with the pathway analysis, reinforcing the idea that in l-Sen state cells, the stem cell program is aberrantly activated.

In identifying the factors critical for the initiation of the e-Sen cell state, we concentrated our attention on e-Sen transcription factors. The activity plot showed that *Tcf7l2*, *Rnf2*, *Fosl1*, *Gata3* and *Hsf1* were all increasing their activity during the q-Young-to-e-Sen transition, with the exception that *Bcl11b* sharply decreased its activity (Fig. 4c; Supplementary Data 3). *Bcl11b* is a previously identified modulator of mammary stem cell self-renewal and quiescence, the loss of which results in stem cell exhaustion[41]. As *Bcl11b* is a predominant transcriptional repressor[64–66], a decreased *Bcl11b* activity is a reflection of the activation of *Bcl11b*-repressed targets. This finding suggests that the initiation of early senescence may be mediated by the loss of *Bcl11b* function.

Accelerated mammary ageing phenotypes was triggered by loss of *Bcl11b*

We then wondered whether *Bcl11b* is a key factor mediating the mammary senescence switch. To determine the mechanism of *Bcl11b* function, *Bcl11b* expression was knocked out in mammary glands, and the *Bcl11b*-knockout (KO) mammary tissues exhibited reduced ductal width (Supplementary Fig. 5d–h), enhanced β-galactosidase activity (Fig. 4d), and elevated $p16^{Ink4a}$ (Fig. 4e and Supplementary Fig. 5a), which are typical markers of senescent cells[7,8]. In multiple aging organs, the stem cell population has been frequently shown to expand with declining functionality[67–70]. Specific to aging mammary glands, a phenotype of an increased basal percentage has been reported to be acquired with age[71], which is consistent with our observation of the mammary gland as the mice aged (Fig. 4f and Supplementary Fig. 5b). We analyzed *Bcl11b*-KO cells in a *Krt14-cre* $Bcl11b^{fl/fl}$ mT/mG mouse, in which the green fluorescent cells represented *Bcl11b*-KO cells, and the red fluorescent cells were the *Bcl11b* WT control cells (Supplementary Fig. 5c). We found that, compared with the control cells in the same mouse, the *Bcl11b*-KO cells exhibited significant increase of basal percentage at a young age (Fig. 4g). Given that our previous data showed that knocking out *Bcl11b* expression significantly reduced mammary stem cell self-renewal ability and promoted stem cell exhaustion[41], consistent with aging mammary stem cells (Supplementary Fig. 5i-j), these apparently expanded basal cells may undergo functional decline. Overall, these data collectively suggest that loss of *Bcl11b* function triggered an accelerated aging process.

Ageing has usually been defined on the basis of certain biomarkers or functional assays; however, these methods are of limited value when aging cells are heterogeneous. Because we built a molecular clock of mammary cell aging, we tried to use a chronological map to gauge the aging grades of the *Bcl11b*-KO cells. When we included the *Bcl11b*-KO cells in the aging pseudotime analysis and reconstructed the trajectory, we found that the vast majority of the *Bcl11b*-KO cells spontaneously accumulated at the later stage and coclustered with l-Sen cells (Supplementary Fig. 6), with a notably diminished q-Young peak (Fig. 4h, $p < 0.000001$). Hence, quantification of each aging stage revealed that the number of a-Young and q-Young cells was drastically reduced compared with that of the age-matched wild-type (WT) cells, while the number of e-Sen and l-Sen cells was profoundly increased (Fig. 4h, i and Supplementary Fig. 6 f). This finding suggested that *Bcl11b*-KO underwent substantially accelerated aging progression, and cells rapidly entered a state very similar to that of senescent cells.

To determine which aging-related molecular pathways were altered by knocking out *Bcl11b*, we performed a pathway analysis (Fig. 4j and Supplementary Fig. 7; Supplementary Data 4). The Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis showed that the activity levels of the typical aging-related pathways, including 'MAPK signaling', 'PI3K-Akt signaling', 'IL17 signaling', 'cellular senescence', 'NF-kB signaling', and 'p53 signaling', were all significantly upregulated in the *Bcl11b*-KO cells, indicating a systemic aging state. In addition, the senescence-related SASP program was also elevated in the *Bcl11b*-KO cells, which was accompanied by increased cellular entropy (Fig. 4k, l). Overall, these data demonstrate that *Bcl11b* may be a molecular switch regulating the q-Young-to-e-Sen transition, ultimately promoting the l-Sen chaos.

Senescent mammary cells induced by Bcl11b KO are susceptible to cancer transformation

Next, we try to address whether the accelerated mammary ageing is intrinsically coupled with cancer in the DMBA based treatment cancer initiation model. In this tumor model, we found that the DMBA induced tumor cells predominantly originated from Krt14+ basal cells (Supplementary Fig. 8a–c), and the mammary cells showed prominent age-related susceptibility to chemical induced cancer transformation (Fig. 5a–c). We therefore determined to address whether the senescent mammary cells caused by *Bcl11b* loss are vulnerable to cancer transformation with the DMBA-induced breast tumor formation assay for WT and *Bcl11b*-KO cells. We tried to prevent microenvironmental factors from influencing the results by transplanting control and *Bcl11b*-KO cells into recipient mice at a saturated dose and then performed DMBA treatment after full mammary tree formation (Fig. 5d). Intriguingly, the mice that received the *Bcl11b*-KO cells acquired tumors much earlier and faster than the control recipient mice, ultimately exhibiting a much higher cancer incidence with age (Fig. 5e, f). We further characterized the WT and *Bcl11b*-KO tumors by scRNA-seq to address the possibility that the tumors originated from cells neighboring senescent cells (Supplementary Fig. 8d). We reasoned that if neighboring cells give rise to tumors, then WT and *Bcl11b* tumor cells likely have a similar phenotype. Interestingly, a gene set enrichment assay

https://doi.org/10.1038/s41467-024-49106-2

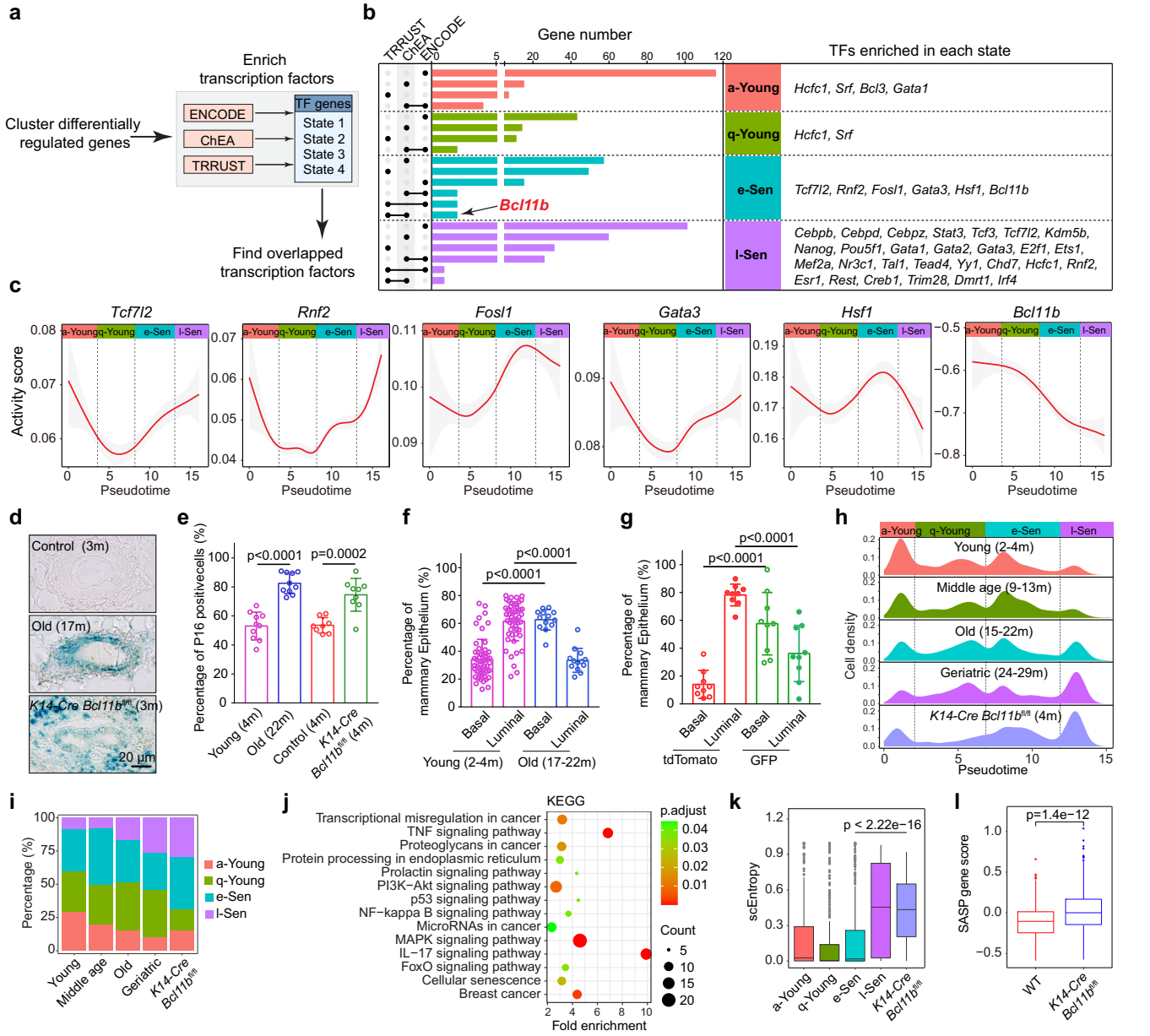

Fig. 4 | q-Young to e-Sen transition is functionally mediated by transcription factor *Bcl11b*. **a** Diagram showing the workflow of the transcription factors enrichment analysis for each cell state using the ENCODE, ChEA, and TRRUST database. Reliable candidates of transcription factors regulating each state were supported by at least two databases. **b** Table showing enriched transcription factors based on differentially expressed genes in each cell state. Black lines indicate TFs enrichment in two databases. **c** Activity score of transcription factors enriched in e-Sen state presented by target gene index over the pseudotime. **d** Representative β-gal staining images of mammary gland from young (3 month), aged (17 month) and *K14-Cre Bcl11b*$^{fl/fl}$ (3 month) mice showing senescent cells. Scale bar, 20 μm. **e** Percentage of p16^{Ink4a} (scale bar, 20 μm) positive cells in mammary epithelial cells from young (4 month, $n = 10$ mice), old (22 month, $n = 10$ mice), control *K14-Cre Bcl11b*$^{wt/wt}$ (4 month, $n = 8$ mice) and *K14-Cre Bcl11b*$^{fl/fl}$ (4 month, $n = 9$ mice) mice. Statistical analysis was performed using two-tailed unpaired *t*-tests; data are presented as mean ± SD. **f** Relative basal luminal proportion of mammary gland epithelia in young ($n = 54$ mice) and old ($n = 13$ mice). Statistical significance was determined by two-tailed unpaired *t*-tests; data are presented as mean ± SD. **g** Quantification of relative basal/luminal proportion in mammary epithelia in *K14-Cre Bcl11b*$^{fl/fl}$ mTmG reporter mice ($n = 9$ mice). Statistical significance was determined by two-tailed unpaired *t*-tests; data are presented as mean ± SD. **h**, **i** Density map (**h**) and percentage (**i**) of mammary cell states in each age group and *K14-Cre Bcl11b*$^{fl/fl}$ group. **j** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis showing pathways significantly enriched in *K14-Cre Bcl11b*$^{fl/fl}$ CD49f^{high}EpCAM^{low} cells. enrichKEGG from clusterProfiler package was used to enrich pathway, and BH adjusted *p*-value was used. **k** Boxplots showing the scEntropy score of *K14-Cre Bcl11b*$^{fl/fl}$ CD49f^{high}EpCAM^{low} cells increased to a level similar to l-Sen cells. Cell number: 360 (a-Young), 666 (q-Young), 664 (e-Sen), 291 (l-Sen), 149 (*K14-Cre Bcl11b*$^{fl/fl}$). Two-sided wilcoxon -test was used. **l** Comparison of SASP gene score in WT ($n = 236$ cells from 3 mice) and *K14-Cre Bcl11b*$^{fl/fl}$ ($n = 149$ cells from 3 mice) CD49f^{high}EpCAM^{low} cells. Two-sided wilcoxon -test was used.

(GSEA) showed that WT tumor cells were significantly enriched with a-Young signature genes, while *Bcl11b*-KO tumor cells were enriched with l-Sen genes (Fig. 5g). In addition, the *Bcl11b*-KO tumor cells themselves exhibited high levels of senescence-associated cytokine IL6 and high NF-kB activity (Fig. 5h, i). *Bcl11b*-KO also increased immune checkpoint pathway (PD-L1 and PD-1) (Supplementary Fig. 8e), consistent with its role in promoting tumor formation. These data support the idea that the l-Sen cells induced by knocking out *Bcl11b* expression were intrinsically susceptible to transformation into cancer cells.

https://doi.org/10.1038/s41467-024-49106-2

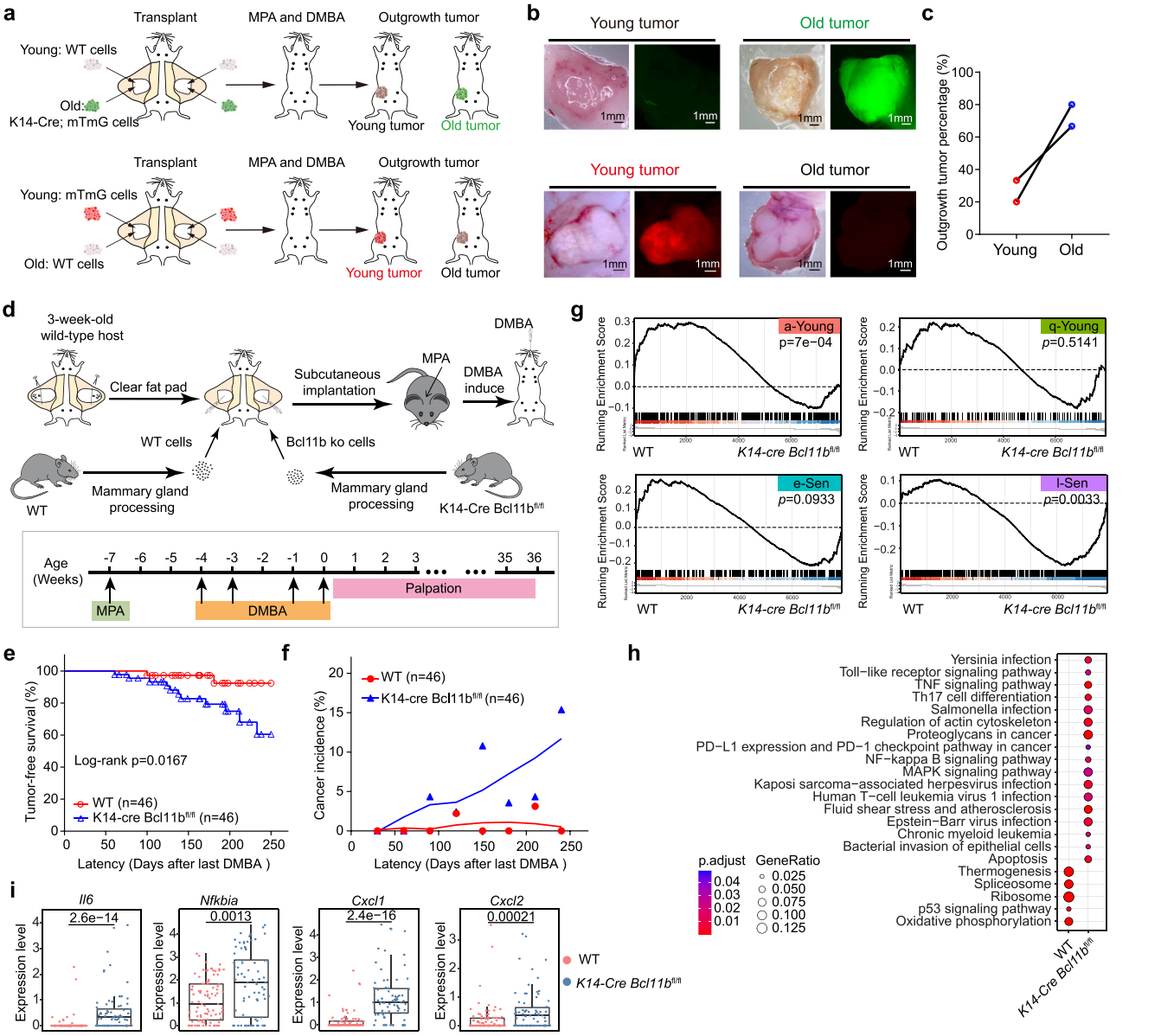

Fig. 5 | The l-Sen cells induced by knocking out *Bcl11b* expression were intrinsically susceptible to transformation into cancer cells. a Schematic diagram showing the DMBA induced cancer assay for young and old transplanted mammary cells. Upper panel: mice transplanted with 2 month old young mammary cells (wt) and 24 month old mammary cells (with GFP reporter) mammary cells at equal MRUs were used for tumor formation assay. After cells grafted, mice were treated with MPA plus DMBA as indicated in the schematic diagram. Lower panel: mice transplanted with 2 month old young mammary cells (with tdTomato reporter) and 12 month old mammary cells (wt) at equal MRUs were used for tumor formation assay. After cells grafted, mice were treated with MPA plus DMBA as indicated in the schematic diagram. **b** Representative images of tumors from young and old mammary gland cells. Scale bar, 1 mm. **c** Outgrowth tumor percentage of young and old mammary gland cells. **d** Schematic diagram showing the DMBA induced cancer assay for wild-type and *K14-Cre Bcl11b*$^{fl/fl}$ transplanted mammary cells. Mice transplanted with WT or *K14-Cre Bcl11b*$^{fl/fl}$ mammary cells were used for tumor formation assay. After cells grafted, mice were treated with MPA plus DMBA as indicated in the schematic diagram. **e** Tumor free survival curve of DMBA induced cancer assay for WT and *K14-Cre Bcl11b*$^{fl/fl}$ transplanted mammary cells. Statistical analysis was determined by log-rank test. Latency was calculated from the day of last DMBA treatment. **f** Cancer incidence of WT and *K14-Cre Bcl11b*$^{fl/fl}$ transplanted mice at each designated latency time (30 days for each period). **g** Gene Set Enrichment Analysis (GSEA) showing WT tumor cells were enriched for a-young signature genes and *K14-Cre Bcl11b*$^{fl/fl}$ tumor cells were enriched for l-Sen signature genes. GSEA from clusterProfiler package was used to perform analysis. **h** KEGG pathway analysis of the differentially expressed genes in DMBA induced tumors in WT and *K14-Cre Bcl11b*$^{fl/fl}$ CD49f^{high}EpCAM^{low} cells. enrichKEGG from clusterProfiler package was used to enrich pathway, BH adjusted *p*-value was used. **i** SASP and NF-kB related genes were up-regulated in *K14-Cre Bcl11b*$^{fl/fl}$ CD49f^{high}EpCAM^{low} tumor cells. WT ($n = 236$ cells from 3 mice) and *K14-Cre Bcl11b*$^{fl/fl}$ ($n = 149$ cells from 3 mice); Two-sided wilcoxon -test was used.

Multiple aging- and longevity-related pathways are governed by *Bcl11b*

We then asked, how does *Bcl11b* play such a striking role in mammary cell senescence? To answer this question, we performed ChIP-seq analysis to profile the targets of *Bcl11b* in mammary cells. We detected 1197 Bcl11b binding sites in the genome, with 918 promoter binding sites and 279 enhancer binding sites (Fig. 6a, b; Supplementary Data 3). The KEGG and GO pathway enrichment analysis revealed that many of the aging-associated pathways were the direct targets of Bcl11b, including the energy/nutrient sensing pathways 'PI3K-Akt signaling', 'mTOR signaling', 'AMPK signaling', which have been implicated in longevity regulation[72]; the inflammation pathway 'TNF signaling'; the fate determination pathway 'Notch signaling", and 'Wnt signaling' and cancer-related pathways (Fig. 6c, d; Supplementary Data 5). We also

https://doi.org/10.1038/s41467-024-49106-2

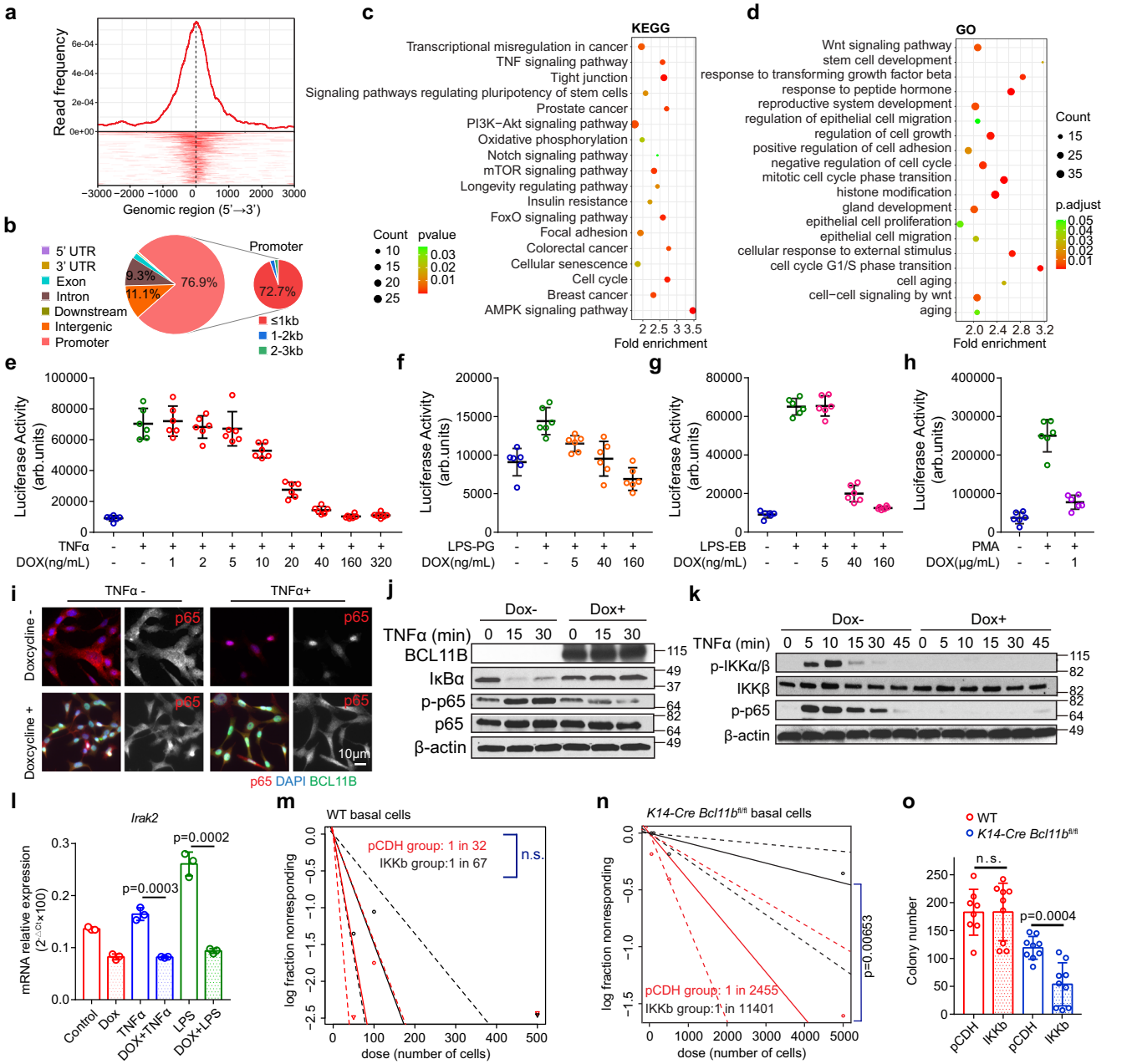

Fig. 6 | Multiple ageing associated pathways are governed by *Bcl11b*. **a** The distribution of global Bcl11b ChIP-seq peaks at transcription start site (TSS). **b** Pie chart showing Bcl11b ChIP-seq peak distribution in different genomic regions. **c**-**d** KEGG (**c**) and GO (**d**) pathway analysis of Bcl11b's ChIP targets. enrichKEGG (**c**) or enrichGO (**d**) from clusterProfiler package was used to enrich pathway, BH adjusted *p*-value was used. **e**–**h** NF-kB luciferase reporter assay showing that *Bcl11b* expression regulates NF-kB activity. The pInducer-*Bcl11b* Comma Dβ cells with NF-kB luciferase reporter were treated with TNFα (**e**; 20 ng/mL, $n = 6$ biological replicates), LPS-PG (**f**; 1 μg/mL, $n = 6$ biological replicates), LPS-EB (**g**; 10^3 EU/mL, $n = 6$ biological replicates), PMA (**h**; 100 ng/mL, $n = 6$ biological replicates). Data are presented as mean ± SD. **i** Immunofluorescence of p65 showing that induced expression of *Bcl11b* efficiently suppressed the nuclear import of p65 induced by TNFa (20 ng/mL) in pInducer-*Bcl11b* Comma Dβ cells. Red: p65; Green: Bcl11b; Blue: DAPI. Repeat 3 times with 3 biologically independent samples. **j** Western blot analysis showing the induced Bcl11b expression by doxycycline (50 ng/mL) inhibited IκBα degradation, p65 phosphorylation upon TNFα (20 ng/mL) treatment in pInducer-*Bcl11b* Comma Dβ cells. $n = 3$ biologically independent samples; repeat 3 times. **k** Western blot analysis showing the induced *Bcl11b* expression by doxycycline (50 ng/mL) inhibited IKKa/b phosphorylation upon TNFα (20 ng/mL) treatment in pInducer-*Bcl11b* Comma Dβ cells. $n = 3$ biologically independent samples, repeat 3 times. **l** Real time PCR confirming *Irak2* mRNA expression is regulated by induced *Bcl11b* expression. $n = 3$ biologically independent samples; data are presented as mean ± SD; two-tailed unpaired *t*-tests. **m** Extreme limiting dilution analysis (ELDA) plot showing the transplant of WT CD49f^{high}EpCAM^{low} cells transduced with pCDH or pCDH-IKKb vectors. n.s., not significant. **n** ELDA plot showing the transplant of *K14-Cre Bcl11b*$^{fl/fl}$ CD49f^{high}EpCAM^{low} cells transduced with pCDH or pCDH-IKKb vectors. **o** Colony formation assay of CD49f^{high}EpCAM^{low} cells from WT and *K14-Cre Bcl11b*$^{fl/fl}$ mice transduced with pCDH ($n = 3$ mice) or pCDH-IKKb ($n = 3$ mice) vectors. Data are presented as mean ± SD; two-tailed unpaired *t*-tests; 3 technical replicates for each mice; n.s., not significant.

identified targets associated with pluripotency and the NF-kB pathway that were regulated by Bcl11b activity (Supplementary Fig. 9a). These pathways were also associated with both the q-Young-to-e-Sen and e-Sen-to-l-Sen cell transitions. These findings suggest that *Bcl11b* is a master regulator of aging progression by comprehensively repressing aging-associated pathways.

As a stress sensing pathway, 'NF-kB signaling' has been recognized as an important contributor to the aging process; therefore, we

wondered whether *Bcl11b* physiologically regulates NF-kB's activity. A NF-kB luciferase activity assay suggested that induced expression of *Bcl11b* efficiently repressed NF-kB activity that had been triggered by TNFα, lipopolysaccharide (LPS) or phorbol myristate acetate (PMA) (Fig. 6e–h). *Bcl11b* regulates NF-kB expression by upstream of the signaling pathway, as indicated by the nuclear transport of RelA, the degradation of IkBa, and the activation of IKKa/b all blocked by *Bcl11b* expression (Fig. 6i–k and Supplementary Fig. 9b). When we analyzed our ChIP-seq data, we found that Bcl11b directly bound to the promoter regions of *Irak2* (an essential NF-kB signaling mediator), *Nfkbia* and *Nfkb2*, suggesting direct transcriptional regulation (Supplementary Fig. 9a). Indeed, when we induced the expression of *Bcl11b*, the mRNA expression of *Irak2* was significantly suppressed (Fig. 6l and Supplementary Fig. 9c). These data collectively demonstrate that *Bcl11b* directly regulates the mammary cell stress response program as one of the mechanisms to slow down aging progression.

NF-kB promotes stem cell exhaustion in the absence of *Bcl11b*

NF-kB signaling is a well-characterized stress-sensing pathway[73] and plays pleiotropic roles in a variety of biological processes[74]. This pathway is aberrantly activated during aging through an unclear mechanism. We asked, under what conditions does NF-kB activation convert a cell from a young state to a senescent state? To answer this question, we first tested stem cell activity after NF-kB activation in WT cells and *Bcl11b*-KO cells and found that the enforced expression of IKKb triggered NF-kB activation (Supplementary Fig. 10a–c) but did not significantly reduce the mammary reconstitution ability in the WT cells (1 in 32 of the control cells vs. 1 in 67 of the IKKb-overexpressing cells, $p > 0.05$) (Fig. 6m, Supplementary Fig. 10d, f, i, j). However, when the activation of NF-kB was performed in cells with a *Bcl11b*-KO background, the mammary reconstitution ability was significantly reduced (1 in 2455 *Bcl11b*-KO cells vs. 1 in 11401 *Bcl11b*-KO+*IKKb*-positive cells, $p < 0.01$) (Fig. 6n, Supplementary Fig. 10e, g, i, k, l). The differential NF-kB activation effects in cells with different backgrounds imply that NF-kB activity alone is not sufficient to drive stem cell exhaustion, and the reduced regeneration ability, which is a frequent hallmark of tissue aging, is initiated through a specific epigenetic modification program involving multiple signaling pathways that depend on *Bcl11b* expression. In accordance with this idea, we observed differential colony formation rates of IKKb-expressing WT and *Bcl11b*-KO cells (Fig. 6o and Supplementary Fig. 10h), indicating that *Bcl11b* may protect mammary cells from aging in response to NF-kB activation.

Manipulation of l-Sen cells regulates aging-related cancer transformation

We then explored the possibility of reversing the progressive aging process as a strategy to reduce cancer vulnerability. As the *Bcl11b* activity declined with age, we screened a small customized drug pool to identify a chemical drug that would enable us to reduce the number of senescent cells and increase the number of young cells. We tested PI3K-Akt-mTOR inhibitors, NF-kB pathway inhibitors, Jak-Stat inhibitors, wnt inhibitors, notch inhibitors and hedgehog inhibitors to evaluate their roles in enhancing *Bcl11b* expression, which is an indication of a young cell state. We found that TPCA-1[75,76], a dual NF-kB and Jak-Stat inhibitor, exerted the most striking effect in restoring *Bcl11b* expression, while other NF-kB or Jak-Stat single pathway inhibitors played minimal roles in *Bcl11b* expression, indicating that aging is a progressive and coordinated process that might require the input of multiple pathways (Fig. 7a and Supplementary Fig. 11a). Consistent with the role played by TPCA-1 in promoting *Bcl11b* expression in vitro (Fig. 7b and Supplementary Fig. 11b), when we administered TPCA-1 to 12 month-old mice continuously for 1 month, mammary $CD49f^{high}EpCAM^{low}$ cells were clearly younger, showing a remarkable increase in a-Young cells accompanied by a dramatic reduction in the number of e-Sen and l-Sen cells (Fig. 7c, d and Supplementary Fig. 11c–i). The *Bcl11b* activity score and metabolic pathways were clearly restored after TPCA-1 treatment, while aging-related pathways, including the senescence, NF-kB, Jak-Stat, MAPK and PI3K-Akt pathways, were all significantly suppressed (Fig. 7e–h; Supplementary Data 6). This molecular profile is very similar to that of the 2 month-old mammary glands, suggesting that TPCA-1 can change the senescence profile of mammary gland cells under physiological conditions. In contrast, this rejuvenation effect cannot be achieved by NF-kB inhibition alone (Supplementary Fig. 13). To test whether TPCA-1's rejuvenation effect is dependent on *Bcl11b* expression, we performed TPCA-1 treatment on 3 month-old *Bcl11b* KO mice continuously for 1 month, we found that the ageing phenotype of *Bcl11b* KO mammary cells on the pseudotime can be efficiently rescued by TPCA-1 treatment (Supplementary Fig. 12). The l-Sen cells of TPCA-1 treated mice were significantly reduced compared with *Bcl11b* KO cells, and the SASP pathway, NF-kB signaling pathway, JAK-STAT signaling pathway were all efficiently suppressed. This suggests that when *Bcl11b*'s downstream target signaling pathways were inhibited, it can play a similar role as the Bcl11b expression and block the accelerated ageing progression.

To determine whether TPCA-1 induced younger mammary gland can efficiently reduce cancer incidence, we performed 7 weeks of intraperitoneal administration of TPCA-1 after medroxyprogesterone acetate (MPA)- and DMBA-treatment, and tracked the rate of cancer incidence (Fig. 7i). We found that TPCA-1 treatment successfully and efficiently decreased the tumor burden and significantly increased tumor-free mouse survival (Fig. 7j, k, $p < 0.05$). These data demonstrate that mammary cancer formation can be controlled with a strategy that reshape the aging tissue transcriptome to that of tissue in the juvenile state to reduce cancer susceptibility.

Discussion

The global aging trend has become increasingly common worldwide, with an anticipated increase in cancer burden. The biological relationship between aging and cancer has been a critical issue to clarify to guide prophylactic measures. Our chronological single-cell transcriptome analysis of the mammary gland enabled us to reconstruct a molecular portrait of the physiological aging process, which revealed heterogeneous senescent cell states and progressive aging processes intrinsic to epithelial cells. This molecular map bridges cellular senescence and cancer initiation and answers the long-standing question, why do aging cells with degenerative activities paradoxically foster cancer formation? Our study implies that the senescent state is a heterogeneous, dynamic and entropic cell state with functional deterioration, accompanied by cell cycle arrest at early stage and activated stem cell/cancer programs at later stage. The paradox can be explained by a progressive ageing model that senescence and cancer are successive biological steps in the same linear developmental trajectory, not bifurcating biological processes (Fig. 7l). This understanding can help us design innovative strategies to block aging and cancer at various intersections.

Our study led us to rethink the prevailing mutation accumulation model for explaining the aging–cancer relationship[77]. In the last century, Carl Nordling proposed the theoretical framework that carcinogenesis is driven by mutation of the genome[78]. This seminal concept was first developed into a multistage model for malignancy transformation[79]. However, this theory does not explain why a substantial portion of mutations occur early in life, while cancers arise exponentially later in life[80,81]. Neither does it explain the disproportion between cancer frequency and animal body size, as well as the scaling of cancer incidence to animal lifespan[82]. These facts implicate that there might be more factors beyond genetic mutations involved in determination of cancer initiation[77]. In this study, we clearly observed transcriptome alterations during aging, and these were biologically

https://doi.org/10.1038/s41467-024-49106-2

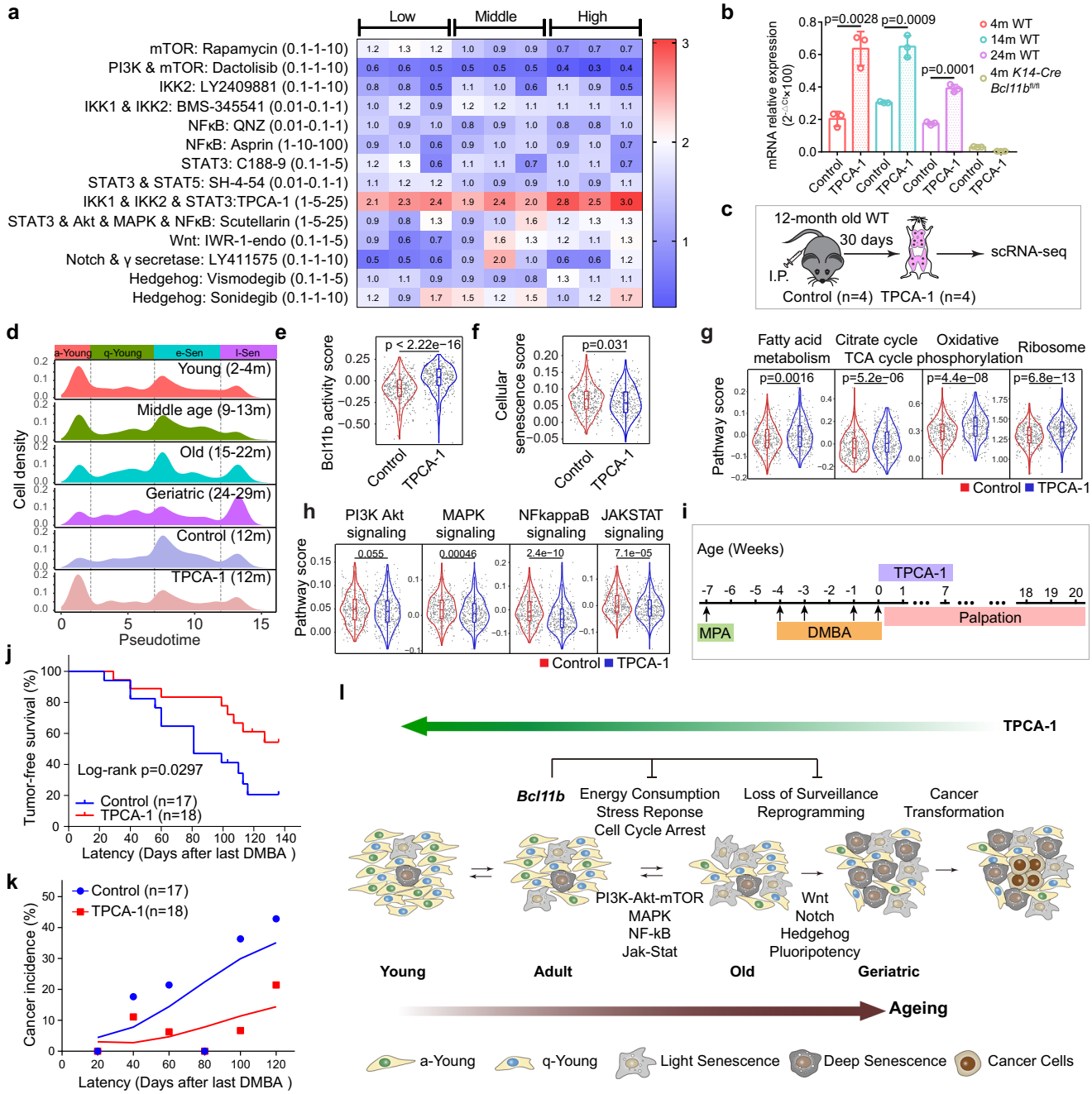

Fig. 7 | TPCA-1 reshapes ageing programs at the transcriptome level. **a** Evaluation of drugs that can efficiently upregulate *Bcl11b* expression. Real time PCR quantification of *Bcl11b* expression levels treated with low-middle-high dose of Rapamycin, Dactolisib, LY2409881, BMS-345541, QNZ, Asprin, C188-9, SH-4-54, TPCA-1, Scutellarin, IWR-1-endo, LY411575, Vismodegib and Sonidegib with β-actin as an internal control ($n = 3$ biologically independent samples) in Comma Dβ cells. Fold changes compared to the control group were shown in the data matrix. Data were compiled and presented as heatmap. **b** Real time PCR confirming *Bcl11b* mRNA expression were upregulated by TPCA-1 (25 μM) treatment in primary mammary CD49f^{high}EpCAM^{low} cells isolated from 4 m WT, 14 m WT, 24 m WT, 4 m *K14-Cre Bcl11b*$^{fl/fl}$ mice. $n = 3$ biologically independent samples; bar and whiskers denote mean ± SD; two-tailed unpaired *t*-tests. **c** Schematic diagram showing the strategy of TPCA-1 (10 mg/kg) treatment in vivo on 12 month-old mice. 373 cells of control group and 367 cells of TPCA-1group were used for analysis. **d** Cell density distribution of mammary CD49f^{high}EpCAM^{low} cells from young, middle age, old, geriatric mice, along with control (12 month-old, $n = 4$) and TPCA-1treated (12 month-old, $n = 4$) mice over pseudotime. **e**, **f** Violin plots showing the increased *Bcl11b* activity score (**e**) and decreased cellular senescence score (**f**) in TPCA−1 treated group. $n = 4$ mice for control and TPCA-1 group. Two-sided wilcoxon -test was used. **g** KEGG analysis of the up-regulated (**g**) and down-regulated (**h**) pathways in CD49f^{high}EpCAM^{low} cells of TPCA-1 treated mammary gland. $n = 4$ mice for control and TPCA-1 group. Two-sided wilcoxon -test was used. **i** Schematic diagram showing the DMBA induced cancer assay for mice treated with DMSO (4%) and TPCA-1 (10 mg/kg). **j** Tumor free survival curve of DMBA induced cancer assay for mice treated with DMSO (4%) and TPCA-1 (10 mg/kg) as indicated. Statistical test was performed by log-rank test. Latency was counted from the date of the last DMBA treatment. **k** Cancer incidence of DMSO (4%) and TPCA-1 treated mice at each designated latency time (20 days for each period). Latency was counted from the date of the last DMBA treatment, $p < 0.05$. **l** Schematic diagram showing the hypothesis of how the progressive senescence programs result in cancer vulnerability.

correlated with cancer initiation, and we identified a master fate determinant, *Bcl11b*, which is involved in epigenetic regulation[83] in aging and cancer. In addition, we demonstrated that after mutation, modulation of epigenetic reprogramming with chemical inhibitors targeting NF-kB and Jak-stat efficiently reduced cancer formation. This outcome suggests that although mutation is essential and may be indispensable for triggering cancer formation, epigenetic programs might be equally crucial to cancer development and may be more manageable. Therefore, modulating epigenetic modifications may be a better and more feasible way than genetic intervention to control cancer incidence at the population level. We envision that a comprehensive molecular understanding of epigenetic regulation in aging may help us find diverse and promising targets to eventually reduce cancer occurrence.

Previous single cell analyses on ageing reported that ageing program is tissue and cell type dependent[33–35], which argues against the point that the Bcl11b is a universal ageing regulator in various tissues. However, data mining of TCGA database suggests, across various tissues/tumors, Bcl11b's activity (indicated by its ChIP-seq targets) is frequently reduced in tumors compared with normal tissues (Supplementary Fig. 14). Consistent with this data, the related senescence pathways, oncogenic factors and pathways are also elevated in tumors compared with normal tissues. Although Bcl11b may not be a global ageing gene, we think Bcl11b's downstream pathways might be conserved in the ageing process across various tissues.

We normally regard 'senescence' as a certain static cellular state associated with cell cycle arrest. However, increasing evidence suggest that the senescent cell cycle arrest may not be terminal, in certain conditions, senescence cells can reenter cell cycle[7,30,84]; in addition, senescent features are heterogeneous including aberrant epigenetics, abnormal secretome, functional decline etc. Cell-cycle arrest is one of the features, but not the only feature. Senescence is also a dynamic process showing different features at different stages, which has been reviewed[85]. According to our data, cell cycle arrest is a feature of early senescent cells, at later senescence stage, these cells acquire ability to reenter cell cycle.

While our study successfully identified senescence heterogeneity using single-cell transcriptomic analysis, it is important to acknowledge its limitations. Future experiments that experimentally separate senescent cells with distinct cell fates and observe their behavior under physiological conditions would provide valuable insights. Additionally, our investigation focused on a specific cell population in the mammary gland during the longitudinal aging process, and we established molecular connections between aging and cancer using a DMBA treatment-based cancer initiation model. It would be intriguing and promising to explore whether the aging dynamics we discovered extend to other mammary populations or even other organs in diverse cancer models. This avenue of research holds great potential for further understanding the broader implications of aging in cancer development.

Methods

Mice and in vivo models

Animals were housed in a specific pathogen-free conditions and fed standard mouse chow. All animal experiments were carried out in compliance with China laws and regulations. The local institutional animal ethics board (Institutional Animal Care and Use Committee of Westlake University) approved all mouse experiments (permission numbers: 19-001-2-CS). Experiments were performed in accordance with government and institutional guidelines and regulations. All mice are housed at 20–24 °C with 40–60% humidity, and 12 h cycle of light/darkness (7 a.m. – 7 p.m.).

The $Bcl11b^{flox/flox}$ mice (C57BL/6 background) were generously provided by Mark Leid's lab and the B6N.Cg-Tg (KRT14-cre)1Amc/J (stock number 018964) were purchased from Jackson Laboratory. mTmG mice (B6.129(Cg)-Gt(ROSA)26Sortm4(ACTB-tdTomato,-EGFP) Luo/J, 007676) were purchased from The Jackson Laboratory (The Jackson Laboratory, Bar Harbor, Maine, USA). Female C57BL/6 mice, 2–6 months old were purchased from Jackson Laboratory, and were maintained till 29 months.

Tissue processing and flow cytometry

Mammary glands were collected from 2nd, 3rd, 4th pair of mammary glands of C57/BL6 mice, and were dissected and processed according to the published protocol[86] with minor revision. Mammary glands were minced into 1 mm^3 size using a tissue cutting blade and digested with 0.5 mg/mL Collagenase type III (Worthington, LS004182) and 50 U/mL hyaluronidase (Worthington, LS002592) for 2 h with gentle pipetting every 30 min. Digested mammary homogenate was collected and treated with ACK lysing buffer for 5 min on ice, then was digested using 0.25% Trypsin-EDTA (GIBCO) for 5 min, followed by DNase I (Worthington, LS002139) digestion. After filtered by 70 µm strainer, the dissociated mammary cells were stained with CD45 (560501, BD, 1:200), CD31 (562939, BD, 1:200), Ter119 (560504, BD, 1:200), EpCAM (118220, BioLegend, 1:200), CD49f (313606, BioLegend, 1:200) for 20 min on ice. Cells were washed and resuspended in HBSS + 2% FBS + 1% PSA + DAPI (1 µg/mL), then were sorted using FACS Aria II (BD Bioscience). FACS data was analyzed by FlowJo (V10).

Transplantation

For transplantation assay of young and aged mammary gland, 3 month-old and 18 month-old WT C57BL/6 mice were lineage depleted using Mammary Epithelium Enrichment Kit (Stem Cell Technologies) and resuspended in injection media (HBSS + 2% FBS + 1% PSA + 50% Matrigel) to 50 k/5 µL, and serially diluted to 20 k/5 µL, 5 k/5 µL and 2 k/5 µL. Cells were subjected to limiting dilution transplant to the cleared fat pad of 3 week-old recipient mice (C57BL/6). Briefly, the recipient mice were anesthetized with pentobarbital sodium at a dose of 70 mg/kg. The inguinal rudimentary tree was removed and 5 µL of cell suspension was injected onto the residual fat pad using 25 µL Hamilton Syringe. 7 weeks later, the recipient mice were analyzed using mammary gland whole mount carmine staining. The MRU frequency and confidence interval were determined by ELDA.

For transplantation of Bcl11b knockout cells, mammary cells from 4 month-old wild-type and Krt14-cre $Bcl11b^{flox/flox}$ mice were resuspended in injection media (HBSS + 2% FBS + 1% PSA + 50% Matrigel) to 200 k/5 µL. Cells were injected into the cleared fat pad of 3 week-old recipient mice (C57BL/6). 7 weeks later, the recipient mice were subsequently mated with WT male mice for 3 rounds. Mammary Tissues were collected at designated time points and subjected to whole mount analysis.

For transplant assay of IKKb overexpression cells, $CD49f^{high}EpCAM^{low}Lin^{-}$ cells from 2 month-old C57BL/6 WT mice were sorted using FACS and resuspended in culture media (DMEM/F12 + 2% FBS + 1% PSA + 2% B27) supplemented with EGF (10 ng/ mL, BD Bioscience), Rspo1(250 ng/mL, R&D), ROCK inhibitor Y27632 (10 uM, Sigma) to 100 k/200 µL/well in 96-well non-adherent culture plate (Corning). Cells were transduced with lentivirus-pCDH or lentivirus-pCDH-IKKb at MOI 20 overnight. The next day, cells were seeded on top of the Matrigel to do the colony culture. One week later, colonies were digested using DispaseII 1 mg/mL (Sigma, D4693) for 1 h followed by treatment of 400 µL TrypLE™ Select (GIBCO)/eppendorf tube at 37 °C for 5 min. Dissociated cells were neutralized with HBSS + 2% FBS + 1% PSA and stained with DAPI. GFP+ cells were sorted and then resuspended with injection media (HBSS + 2% FBS + 1% PSA + 50% Matrigel) to 2.5k/5 µL, and serially diluted to 500 cells/5 µL, 100 cells/5 µL and 50 cells/5 µL. Cells were subjected to limiting dilution transplant to recipient mice. Mice were maintained in aseptic sterile condition for 7 weeks before whole mount analysis. For secondary transplant, mammary fat pad at 2.5 k/5 µL dilution, which exclusively

gave rise to full tree, were collected and digested to single cell suspension. Cells from one fat pad of 2.5k/5 µL dilution group were divided equally to six parts and transplanted to 6 cleared fat pad of recipient mice, respectively. Mice were maintained for 7 weeks before whole mount analysis. The mammary gland area percentage was determined by the outgrowth area divided by cleared fat pad area. The area was measured by Image J software.

Colony formation assay

For colony formation assay, 35 µL/well growth factor reduced Matrigel (Corning, 356231) was overlaid on the 96-well plate and solidified at 37 °C for 10 min. CD49f^{high}EpCAM^{low}Lin$^-$ cells were collected from FACS and cultured in 200 µL culture media (DMEM/F12 + 2% FBS + 1% PSA + 2% B27) supplemented with 10 ng/ mL EGF (BD Bioscience), 250 ng/ mL Rspo1(R&D), 10 µM ROCK inhibitor Y27632 (Sigma), and then were plated on top of the Matrigel. Cells were cultured at 37 °C incubator with 5% CO_2 for 7 days.

For colony formation ability test of aging CD49f^{high}EpCAM^{low}Lin$^-$ cells, CD49f^{high}EpCAM^{low}Lin$^-$ cells were sorted from 2–4 months, 12-13 months and 22–29 months mice, seeded on Matrigel in 96-well plates, 3000 cells per well, three replicates for each mouse sample. Colony number was counted after 7 days culture.

For colony formation assay of IKKb overexpression cells, mammary CD49f^{high}EpCAM^{low}Lin$^-$ cells were sorted from 2–4 months WT or K14-Cre Bcl11b$^{fl/fl}$ mice and transduced with pCDH and pCDH-IKKb virus for 7 days. 9000 GFP positive cells were sorted and seeded to 3000 cells/200 µL/well equally in three wells in the culture medium. Colony number was counted after 7 days culture. Each group was repeated with three biological replicates.

Mammary gland whole mount carmine staining

Mammary gland was dissected and fixed in Carnoy's solution (60% Ethanol, 30% $CHCl_3$,10% Acetic Acid) for 4hrs, washed with 70% ethanol and ddH_2O, and then stained with carmin-alum staining solution (0.2% wt/vol carmine (Sigma, C1022), 0.5% wt/vol aluminum potassium sulfate (Sigma, A7176) and 0.01% wt/vol thymol in ddH_2O) overnight. Tissue was dehydrated through 75%, 95% and 100% ethanol, and then placed in xylene to remove the fat tissue. For long-term preservation, tissue was mounted by Permount® (Fisher Scientific). Images were obtained using a stereomicroscope (Nikon, SMZ18).

β-Galactosidase staining

Mammary gland was dissected and immediately fixed by 2% formalin containing 0.25% glutaraldehyde for 1.5 h followed by 30% sucrose infiltration overnight, and then was embedded in O.C.T. compound (Tissue-Tek) and frozen at −80 °C. Frozen tissue was sectioned to 14 µm using Cryostat Leica CM3050 S (LEICA). Frozen sections were hydrated with PBS for 10 min at RT and stained using Senescence β-Galactosidase staining kit according to manufacturer's instructions (Cell signaling technology, 9860). The sections were incubated at 37 °C in a dry incubator (no CO_2) for 48 h and photographed.

Western blot

Comma D beta cell line[87] was kindly provided by Dr. Medina, and was described previously. Comma D beta cell line was cultured in DMEM-F12 (Invitrogen) supplemented with 2% of Fetal Bovine Serum (Hyclone), 1% PSA (Invitrogen), 10 ng/mL EGF(BD) and 5 µg/mL Insulin (Sigma), at 37 °C with 5% CO2. Comma D beta cells were harvested and lysed using 2X Laemmli SDS sample buffer (100 mM Tris pH6.8, 10% glycerol, 4% SDS, 0.01% Bromophenol Blue), and boiled on heat block at 100°C for 15 min. Samples were loaded to 4%–20% precast gradient gel (Bio-Rad) and electrophoresed at 200 v for 45 min, and transferred to Odyssey® nitrocellulose membrane (LI-COR). After being blocked by PBS + 0.1% tween 20 + 5% Non-fat dry milk for 1 h at RT, the membrane was subjected to primary antibody staining beta Actin (Santa Cruz), Rat anti-Bcl11b (abcam, ab18465, 1:1000), Rb anti-IKKβ (Cell signaling technology, 2370, 1:1000), Rb anti-p-IKKα/β (Cell signaling technology, 2697, 1:1000), Rb anti-p-p65 (Cell signaling technology, 3033, 1:1000), Rb anti-p65 (Cell signaling technology, 8242, 1:1000), Mouse anti-IκBα (Cell signaling technology, 9247, 1:1000) at 4 °C overnight. Membrane was washed using PBST (PBS + 0.1% Tween 20) 3 × 10 min, and stained with secondary antibodies HRP-Donkey anti mouse (Cell signaling technology, 7076 S, 1:10000), HRP-Donkey anti rat (Cell signaling technology, 7077 S, 1:1000) or rabbit (Cell signaling technology, 7074 S, 1:1000) at RT for 1 hr. Membrane was subsequently washed 3 × 10 min by PBST and developed using SuperSignal® West Dura Extended Duration Substrate (Thermo Scientific, 34094) and imaged by Gel imaging system (GE, AI680RGB).

RNA extraction and Real-Time PCR

For the TNFα and LPS treatment assay, after starvation for 12 h, pInducer-Bcl11b Comma D beta cells were treated with 50 ng/mL doxycycline for 12 h to induce Bcl11b expression. Cells were then treated 20 ng/mL TNFα or 10^3 EU/mL LPS-EB for 10 h before lysed by 400 µL Trizol (Life Technologies),. RNA was extracted according to the manufacturer's instruction with addition of ultrapure glycogen (Thermo Scientific, R0551) as carrier. RNA was reverse transcribed to cDNA using PrimeScript™ RT reagent kit (TaKaRa, RR037A) according to the manufacturer's instructions. cDNA was then subjected to the real time PCR for specific gene target by TB Green Premix Ex Taq (TaKaRa, RR420B) according to manufacturer's instructions using Real Time PCR system (SIS-PCR005, Jena).

For the drug screening assay, Comma D beta cells were treated with Rapamycin (0.1 µM, 1 µM, 10 µM, Selleck, S1039), Dactolisib (0.1 µM, 1 µM, 10 µM, Selleck, S1009), LY2409881 (0.1 µM, 1 µM, 10 µM, Selleck, S7697), BMS-345541 (0.01 µM, 0.1 µM, 1 µM, Selleck, S8044), QNZ (0.01 µM, 0.1 µM, 1 µM, Selleck, S4902), Asprin (1 µM, 10 µM, 100 µM, Selleck, S3017), C188-9 (0.1 µM, 1 µM, 5 µM, Selleck, S8605), SH-4-54 (0.01 µM, 0.1 µM, 1 µM, Selleck, S7337), TPCA-1 (1 µM, 5 µM, 25 µM, Selleck, 2824), Scutellarin (1 µM, 5 µM, 25 µM, Selleck, S3810), IWR-1-endo (0.1 µM, 1 µM, 5 µM, Selleck, S7086), LY411575 (0.1 µM, 1 µM, 10 µM, Selleck, S2714), Vismodegib (0.1 µM, 1 µM, 5 µM, Selleck, S1082) and Sonidegib (0.1 µM, 1 µM, 10 µM, Selleck, S2151) for 24 h. For primary mammary cell validation experiment, CD49f^{high}EpCAM^{low}Lin$^-$ cells from 4 month-old, 14 month-old, 24 month-old WT mice and 4 month-old K14-Cre Bcl11b$^{fl/fl}$ mice were sorted into the 96-well Ultra-Low attachment culture plate (Corning, 3474) (10,000 cells/200 µL/well) and cultured in the culture media (DMEM/F12 + 2% FBS + 1% PSA + 2% B27 + 10 ng/mL EGF + 250 ng/ mL Rspo1 + 10 µM Y27632) with 25 µM TPCA-1 or DMSO for 24 h. DAPI negative cells were then sorted to perform RT-PCR assay.

Data was analyzed by Excel and GraphPad Prism 7.00. The relative gene expression was normalized by β-actin expression. Target gene Sybrgreen primers were designed by IDT (Integrated DNA Technologies) and the primers are listed below: Bcl11b (Exon1-2): Forward ATGCCAGAATAGATGCCGG, Reverse CTCTATCTCCAGACCCTCGTC; Bcl11b (Exon2-4): Forward AGGAGAGTATCTGAGCCAGTG, Reverse GTTGTGCAAATGTAGCTGGAAG; Irak2: Forward TGTCACCTGGAACTCTACCG, Reverse TTTCTCCTGTTCATCCTTGAGG; Tnfrsf1b: Forward ACTCCAAGCATCCTTACATCG, Reverse TTCACCAGTCCTAACATCAGC; β-actin: Forward ACCTTCTACAATGAGCTGCG, Reverse CTGGATGGCTACGTACATGG.

Luciferase assay

Comma D beta cell line which was stably expressed *pInducer Bcl11b* plasmid and NFκB-inducible Luciferase reporter plasmid was used for this experiment. After starvation for 12 h, cells were treated with Doxycycline (Sigma, D9891) at various doses followed by 20 ng/mL TNFα (Biolegend, 575204), 1 µg/mL LPS-PG (Invivogen, tlrl-ppglps), 10^3 EU/mL LPS-EB (Invivogen, tlrl-3pelps), and 100 ng/mL PMA

(InvivoGen, tlrl-pma) treatment, as indicated. Cells were washed with 1 × PBS and lysed with PLB according to the manufacturer's instruction for 15 min (Promega, E1910). Cells were transferred to a 96-well plate, 20 µL/well. After added with 100 µL/well LAR II, the sample was measured by the luciferase activity using the microplate reader (Thermo, Varioskan LUX).

Immunofluorescence

For frozen section, mammary gland was dissected and immediately fixed using 4% formalin for 2 hrs followed by PBS washing and 30% sucrose infiltration overnight. The fixed mammary tissue was then embedded in O.C.T. compound (Tissue-Tek) and frozen at −80 °C. Frozen tissue block was sectioned to 14 µm at −35 °C using Cryostat Leica CM3050 S (LEICA). For immunofluorescence assay, frozen sections were rehydrated with PBS for 10 min at RT. Sections were blocked with TBS + 0.1% Triton X-100 + 2% BSA + 10% Donkey serum for 1 hr at RT, and then stained with primary antibody mouse anti-$p16^{Ink4a}$ (Santa Cruz, sc-1661), Rb anti-p65 (Cell signaling technology, 8242), Rat anti-Bcl11b (Abcam, ab18465) 1:200 overnight at 4 °C. After washed by TBST (TBS + 0.1%Triton X-100) three times, sections were stained with secondary antibody Donkey anti-mouse, rat, rabbit 1:200 (Jackson ImmunoResearch) for 1 hr at RT. After 3 × TBST washing and brief 1 µg/mL DAPI staining, sections were mounted with Antifade Mounting Medium (Beyotime).

Immunohistochemistry and HE staining

Mammary tissue was collected and immediately fixed using 4% formalin overnight at 4 °C and dehydrated by gradient ethanol solution (70%, 85%, 95%, 100%). Dehydrated tissue was infiltrated by Xylene solution and embedded with paraffin. Tissue block was sectioned to 5 µm using Rotary Microtome Leica RM2255 (LEICA). Paraffin section was de-paraffinized using Xylene and rehydrated followed by gradient ethanol solution (100%, 95%, 85%, 70%, 0%) and subjected to immunohistochemistry staining according to the Histostain-Plus IHC Kit (NeoBioscience, ENS003.120, ENS004.300). Antigen was retrieved in citrate buffer (10 mM Sodium Citrate, 0.05% Tween 20, pH 6.0) for 20 min at 100 °C in microwave. The sections were treated with 3% H_2O_2 for 10 min, washed and blocked for 1 h, sections were incubated with Rb anti-p-p65 (Abcam, ab131100, 1:50), Rb anti-Il-6 (NOVUS, NB600-1131, 1:100), Mouse anti-Ssea1 (Abcam, ab16258, 1:200) and Rb anti-Oct4 (Abcam, ab19857, 1:200) overnight at 4 °C. Then sections were washed using TBST, secondary antibody incubation, HPR incubation, DAB incubation, Hematoxylin dyeing, gradient dehydration and mounted by CV5030 CoverSlipper (LEICA). Images were obtained using Eclipse Ti2 inverted microscope (Nikon).

For the HE staining assay, after de-paraffinize and rehydration, paraffin section was stained using ST5020 muti-stainer (LEICA) and mounted by CV5030 CoverSlipper (LEICA).

ChIP-seq

8×10^7 pInducer *Bcl11b* Comma D beta cells were treated with 100 ng/mL Doxycycline overnight and harvested. Rabbit anti-IgG (Abcam, ab172730) and rabbit anti-Bcl11b (Benthyl laboratories.inc, A300-384A) were used for ChIP-seq pull-down. Briefly, after cross-linked by 1% (wt/vol) formaldehyde solution, cells were quenched by glycine (0.12 M), washed one time with PBS and resuspended in PBS. Cells were lysed using SDS lysis buffer (50 mM Tris-HCl 8.0, 5 mM EDTA 8.0, 0.1%SDS and 1 × protease/phosphatase Inhibitor Cocktail (CST, 5872 S)). Then chromatins were sheared with AFA Focused-ultrasonicator using Covaris ME220 with 70 peak power, 20% duty factor, 14 average power for 3 min at 1×10^7 cells/tube. Add 9 × ChIP dilution buffer (50 mM Tris-HCl 8.0, 167 mM NaCl, 0.11% Triton X-100, 0.11% Sodium Deoxycholate and 1 × protease/phosphatase Inhibitor Cocktail) to the sonicated chromatin. 10% of the slurry was taken as input and 90% sonicated chromatin were divided equally to two parts. The sonicated chromatin was incubated in the cold room overnight with 50 µL protein G-Dynabeads (Invitrogen,10004D) which had been conjugated with 50 µg appropriate IgG or Bcl11b antibody. Beads were then washed with RIPA buffer 1(50 mM Tris-HCl 8.0, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.1% Sodium Deoxycholate and 1 × protease/phosphatase Inhibitor Cocktail), RIPA buffer 2 (50 mM Tris-HCl 8.0, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.1% Sodium Deoxycholate and 1 × rotease/phosphatase Inhibitor Cocktail), LiCl buffer (100 mM Tris-HCl 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% NP-40, 0.5% Sodium Deoxycholate and 1× protease/phosphatase Inhibitor Cocktail) and TE buffer (10 mM Tris-HCl 8.0, 1 mM EDTA), and were eluted into 200uL of ChIP direct elution buffer (10 mM Tris-HCl 8.0, 300 mM NaCl, 5 mM EDTA, 0.5% SDS, 0.2% Sodium Deoxycholate). Samples were reverse cross-linked at 65 °C overnight, treated with 4uL RNase A (1 mg/mL) at 37 °C for 30 min, incubated with 1 uL proteinase K (10 mg/mL) at 55 °C for 1 h, and extracted by phenol/chloroform. ChIP DNA library was constructed by VAHTS Universal DNA Library Prep Kit for Illumina V2 (Vazyme, ND606). Briefly, samples were subjected to adapter ligation, ChIP DNA was then amplified 12 cycles and purified using AMpure XP beads (Beckman, A63881) twice and then were submitted for150 bp paired-end sequencing on an Illumina novaseq 6000 platform (Novogene).

Single cell RNAseq library preparation and sequencing

A modified Smart-seq2 protocol was applied for single-cell RNA-seq[45,46] according to previously reported protocol. Briefly, single $CD49f^{high}EpCAM^{low}Lin^-$ cell from various age (2,4,9,11,13,15,17,19,22,24, 29 months old) mice or DMBA tumors were directly sorted into 96-well plate containing lysis buffer (0.05 uL RNase Inhibitor (40 U/µL), 0.095 uL 10% Triton X-100, 0.5 uL dNTP (10 mM), 0.1 uL ERCC (3×10^5) and 0.555 uL Nuclease-free water) using FACS. Single cell was immediately lysed at 70 °C for 3 min in the PCR system (TAdvanced 96SG, analytikjena). The sample was reverse transcribed to cDNA using SuperScriptIIreverse transcriptase (invitrogen, 18064-071) with a template switch oligo (TSO) primer and a sample-specific 25 nt oligo dT reverse transcription primer (TCAGACGTGTGCTCTTCCGATCTXX XXXXXX-NNNNNNNN-T25, X representing sample-specific barcode and N representing unique molecular identifier (UMI)). Then the cDNA was amplified by 18 cycles of PCR with 3'P2 primer and IS primer using KAPA HiFi HotStart Ready Mix (Kapa Biosystem, KK2602). After being pooled together and purified by AMpure XP beads (Beckman, A63881) twice, the barcoded DNAs were amplified using biotinylated pre-index primers by 4 cycles of PCR to introduce biotin tags to the 3′ ends of the amplified cDNAs. After purified by AMpure XP beads, cDNA was sonicated to ~300 bp fragments using Covaris ME220. 3′ terminal of the cDNA was enriched using Dynabeads® MyOne Streptavidin C1 beads (Invitrogen, 65001). The RNAseq libraries were constructed using the Kapa Hyper Prep Kit (Kapa Biosystem, KK8504) according to the manufacturer's instructions. Briefly, after end repair and A-tailing, Streptavidin conjugated DNA was ligated to the appropriate concentration adapter (1:10) (NEB, 7335 L). Then, after USER enzyme treatment and post-ligation cleanup, DNA was amplified with QP2 primer and short universal primer by 6 cycles of PCR and released from the streptavidin beads. Finally, AMpure XP beads were used to purify the DNA and DNA library quality was verified by Fragment Analyzer-12/96 (AATI) and then the DNA library was submitted to 150 bp paired-end sequencing on an Illumina NovaSeq 6000 platform (Novogene).

In vivo ageing clock rescue assay

To test TPCA-1's (GW683965) (Selleck, S2824)[75,76] effect on mammary ageing, 12 month-old virgin female WT mice or 3 month-old K14-cre $Bcl11b^{fl/fl}$ mice were treated by 10 mg/kg TPCA-1 or 4% DMSO in PBS intraperitoneally every day for 30 days. Mammary cells were then

dissociated and harvested for single cell RNAseq and pseudotime analysis.

To test BMS-345541 (Selleck, S8044)[88] effect on mammary ageing, 3 month-old K14-cre Bcl11b$^{fl/fl}$ mice were gavaged with 35 mg/kg BMS-345541 or 4% DMSO in PBS. Mammary cells were then dissociated and harvested for single cell RNAseq.

DMBA induced tumor formation

Mouse mammary tumors induced by MPA and DMBA assay were performed according to the previous published paper[55,56] with minor modifications. Briefly, the cleared fat pad of 3 week-old female mice (C57BL/6) were transplanted with WT or K14-cre Bcl11b$^{fl/fl}$ cells according to the transplantation assay methods described above in the Transplantation session. The recipient mice were implanted subcutaneously with a 50 mg 90 day-release MPA pellet (Innovative Research of America, NP-161-50 mg). Three weeks later, DMBA (200 µL, 5 mg/mL) (Sigma-Aldrich, D3254-1G) was administered by oral gavage 4 times throughout the following 5 weeks at -4, -3, -1, 0 weeks. Tumors were determined by manual palpation. Cancer incidence was calculated by the number of tumor cases within a designated latency period. The latency time was calculated from the last DMBA treatment day. Mice were euthanized before the tumor size reaching 2 cm in diameter according to our animal protocol 19-001-2-CS approved by IACUC (Institutional Animal Care and Use Committee of Westlake University).

To test TPCA-1's effect on DMBA induced tumor formation, female mice induced by MPA and DMBA, then followed by treatments with 10 mg/kg TPCA-1 or 4% DMSO in PBS intraperitoneally every day for 7 weeks. Tumors were determined by manual palpation. Cancer incidence was calculated by the number of tumor cases within a designated latency period. The latency time was calculated starting from the last DMBA treatment day.

To test the vulnerability of ageing mammary cells to cancer, we transplanted 3 month old young mixed with old mammary cells to syngeneic recipient mice. Briefly, to avoid the influence of the environmental stromal cells, we transplanted 2 month old young mammary cells (wt) with 24 month old mammary cells (with GFP reporter) to syngeneic recipient mice ($n = 11$) at equal MRUs and induced cancer formation by DMBA treatment. We repeat this experiment with 2 month old young mammary cells (with tdTomato reporter) and 12 month old mammary cells (wt). The recipient mice number is 22.

To test whether basal cells could be the origin of DMBA induced cancer using Krt14rtTA-TetOcre-mTmG mice ($n = 33$). We first labeled basal cells with doxycycline induction, and then treat mice with DMBA.

Single cell RNA-seq analysis

Raw reads were first processed using TrimGalore (Ver.0.6.7) (https://github.com/FelixKrueger/TrimGalore) to remove adapter sequences with paired end mode and default parameter. Quality control was evaluated with FastQC (Ver.0.11.9) (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). For each sequencing batch, the top 100 barcode ranked by read counts were retained using whitelist tool in UMI-tools (Ver.1.1.1)[89]. Then barcode and UMI information were extracted from read2 using extract tool in UMI-tools and added to read1. Subsequently, read1 were aligned to the mm10 genome using STAR aligner with default parameter except for the outFilterMultimapNmax = 1. Mapped reads were assigned to genes using featureCounts[89]. Finally, we used count tools in UMI-tools to generate count matrix, which the number of UMIs represents the transcript number of each gene within each individual cell.

After getting the count matrix, we applied four criterions to further exclude cells with low data quality: first, cells with barcode not included in the barcode sequence list were removed; second, cells with ERCC percentage larger than 10% or mitochondria percentage larger than 5% were filtered out; third, cells with gene number less than 200 or more than 6000 were removed; lastly, cells with stromal gene (Pecam1, Ptprc, Lyve1, Col1a1) expression level higher than 0.1 were removed. As for gene, we removed those detected in less than 10 cells. Finally, for young and aged single cell RNA-seq assay, 1981 cells passed the quality control. The same filtering criterion was used for K14-Cre Bcl11b$^{fl/fl}$, tumor and TPCA-1 treatment single cell RNA-seq assay. 149 K14-Cre Bcl11b$^{fl/fl}$ CD49f^{high}EpCAM^{low}Lin$^-$cells, 310 CD49f^{high}EpCAM^{low}Lin$^-$ tumor cells, 373 DMSO treated and 367 TPCA-1 treated CD49f^{high}EpCAM^{low}Lin$^-$ cells were used for analysis.

We used the Seurat (Ver 4.2.0)[90] to carry out the abovementioned filtering, data normalization, and all downstream analysis including dimensionality reduction, clustering, tSNE plot overlaying, and differential gene expression. More specifically, UMI counts in each cell were normalized with NormalizeData function with default parameter and vars.to.regress parameter was used to regress out the cell cycle, ERCC percentage, mitochondria percentage, batch and gene number effect. For dimension reduction, RunPCA function was used and the top 10 principal components were passed to tSNE analysis by RunTSNE function. Finally, clustering was performed by FindNeighbors and FindClusters functions with original Louvain algorithm and the resolution was set to 1.

Single cell trajectory analysis

Single cell trajectory was inferred by Monocle2 (Ver 2.24.1)[91]. Count matrix was input to Monocle2 and negative binomial distribution was used for building statistical distribution for read counts with lower detection limit set as 0.5. PhenoData was exported from Seurat object. Dimension reduction was done with DDRTree and the effect of Size_Factor, num_genes_expressed, ERCC percentage, mitochondria percentage, cell cycle and batcheffect were regressed out. The root of the pseudo-time trajectory was selected as the most abundant cell state in the cells from the 2 month-old batch. Genes significantly changed along pseudo-time were selected using differentialGeneTest function with qval < 0.05, resulting in 1932 differentially expressed genes. A new count matrix was generated from the original to include only differentially expressed genes. The count matrix was then smoothed using genSmoothCurves function, log transformed using a base of 10 and pseudo-count of 1 to prevent logarithm of zero value. All elements in the transformed count matrix were further truncated by a straightforward way that all elements larger than 3 or smaller than -3 were set as 3 or -3 respectively. Gene hierarchical clustering was performed on the transformed matrix with Heatmap in ComplexHeatmap package[92]. Genes were clustered into 4 clusters according to gene expression in cells along with pseudo-time increasing. Based on intersecting point of the average expression of genes in each cluster, cells were separated into 4 states. The same pipeline for combined aging & Bcl11b ko data and aging & TPCA-1 treatment data.

Pathway enrichment analysis

All gene symbols were mapped to their Entrez gene ids using biomaRt[93,94]. Then, both GO and KEGG pathway enrichment analysis were performed by clusterProfiler[95]. For simplicity, only biological process (BP) terms in GO were used for enrichment analysis. To do GSEA analysis, genes ordered according to decreased foldchange were fed to GSEA function. To do GSEA enrichment of Bcl11b-KO vs WT tumor CD49f^{high}EpCAM^{low}Lin$^-$ cells on 4 state marker genes, marker genes of each state were considered as one pathway. For enriched results from enrichGO, enrichKEGG or GSEA, only pathways with BH-adjusted p-value <0.05 were retained. Pathway activity of each cell was got with AddmoduleScore function in Seurat package[90]. When showing pathway activity along with pseudotime, the activity was also smoothed with genSmoothCurves function in monocle package[91].

Cell cycle state determination

We re-implemented the method previously used in Kowalczyk et al.[49] to determine the cell cycle state for each cell. First, filtered count matrix after Seurat was transferred into TPM using calculateTPM function in scater package[96] and then log-transformed by $\log_2(TPM+1)$. The cycle gene list in human was taken from the previously published paper by Whitfield[97]. All human gene symbols were transformed to mouse using biomaRt package[93,94]. Genes were filtered by the correlation with the average gene expression of corresponding cell cycle stage. To retain mammary gland specific cell cycle genes, the correlation threshold was set to 0.25. Finally, we got 14, 13, 21, 19, 13 genes for G1/S, G2, G2/M, M/G1 and S phase, respectively. The average cell cycle gene expression of each stage was calculated. We identified cells with G1/S score < 0 & G2/M < 0 as G0 cells. Other cells were defined as the corresponding stage according to gene maximum expression.

SASP gene score calculation

The SASP gene list in human was taken from the previously published paper by Coppe et al.[11]. All human gene symbols were mapped to mouse symbols by biomaRt package[93,94]. Then, the SASP score for each cell were calculated using AddModuleScore function in Seurat package[90].

Pathway score and transcription factor activity score calculation in human BRCA

BRCA fpkm data was downloaded with TCGAbiolinks[98], in which we only used the paired samples, then pathway score was calculated as mean expression of pathway genes. The result was shown with ggplot2 (https://ggplot2.tidyverse.org). Statistical analysis was performed using two-tailed paired *t*-tests. To calculate transcription factor activity, we used the transcription factor target genes from Transcription Factor Target Gene Database[99].

scEntropy analysis

We compute single cell entropy according to the public paper[58,100]. Briefly, single cell count expression profile from R was exported as.mat format which can be loaded into MATLAB. For calculating the entropy of cells from different ages, we constructed the gene co-expression network and apply it to Bcl11b ko cells. After computing the entropy of each cell, we visualize these results in R with ggplot2.

Transcription factor activity analysis

We got transcription factor target genes from the following three public databases: ENCODE[101], ChEA[102], and TRRUST v2[103]. For each cluster of genes changing along with pseudo-time, TFs were enriched with hypergeometric test which was implemented with phyper function in R. Transcription factor activity was calculated with transcription factor target genes from the combined databases using AddModuleScore function in Seurat package[90]. To calculate Bcl11b activity score, we did ChIP-seq assay for Bcl11b to find Bcl11b target genes. Based on scRNA-seq of wild type $CD49f^{high}EpCAM^{low}Lin^-$ cells and Bcl11b ko $CD49f^{high}EpCAM^{low}Lin^-$ cells from 4 month-old mice, we got Bcl11b positively regulated genes (genes with pval<0.05, ko <wt intersect with Bcl11b targets) and negatively regulated genes (genes with pval < 0.05, ko > wt intersect with Bcl11b targets). Then Bcl11b activity was computed as weighted average expression of the negatively regulated genes.

ChIP-seq data processing

ChIP-seq reads were trimmed with TrimGalore(https://github.com/FelixKrueger/TrimGalore) with paired end mode and default parameter and QC was done with FastQC(https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Trimmed reads were aligned to the mm10 mouse genome with bowtie2 (Ver.2.4.2) (https://github.com/BenLangmead/bowtie2) with default parameter. Then, the alignment was filtered using samtools[104] with the following parameter: -F 1804 -f 2 -q 30. Duplicated reads were marked with MarkDuplicates in picard-tools (http://broadinstitute.github.io/picard/) and filtered out with samtools. Finally, we found bcl11b binding sites on the genome based on input signal by MACS2 (Ver.2.2.7.1) using following parameter: -f BAMPE -g mm --keep-dup all --nomodel. Bcl11b binding sites were analyzed using ChIPseeker package[105]. ChIP-seq signals were visualized using Integrative Genomics Viewer (IGV) software[106].

Quantification and statistical analysis

Log-rank test was used between WT and K14-Cre $Bcl11b^{fl/fl}$ group tumor formation kinetics by GraphPad Prism 7.00. For limiting dilution analyses, the frequency of mammary repopulating unit was calculated using ELDA software[107]. Statistical analyses were performed using GraphPad Prism 7.00 with unpaired or paired two-tailed Student's *t*-test, as indicated in the figure legends. Bar graphs represent mean ± SD or mean ± SEM, as indicated. The box plots in Figs. 2b, e, f, 3b, l, r, 4k, l, 5i, 7e–g, and Supplementary Fig. 3a–b, 4a–c, 8e, 11d, 12b, 13b–d, 14 show the center line corresponds to the median, the lower and upper lines correspond to the first and third quartiles, and the whiskers extend to 1.5 times IQR (interquartile range), and each dot represents one cell. For Supplementary Fig. 1d-1g, stat_smooth was used to get non-linear regression with method set to loess. Pvalue was added by stat_regline_equation. For Supplemental Data 4, Supplemental Data 5, and Supplemental Data 6, pathways with p.adjust(BH-adjusted) < 0.05 were considered as significantly enriched.

Reporting summary

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

Data availability

scRNA-seq and ChIP-seq data were deposited in GEO database under the accession number GSE195647. To review GEO accession GSE195647: Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE195647. TCGA data is from https://www.cancer.gov/ccg/research/genome-sequencing/tcga. The datasets used to perform transcription factor enrichment analysis are from the following site: ENCODE, https://maayanlab.cloud/Harmonizome/dataset/ENCODE+Transcription+Factor+Binding+Site+Profiles; TRRUST v2, https://www.grnpedia.org/trrust; CHEA, https://maayanlab.cloud/chea3/. The mouse genome mm10(M25) was obtained from GENCODE database. Source data are provided as a Source Data file (https://doi.org/10.6084/m9.figshare.25521706). The remaining data are available within the Article, Supplementary Information or Source Data file.

Code availability

The custom code used for this paper is available on GitHub (https://github.com/Liuxiaoqin95/MaSC_aging).

References

1. Sung, H. et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer J. Clin.* **71**, 209–249 (2021).
2. Kocarnik, J. M. et al. Cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life years for 29 cancer groups from 2010 to 2019: a systematic analysis for the global burden of disease study 2019. *JAMA Oncol.* **8**, 420–444 (2022).
3. DePinho, R. A. The age of cancer. *Nature* **408**, 248–254 (2000).
4. Frank, S. A. In *Dynamics of Cancer: Incidence, Inheritance, and Evolution* 1st edn, 400, pp. 17–20 (Princeton University Press, 2007).

5. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2021. *CA: A Cancer J. Clin.* **71**, 7–33 (2021).
6. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88**, 593–602 (1997).
7. Beauséjour, C. M. et al. Reversal of human cellular senescence: roles of the p53 and p16 pathways. *Embo J.* **22**, 4212–4222 (2003).
8. Dimri, G. P. et al. A biomarker that identifies senescent human cells in culture and in aging skin in vivo. *Proc. Natl Acad. Sci. USA* **92**, 9363–9367 (1995).
9. Narita, M. et al. Rb-Mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* **113**, 703–716 (2003).
10. Zhang, W. et al. Aging stem cells. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. *Science* **348**, 1160–1163 (2015).
11. Coppe, J. P. et al. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol.* **6**, 2853–2868 (2008).
12. Ma, S. et al. Caloric restriction reprograms the single-cell transcriptional landscape of Rattus norvegicus aging. *Cell* **180**, 984–1001.e1022 (2020).
13. Akala, O. O. et al. Long-term haematopoietic reconstitution by Trp53-/-p16Ink4a-/-p19Arf-/- multipotent progenitors. *Nature* **453**, 228–232 (2008).
14. Freund, A., Orjalo, A. V., Desprez, P. Y. & Campisi, J. Inflammatory networks during cellular senescence: causes and consequences. *Trends Mol. Med.* **16**, 238–246 (2010).
15. Baar, M. P. et al. Targeted apoptosis of senescent cells restores tissue homeostasis in response to chemotoxicity and aging. *Cell* **169**, 132–147.e116 (2017).
16. Adams, P. D., Jasper, H. & Rudolph, K. L. Aging-induced stem cell mutations as drivers for disease and cancer. *Cell Stem Cell* **16**, 601–612 (2015).
17. Sperka, T., Wang, J. & Rudolph, K. L. DNA damage checkpoints in stem cells, ageing and cancer. *Nat. Rev. Mol. Cell Biol.* **13**, 579–590 (2012).
18. Martincorena, I. & Campbell, P. J. Somatic mutation in cancer and normal cells. *Science* **349**, 1483–1489 (2015).
19. Vogelstein, B. et al. Genetic alterations during colorectal-tumor development. *N. Engl. J. Med.* **319**, 525–532 (1988).
20. Rozhok, A. I. & DeGregori, J. The evolution of lifespan and age-dependent cancer risk. *Trends cancer* **2**, 552–560 (2016).
21. Christensen, K. et al. Cancer and longevity–is there a trade-off? a study of cooccurrence in Danish twin pairs born 1900-1918. *J. Gerontol. A Biol. Sci. Med Sci.* **67**, 489–494 (2012).
22. Coppé, J. P., Desprez, P. Y., Krtolica, A. & Campisi, J. The senescence-associated secretory phenotype: the dark side of tumor suppression. *Annu. Rev. Pathol.* **5**, 99–118 (2010).
23. Storer, M. et al. Senescence is a developmental mechanism that contributes to embryonic growth and patterning. *Cell* **155**, 1119–1130 (2013).
24. Muñoz-Espín, D. et al. Programmed cell senescence during mammalian embryonic development. *Cell* **155**, 1104–1118 (2013).
25. Jun, J. I. & Lau, L. F. The matricellular protein CCN1 induces fibroblast senescence and restricts fibrosis in cutaneous wound healing. *Nat. Cell Biol.* **12**, 676–685 (2010).
26. Demaria, M. et al. An essential role for senescent cells in optimal wound healing through secretion of PDGF-AA. *Dev. cell* **31**, 722–733 (2014).
27. Ritschka, B. et al. The senescence-associated secretory phenotype induces cellular plasticity and tissue regeneration. *Genes Dev.* **31**, 172–183 (2017).
28. Mosteiro, L. et al. Tissue damage and senescence provide critical signals for cellular reprogramming in vivo. *Science* **354**, aaf4445 (2016).
29. Chiche, A. et al. Injury-induced senescence enables in vivo reprogramming in skeletal muscle. *Cell Stem Cell* **20**, 407–414.e404 (2017).
30. Milanovic, M. et al. Senescence-associated reprogramming promotes cancer stemness. *Nature* **553**, 96–100 (2018).
31. Tabula Muris, C. A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. *Nature* **583**, 590–595 (2020).
32. Li, C. M. et al. Aging-associated alterations in mammary epithelia and stroma revealed by single-cell RNA sequencing. *Cell Rep.* **33**, 108566 (2020).
33. Zhang, M. J., Pisco, A. O., Darmanis, S. & Zou, J. Mouse aging cell atlas analysis reveals global and cell type-specific aging signatures. *eLife* **10**, e62293 (2021).
34. Ximerakis, M. et al. Single-cell transcriptomic profiling of the aging mouse brain. *Nat. Neurosci.* **22**, 1696–1708 (2019).
35. Kimmel, J. C. et al. Murine single-cell RNA-seq reveals cell-identity- and tissue-specific trajectories of aging. *Genome Res.* **29**, 2088–2103 (2019).
36. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111 (2001).
37. Visvader, J. E. Cells of origin in cancer. *Nature* **469**, 314–322 (2011).
38. Stingl, J. et al. Purification and unique properties of mammary epithelial stem cells. *Nature* **439**, 993–997 (2006).
39. Shackleton, M. et al. Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84–88 (2006).
40. Fu, N. Y. et al. Identification of quiescent and spatially restricted mammary stem cells that are hormone responsive. *Nat. Cell Biol.* **19**, 164–176 (2017).
41. Cai, S. et al. A quiescent Bcl11b high stem cell population is required for maintenance of the mammary gland. *Cell Stem Cell* **20**, 247–260.e245 (2017).
42. Chakrabarti, R. et al. Notch ligand Dll1 mediates cross-talk between mammary stem cells and the macrophageal niche. *Science* **360**, eaan4153 (2018).
43. Wang, D. et al. Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* **517**, 81–84 (2015).
44. Picelli, S. et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).
45. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
46. Li, L. et al. Single-cell RNA-seq analysis maps development of human germline cells and gonadal niche interactions. *Cell Stem Cell* **20**, 858–873.e854 (2017).
47. Lodato, M. A. et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* **359**, 555–559 (2018).
48. Schumacher, B., Pothof, J., Vijg, J. & Hoeijmakers, J. H. J. The central role of DNA damage in the ageing process. *Nature* **592**, 695–703 (2021).
49. Kowalczyk, M. S. et al. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Res.* **25**, 1860–1872 (2015).
50. Tyner, S. D. et al. p53 mutant mice that display early ageing-associated phenotypes. *Nature* **415**, 45–53 (2002).
51. Adler, A. S. et al. Motif module map reveals enforcement of aging by continual NF-kappaB activity. *Genes Dev.* **21**, 3244–3257 (2007).
52. Hoare, M. et al. NOTCH1 mediates a switch between two distinct secretomes during senescence. *Nat. Cell Biol.* **18**, 979–992 (2016).
53. Schmitt, C. A. The persistent dynamic secrets of senescence. *Nat. Cell Biol.* **18**, 913–915 (2016).

54. Pereira, B. et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat. Commun.* **7**, 11479 (2016).
55. Aldaz, C. M., Liao, Q. Y., LaBate, M. & Johnston, D. A. Medroxyprogesterone acetate accelerates the development and increases the incidence of mouse mammary tumors induced by dimethylbenzanthracene. *Carcinogenesis* **17**, 2069–2072 (1996).
56. Gonzalez-Suarez, E. et al. RANK ligand mediates progestin-induced mammary epithelial proliferation and carcinogenesis. *Nature* **468**, 103–107 (2010).
57. Abba, M. C. et al. DMBA induced mouse mammary tumors display high incidence of activating Pik3caH1047 and loss of function Pten mutations. *Oncotarget* **7**, 64289–64299 (2016).
58. Jin, S., MacLean, A. L., Peng, T. & Nie, Q. scEpath: energy landscape-based inference of transition probabilities and cellular trajectories from single-cell transcriptomic data. *Bioinformatics* **34**, 2077–2086 (2018).
59. Liu, Z. et al. Large-scale chromatin reorganization reactivates placenta-specific genes that drive cellular aging. *Dev. cell* **57**, 1347–1368.e1312 (2022).
60. Li, J. et al. Caenorhabditis elegans HCF-1 functions in longevity maintenance as a DAF-16 regulator. *PLoS Biol.* **6**, e233 (2008).
61. Lahoute, C. et al. Premature aging in skeletal muscle lacking serum response factor. *PLoS One* **3**, e3910 (2008).
62. Stein, B. & Yang, M. X. Repression of the interleukin-6 promoter by estrogen receptor is mediated by NF-kappa B and C/EBP beta. *Mol. Cell Biol.* **15**, 4971–4979 (1995).
63. Chicas, A. et al. H3K4 demethylation by Jarid1a and Jarid1b contributes to retinoblastoma-mediated gene silencing during cellular senescence. *Proc. Natl Acad. Sci. USA* **109**, 8971–8976 (2012).
64. Avram, D. et al. Isolation of a novel family of C(2)H(2) zinc finger proteins implicated in transcriptional repression mediated by chicken ovalbumin upstream promoter transcription factor (COUP-TF) orphan nuclear receptors. *J. Biol. Chem.* **275**, 10315–10322 (2000).
65. Krasteva, V. et al. The BAF53a subunit of SWI/SNF-like BAF complexes is essential for hemopoietic stem cell function. *Blood* **120**, 4720–4732 (2012).
66. Cismasiu, V. B. et al. BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter. *Oncogene* **24**, 6753–6764 (2005).
67. Liang, Y., Van Zant, G. & Szilvassy, S. J. Effects of aging on the homing and engraftment of murine hematopoietic stem and progenitor cells. *Blood* **106**, 1479–1487 (2005).
68. Morrison, S. J., Wandycz, A. M., Akashi, K., Globerson, A. & Weissman, I. L. The aging of hematopoietic stem cells. *Nat. Med.* **2**, 1011–1016 (1996).
69. de Haan, G., Nijhof, W. & Van Zant, G. Mouse strain-dependent changes in frequency and proliferation of hematopoietic stem cells during aging: correlation between lifespan and cycling activity. *Blood* **89**, 1543–1550 (1997).
70. Biteau, B., Hochmuth, C. E. & Jasper, H. JNK activity in somatic stem cells causes loss of tissue homeostasis in the aging Drosophila gut. *Cell Stem Cell* **3**, 442–455 (2008).
71. Dong, Q. et al. Aging is associated with an expansion of CD49f(hi) mammary stem cells that show a decline in function and increased transformation potential. *Aging* **8**, 2754–2776 (2016).
72. Kenyon, C. J. The genetics of ageing. *Nature* **464**, 504–512 (2010).
73. Mercurio, F. & Manning, A. M. NF-kappaB as a primary regulator of the stress response. *Oncogene* **18**, 6163–6171 (1999).
74. Taniguchi, K. & Karin, M. NF-kappaB, inflammation, immunity and cancer: coming of age. *Nat. Rev. Immunol.* **18**, 309–324 (2018).
75. Podolin, P. L. et al. Attenuation of murine collagen-induced arthritis by a novel, potent, selective small molecule inhibitor of IkappaB Kinase 2, TPCA-1 (2-[(aminocarbonyl)amino]-5-(4-fluorophenyl)-3-thiophenecarboxamide), occurs via reduction of proinflammatory cytokines and antigen-induced T cell Proliferation. *J. Pharm. Exp. Ther.* **312**, 373–381 (2005).
76. Nan, J. et al. TPCA-1 is a direct dual inhibitor of STAT3 and NF-kappaB and regresses mutant EGFR-associated human non-small cell lung cancers. *Mol. Cancer Ther.* **13**, 617–629 (2014).
77. Marongiu, F. & DeGregori, J. The sculpting of somatic mutational landscapes by evolutionary forces and their impacts on aging-related disease. *Mol. Oncol.* **16**, 3238–3258 (2022).
78. Nordling, C. O. A new theory on cancer-inducing mechanism. *Br. J. Cancer* **7**, 68–72 (1953).
79. Armitage, P. & Doll, R. The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br. J. Cancer* **8**, 1–12 (1954).
80. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
81. Vijg, J., Busuttil, R. A., Bahar, R. & Dollé, M. E. Aging and genome maintenance. *Ann. N. Y. Acad. Sci.* **1055**, 35–47 (2005).
82. Vincze, O. et al. Cancer risk across mammals. *Nature* **601**, 263–267 (2022).
83. Daher, M. T., Bausero, P., Agbulut, O., Li, Z. & Parlakian, A. Bcl11b/Ctip2 in Skin, Tooth, and craniofacial system. *Front Cell Dev. Biol.* **8**, 581674 (2020).
84. Sage, J., Miller, A. L., Pérez-Mancera, P. A., Wysocki, J. M. & Jacks, T. Acute mutation of retinoblastoma gene function is sufficient for cell cycle re-entry. *Nature* **424**, 223–228 (2003).
85. Lee, S. & Schmitt, C. A. The dynamic nature of senescence in cancer. *Nat. Cell Biol.* **21**, 94–101 (2019).
86. Prater, M., Shehata, M., Watson, C. J. & Stingl, J. Enzymatic dissociation, flow cytometric analysis, and culture of normal mouse mammary tissue. *Methods Mol. Biol.* **946**, 395–409 (2013).
87. Campbell, S. M., Taha, M. M., Medina, D. & Rosen, J. M. A clonal derivative of mammary epithelial cell line COMMA-D retains stem cell characteristics of unique morphological and functional heterogeneity. *Exp. Cell Res.* **177**, 109–121 (1988).
88. Yang, J., Amiri, K. I., Burke, J. R., Schmid, J. A. & Richmond, A. BMS-345541 targets inhibitor of kappaB kinase and induces apoptosis in melanoma: involvement of nuclear factor kappaB and mitochondria pathways. *Clin. Cancer Res.* **12**, 950–960 (2006).
89. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
90. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e3529 (2021).
91. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
92. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
93. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
94. Durinck, S. et al. BioMart and bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**, 3439–3440 (2005).
95. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: A J. Integr. Biol.* **16**, 284–287 (2012).
96. McCarthy, D. J., Campbell, K. R., Lun, A. T. & Wills, Q. F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179–1186 (2017).

97. Whitfield, M. L. et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell* **13**, 1977–2000 (2002).
98. Colaprico, A. et al. TCGAbiolinks: an R/bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **44**, e71 (2016).
99. Plaisier, C. L. et al. Causal mechanistic regulatory network for glioblastoma deciphered using systems genetics network analysis. *Cell Syst.* **3**, 172–186 (2016).
100. Teschendorff, A. E. & Enver, T. Single-cell entropy for accurate estimation of differentiation potency from a cell's transcriptome. *Nat. Commun.* **8**, 15599 (2017).
101. Rouillard, A. D. et al. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database (Oxf.)* **2016**, baw100 (2016).
102. Lachmann, A. et al. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* **26**, 2438–2444 (2010).
103. Han, H. et al. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* **46**, D380–D386 (2018).
104. Li, H. et al. The sequence alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
105. Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
106. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
107. Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **347**, 70–78 (2009).

Acknowledgements

This work was supported by National Natural Science Foundation of China NSFC grant 32170803, 81872405. This work was supported by Westlake Education Foundation. We thank Dr. Fuchou Tang for sharing single cell RNAseq protocol. We thank Peng Guo, Dandan Liao and Yanan Tang at the Flow Cytometry Facility of Westlake University for assistance in FACS sorting. We thank Haiyan Zhang at the Genomics Core Facility of Westlake University for assistance in library quality testing. We thank Shuying Li for assistance with the TPCA-1 treatment in Bcl11b KO mice. We thank Tingting Dong and Yue Zhang for their assistance in p16 knockout cell construction. We thank the Westlake Animal Facility for their support in mouse husbandry. We are grateful for Dr. Dangsheng Li, Dr. Zhen Qi for valuable discussions and suggestions.

Author contributions

S.C. and H.B. conceptualized the project and developed the methodology; H.B. performed most of the experiments; S.C. and H.B. conceptualized the bioinformatic analysis. X.L. and N.L. performed scRNA-seq and ChIP-seq data analysis; Y.M. and M.L. conducted the experiments for revision, including immunofluorescence, and immunohistochemical analysis; R.T. executed tumor induction assay in Krt14rtTA-TetOcre-mTmG mice; Y.G performed mouse genotyping; M.F.C. revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-49106-2.

Correspondence and requests for materials should be addressed to Shang Cai.

Peer review information *Nature Communications* thanks Martin Zhang and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.