Heliyon 10 (2024) e35529

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

The uncertainties and certainties of gene transcription in a human tumor cell

Yinchun Lv [a], Yulin Chen [a], Xue Li [a], Siying Li [a], Qiaorong Huang [a], Ran Lu [a,b], Junman Ye [a], Wentong Meng [a], Xiaolong Chen [a], Xianming Mo [a,*]

[a] *Department of General Surgery, Gastric Cancer Center, Laboratory of Stem Cell Biology, West China Hospital, Sichuan University, Chengdu, China*
[b] *Department of Urology and Pelvic Surgery, West China-PUMC C.C. Chen Institute of Health, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China*

A B S T R A C T

Previously we have identified that the expression number and levels of oncogenes and antioncogenes are highly positively or negatively associated with major cellular progress in a cancer cell. However, we have not defined any cellular potentials of a human tumor cell at the level of the overall gene expression. Here, we counted the overall number of expression genes and overall counts of mRNA in depth and revealed that the expression levels of mRNA were directly associated with the expression number of genes in a human tumor cell. Gene expression networks revealed steady states of tricarboxylic acid (TCA) cycle and ATP production, differentiation potentials that might be disturbed and blocked by uncertain gene expressing networks, and potential capabilities to undergo epithelial-mesenchymal transition (EMT), neurogenesis, angiogenesis, inflammatory response, immune evasion, and metastasis in a human tumor cell. Our analysis identifies unpredictable gene expression characteristics in human tumor cells. The results might profoundly influence mechanisms how a human tumor cell generates and undergoes its progresses.

1. Introduction

Tumor related cellular specialization plays a central role in controlling behaviors of the tumor tissues. On a whole tumor tissue level, cells of different classes (for example, tumor cells, fibroblasts, macrophages, lymphocytes, and endothelial cells) interact to maintain tumor mass and enable the progression of tumor tissues in human body [1,2]. On a tumor cell level, subtle specializations can control behaviors that distinguish tumor cells, including proliferation [3], resisting cell death [4], altered autophagy [5], reassessing senescence [6], activating invasion and metastasis [7], conducting epithelial-mesenchymal transition (EMT) [8], gaining self-renewal capabilities and stemness [9], promoting inflammation [10], reprogramming energy metabolism [11], and inducing angiogenesis [12] and neurogenesis [13]. Some tumor cell populations have been characterized in detail; many remain uncharacterized or have yet to be discovered [14,15]. To date, high-throughput single-cell RNA-seq (scRNA-seq) discloses plenty of cell-type classifications in the tumor tissues and reveals the RNA repertoires of every tumor cell type and state and identifies molecular markers for each population of tumor cells [16,17]. The systemic endeavors would help to understand the roles and interactions of cells in tumor tissue, including the roles of distinct tumor cell types in the disease.

Gene expression networks (GENs) or gene expression patterns in a tumor cell provide hierarchic evaluation of cellular behaviors in tumor tissues [18,19]. GENs generate and guide specific roles of cells involved in the initiation and progression of tumor tissues, including tumor tissue genesis and formation of tumor cell types [20]. GENs also determine the main events of activating invasion and metastasis [21]. Beyond that, GENs control a vast array of capabilities and modes of response of a tumor cell to environmental

* Corresponding author.
E-mail address: xmingmo@scu.edu.cn (X. Mo).

https://doi.org/10.1016/j.heliyon.2024.e35529
Received 11 May 2024; Received in revised form 25 July 2024; Accepted 30 July 2024
Available online 31 July 2024
2405-8440/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

fluctuations and challenges [22]. Thus, gene expression plays a key role in a wide variety of core biological processes in tumor cells, ranging from tumor tissue development, homeostasis, and tumor cell specialization to cellular stress response [23]. Previously, we have analyzed single-cell sequencing datasets of human hepatocellular carcinoma and identified that the expression number and levels of oncogenes and antioncogenes are highly positively or negatively associated with many major cellular progresses, including proliferation, apoptosis, autophagy, EMT, angiogenesis, neurogenesis, and cellular differentiation [24]. These findings indicate that the overall transcriptional activation of oncogenes and antioncogenes transform the activation of genes involved in these cellular processes. In addition, we identified that the overall expression levels of genes is weakly inked to the expression number and levels of oncogenes and antioncogenes. In the present work, we counted total expression number of genes and total counts of mRNA in depth to define the gene transcriptional activation in each human tumor cell and identified unpredictable gene expression characteristics in human tumor cells.

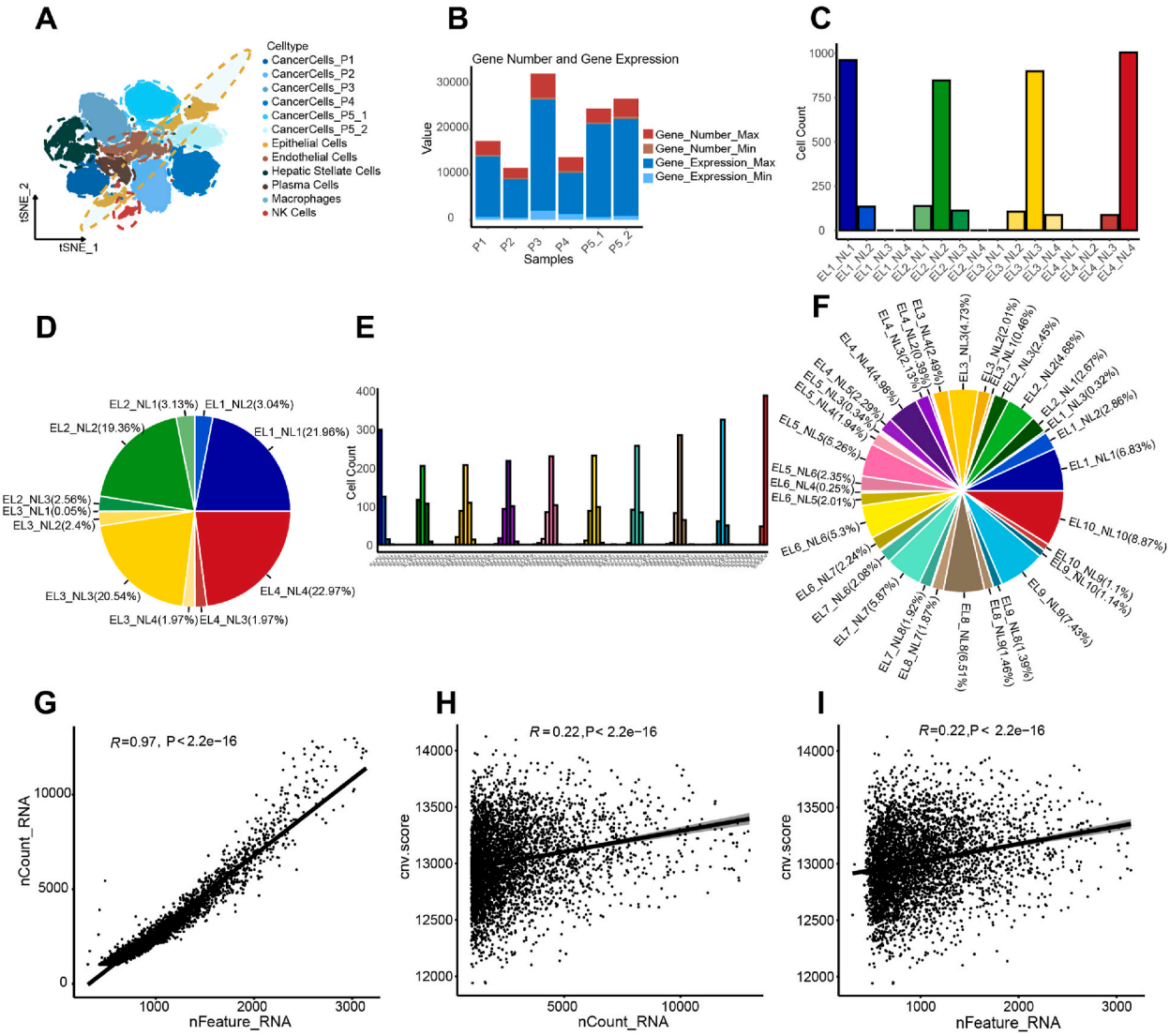


Fig. 1. Association of the copies of mRNA and the gene expression number in a human tumor cell. A. The tSNE shows all types of cells derived into 6 HCC samples. **B.** The stacked bar charts show max/min gene expression numbers and levels in different HCC samples. **C.D.E.F.** The bar plots show cell counts for 16 and 100 gene classifications of the overall gene in the P1 sample of HCC, while pie charts display their percentages. **G.** The scatter plot shows the correlation between the gene expression numbers and gene expression levels. **H.I.** The scatter plots show the correlation between the copy number variation (CNV) score and the gene expression levels or gene expression levels, respectively. All plots with Pearson correlation coefficient (R-value) and *P*-value labeled at the top.

2. Results

2.1. Association of mRNA counts and the gene expression number in a human tumor cell

We collected single-cell sequencing datasets of 6 samples of human hepatocellular carcinoma from 5 patients as described [24]. Cells with more than 20 % mitochondrial gene expression were excluded from the datasets. We obtained single-cell transcriptomes of 30,429 tumor cells after quality control defined by gene expression patterns in each cell. Principle-component analysis (PCA) and t-distributed stochastic neighbor embedding (tSNE) analysis showed that human hepatocellular carcinoma (HCC) had high molecular heterogeneity of inter-samples and intra-sample (Fig. 1A), consistent with previous observations [24]. We counted total expression number of genes and total counts of mRNA in each HCC cell in each HCC sample. The expression levels and expression numbers of genes expressed in each single cancer cell were measured and counted separately in an HCC sample. For instance, we captured the maximum number of gene expressed was 3144 in a tumor cell and the minimum number of gene expressed was 307 in a tumor cell in P1 sample. In addition, the maximum count of mRNA was 13314 in a tumor cell and the minimum counts of mRNA was 655 in a tumor cell in the P1 sample. The largest expression number of genes in a tumor cell was about 10-fold higher, in comparison to the smallest expression number of genes in a tumor cell. The largest number of mRNA counts in a tumor cell were about 20-fold higher, in comparison to the smallest number of mRNA counts in a tumor cell (Fig. 1B). The mRNA count (expression level (EL)) and number level (NL) of genes were classified into four classes according to their respective quartiles. EL or NL- 1 to 4 indicated increasing expressing level or number of genes, respectively. Combining the EL and NL labels of genes respectively in an individual cell, the cancer cells of human HCC were theoretically clustered into 16 groups (Fig. 1C). The cancer cells were clustered into four large groups and a few small groups. The four large groups are EL1NL1 (22.0 %), EL2NL2 (19.4 %), EL3NL3 (20.5 %), EL4NL4 (23.0 %). The small groups consisted of EL1NL2 (3.0 %), EL2NL1 (3.1 %), EL2NL3 (3.0 %), EL3NL2 (2.4 %), EL3NL4 (2.0 %), EL4NL3 (2.0 %) (Fig. 1D). To examine the data further, we classified the expression level (EL) and number level (NL) of cancer cells into eight (data not shown) and ten classes according to their respective octuple and decile (Fig. 1E and F). The similar gene expression pattern in an HCC cell was obtained. We obtained the same gene expression pattern from further measurements from the rest of the HCC samples. Figure S1 A, B,

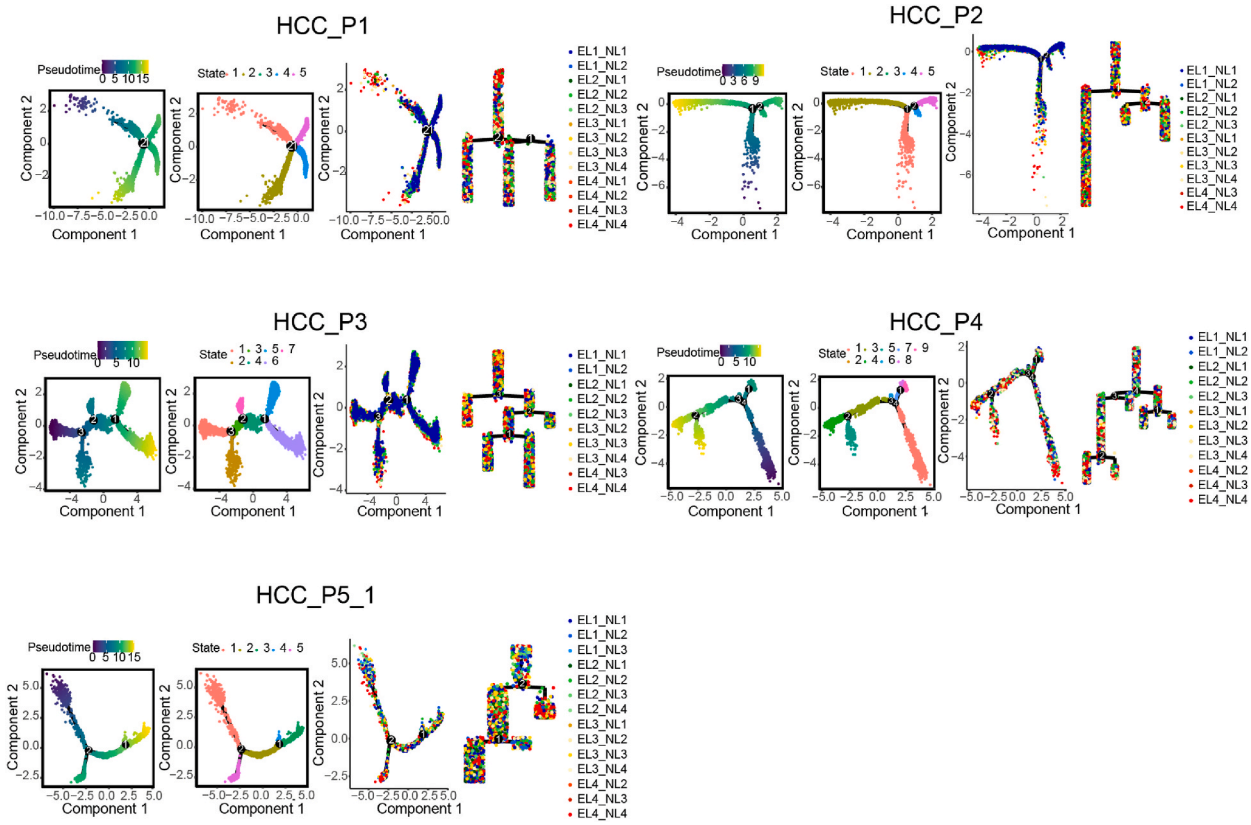


Fig. 2. Constitutive relationships of human tumor cells with differential gene expression. These five pseudotime plots display the potential differentiation trajectories of cancer cells for each HCC sample. The first two plots for each sample are derived from calculated pseudotime and states. The third plot maps the HCC cells onto the Monocle 2 inferred cell development trajectory based on a four-level classification of expression count and expression level, combined pairwise. Except for sample P1, the other samples do not show a clear differentiation trajectory under the four-level classification of expression count and expression level.

C, and D). In addition, the correlation analysis revealed that the mRNA count was directly associated with the gene number of expression in an HCC cancer cell (Fig. 1G and data not shown).

Comprehensive analysis of copy number variations (CNV) showed that the gene expressing regions of chromatins in HCC cancer cells of an individual sample manifest a unique pattern as we previously described [24]. A distinctive CNV pattern was visualized in each cancer cell in a sample of cancer tissues. In the present works, we analyzed the relationship between the CNV and overall gene expression in an HCC cell. The results showed that the CNVs were not associated with the overall gene expression, including the expression number of genes (Fig. 1I) and mRNA count number (Fig. 1H) in a cancer cell. Thus, the chromatin variations do not affect overall gene expression in a cancer cell.

To further visualize the gene expression patterns in a human tumor cell, we collected single-cell sequencing datasets from three human colorectal carcinoma (CRC) patients, with each patient corresponding to one sample. Cells with more than 20 % mitochondrial gene expression were excluded from the datasets. After quality control, single-cell transcriptomes of 5560 tumor cells were analyzed. The gene expression patterns were the same as the ones of HCC cells (Figs. S1E, F, and G). We also verified our findings in published

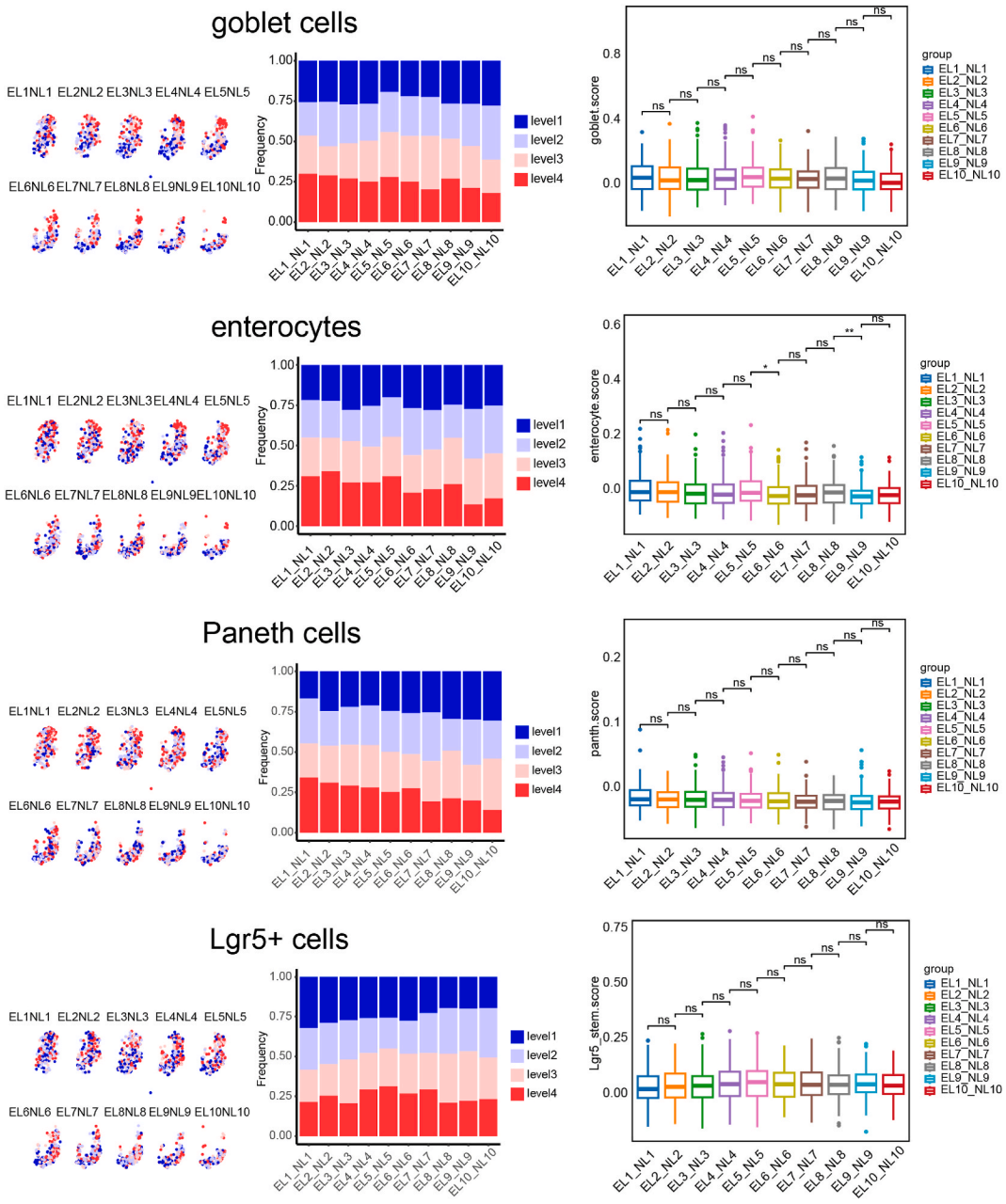


Fig. 3. Differentiational potential of a human tumor cell with a given gene expressing pattern.

datasets (Figs. S2A–H and data not shown), including published single cell sequencing data from 6 samples of human acute lymphoblastic leukemia [25] (Fig. S2A), 7 samples of human uveal melanoma [26] (Fig. S2B), 5 samples of human osteosarcoma [27] (Fig. S2C), 3 samples of human prostatic carcinoma [28] (Fig. S2D), 5 samples of human pancreatic ductal adenocarcinoma [29] (Fig. S2E), 4 samples of human breast cancer [30] (Fig. S2F), and 5 samples of human non-small cell lung cancer [31] (Fig. S2G). The same tendencies of gene expression patterns were identified in a human tumor cell. We also analyzed published datasets of single-cell chromatin accessibility assays of chronic lymphocytic leukemia [32] and revealed the same gene expression tendencies as the ones found in a human tumor cell (Fig. S2H). All results showed a tendency in the distribution of gene expression in terms of gene number rather than gene expression level among human tumor cells. Based on the number of genes expressed in an individual tumor cell, the mRNA level is associated with the expression number of gene in a cancer cell. Thus, the expression number of genes determines mRNA counts in a human tumor cell.

2.2. Constitutive relationships of human tumor cells with differential gene expression activations

The transcriptome expression data obtained from the large number of HCC cells allowed us to obtain insight into the functional and constitutive relationships among these cells. We explored the relationships and connectivity among HCC cells by inferring the state trajectory using Monocle2 in each HCC sample. The analysis showed that the types of HCC cells were aggregated into several cellular states in each HCC sample (Fig. 2). For instance, HCC cells of sample P1 displayed 5 cellular states and HCC cells of sample P3 were 7 cellular states. No distribution tendencies of the cancer cells labeled by different classifications of gene expression profiling in the HCC samples were manifested by the moncle trajectories, except for sample P1. In P1, the transcriptional activation in cancer cells showed somewhat distribution tendencies. The cancer cells in the cellular state 1 carried higher transcriptional activities. In contrast, the lower transcriptional activities were identified in cancer cells of cellular states 4 and 5. However, a significant fraction of cancer cells with higher gene expression levels were revealed to be in the cellular state 4 and 5. In the rest of HCC samples, the cancer cells with lower or higher transcriptional activations were identified in all the cellular states. We then analyzed the functional and constitutive relationships of cancer cells in the 3 CRC samples (Figs. S3A, B, and C). The results showed that cancer cells labeled by different classifications of gene expression networks in two samples displayed similar patterns as ones in HCC sample P1 and had somewhat distribution tendencies. Cancer cells of one CRC sample labeled by classifications of gene expression networks did not show any tendencies(Fig. S3A,B,C).

Next, we examined published datasets (Figs. S3D–J and data not shown), including published single cell sequencing data of human acute lymphoblastic leukemia (Fig. S3D), uveal melanoma (Fig. S3E), osteosarcoma (Fig. S3F), prostatic carcinoma (Fig. S3G), pancreatic ductal adenocarcinoma (Fig. S3H), breast cancer (Fig. S3I), and non-small cell lung cancer (Fig. S3J) to determine the state trajectory of tumor cells with the differential gene expressing profiling in each sample. Some samples, especially carcinoma samples, showed no distribution tendencies of the cancer cells with differential gene expression profiling. The rest samples analyzed displayed tendencies of the tumor cells labeled by different classifications of gene expressing profiling. Together, the results suggest that the total gene expression level or number may not define the cellular states of a human tumor cell.

2.3. Differentiational potential of a human tumor cell with a given gene expressing pattern

We detected the cellular statuses of cancer cells with differential expression activities of genes in HCC samples via gene expression patterns. We first measured the hepatocytic signatures of a cancer cell to ascertain the differentiation properties of cancer cells in HCC (Figs. S4 and S5). Based on the averaged expression of genes, the hepatocytic signatures were scored in each cancer cell. The distribution of the hepatocytic signatures showed a clear tendency in tSNE projection as previously described. A cancer cell with higher gene expression activities had more possibilities to carry differentiated hepatocytic signatures. To be more convincing, we scored the potentialities of cellular response to toxic substances, which represent the physiological functions of hepatocytes, in cancer cells and identified that toxic responses were highly correlated with hepatocytic signatures and gene expressing activities in an HCC cancer cell.

We analyzed the differentiation statuses of cancer cells with differential expression activities of genes in CRC samples via gene expression patterns (Fig. 3, S6-S7). The epithelial signatures of a CRC cancer cell were analyzed to ascertain the differentiation properties of cancer cells with differential gene expression. A CRC cancer cell with the signatures of goblet cells, enterocytes, and Paneth cells, all the differentiated epithelial cell types in colon mucosa, were identified in the populations of CRC cancer cells with any levels of gene expression. The cancer cells with the expression of Lgr5, a marker of intestinal stem cell, was also found in the populations of CRC cancer cells with any levels of gene expression (Fig. 3).

The tSNE projections and stacked bar charts show 4 levels of goblet cells, enterocytes, Paneth cells, and Lgr5+ cells potentialities, while the subsequent box plots illustrate the differences in these characteristics under varying gene expression counts and levels (Fig. 3). The tSNE projections are split by 10 categories of overall genes, formed by integrating 10 gene expression levels and 10 number levels, with lower levels consolidated into broader groups. The stacked bar charts show the 4 levels of differentiational potential in the P1 sample of CRC, and from blue to red indicating increasing levels. Statistical differences are indicated by asterisks. *p*-values were calculated by Wilcoxon test. ns: not significant. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$.

We analyzed published datasets of human B-acute lymphoblastic leukemia to define the differentiation status of leukemia cells (Fig. S8). The ratio of leukemia cells with marker gene expression of immature B cells or mature B cells were not significantly altered in the populations of leukemia cells with gene expression activations enhanced. However, the ratio of $CD38^-$ leukemia cells and leukemia cells with marker gene expression of hematopoietic stem cells were increased in the populations of leukemia cells with higher gene expression activities. In all samples of human B-acute lymphoblastic leukemia analyzed, we identified that a fraction of tumor cells did

not carry any gene expression networks involved in known cellular programs.

We verified our findings in published datasets of human tumors including uveal melanoma (Figure S9 A), osteosarcoma (Figure S9 B), prostatic carcinoma (Figure S9 C), pancreatic ductal adenocarcinoma (Figure S9 D), breast cancer (Figure S9 E), non-small cell lung cancer (Figure S9 F) and obtained the differentiation pattern in those tumor cells. All the data indicate that a human tumor cell carries potentials to undergo differentiation via the pathways of a cell that the human tumor cell is generated from.

2.4. Energy production in a human tumor cell with a given gene expressing pattern

We evaluated the gene expression networks involved in the processes of glycolysis, tricarboxylic acid (TCA) cycle, ATP production, and oxidative phosphorylation, key programs of energy metabolism in an HCC cancer cell (Figs. S10–13). When comparing changes in energy metabolism in cells, the results showed that increasing gene expression activations in a cancer cell greatly enhanced the activities of oxidative phosphorylation. Increasing gene expression activations significantly enhanced the activations of glycolysis and ATP production in a HCC cancer cell in some samples, but not all HCC samples. The gene expression activities did not altere the activities of tricarboxylic acid (TCA) cycle in a cancer cell in most HCC cells. However, in some samples, the HCC cells with high transcription activations might carry higher TCA activations.

We next evaluated the gene expression networks involved in the key programs of energy metabolism in a cancer cell of three CRC samples (Fig. 4, Figs. S14 and S15). The results showed that increasing gene expression activations in a cancer cell also greatly enhanced the activities of oxidative phosphorylation in a CRC cell (Fig. 4). In addition, enhanced activations of glycolysis might be detected in a CRC cell with higher transcription activities. In two samples, the gene expression network involved in ATP production was not significantly altered in a cancer cell with any level of transcriptional activation. In one sample, enhanced gene expression network involved in ATP production was identified in a CRC cell with high transcription activities. The gene expression activities did not alter the activities of tricarboxylic acid (TCA) cycle in a CRC cell in all samples (Fig. 4). The data indicate that a cancer cell always carries steady state of gene expression networks involved in ATP production and tricarboxylic acid (TCA) cycle regardless gene transcription activation. The measurements of published datasets of human tumors confirmed the observations (Figs. S16–22). Thus, most human tumor cells maintain steady state of energy supply in spite of transcription activities, especially the activities of gene expressing networks involved in glycolysis and oxidative phosphorylation.

2.5. Cellular programs of a human tumor cell with a given gene expressing pattern

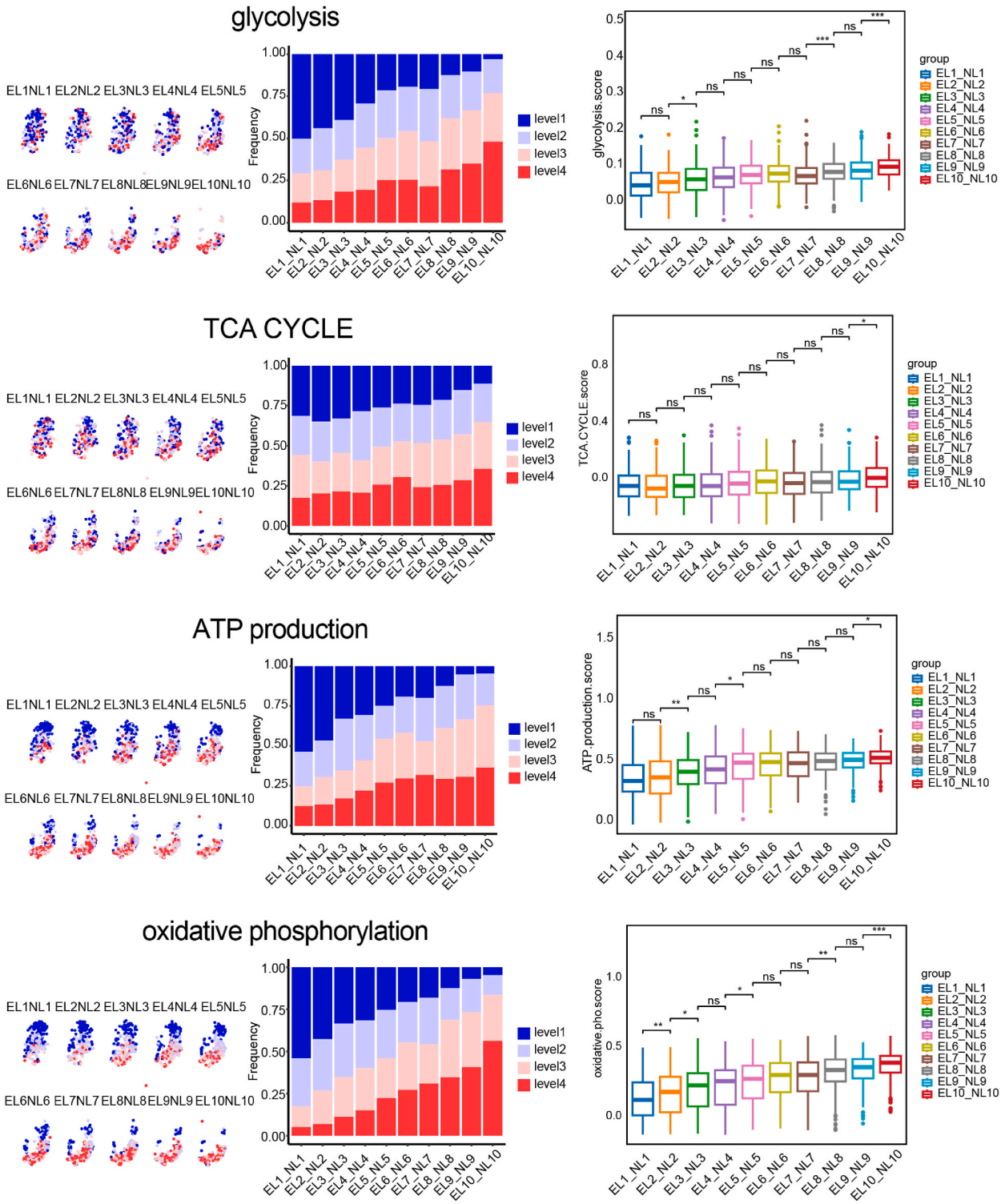
Next, we anayzed the proliferation, apoptosis, autophagy, stemness, and senescence statuses of cancer cells via gene expression patterns in an HCC cancer cell (Figs. S23–26). The results showed that those potentialities of cancer cells were heterogeneous among samples as described [24]. In all HCC samples, individual cancer cells increased their proliferative abilities when they exhibited higher gene expression activities. Additionally, in some samples, individual cancer cells showed increased possibilities of undergoing apoptosis and autophagy with higher gene expression activities. In the rest of the samples, the possibilities of individual cancer cells undergoing apoptosis and autophagy were similar regardless of their gene expression activities. In an HCC cancer, the stemness and senescence statuses of cancer cells via gene expression patterns showed that increasing gene expressing activities decreased or did not change the possibilities to display stemness and senescence statuses.

We analyzed the cellular statuses of cancer cells with differential expression activities of genes in CRC samples via gene expression patterns (Fig. 5, S27-S28). The potentials of proliferation, apoptosis, autophagy, stemness, and senescence statuses of a CRC cancer cell were found to be highly similar to the ones of an HCC cancer cell. The measurements of human tumor cells in published datasets also showed similar patterns to those observed in HCC cancer cells (Figs. S29–35). The results indicate that a human cancer cell could carry gene expression networks of multiple cellular programs. However, a given overall gene expression activation indefinitely defines cellular statuses of a human tumor cell.

2.6. Cellular reaction programs of a human tumor cell with a given gene expressing pattern

We measured the abilities of epithelial-mesenchymal transition (EMT), neurogenesis, angiogenesis, inflammatory response, immune evasion, migration, and metastasis in an HCC cancer cell with differential gene expressing activities (Figs. S36–39). The results showed that an HCC cancer cell with higher gene expression activities might have increased possibility to undergo neurogenesis and decreased ability to display the capabilities to undergo EMT and to stimulate angiogenesis. The gene expression activities in a cancer cell did not significantly alter the abilities of a cancer cell to display activities involved in its inflammatory response, immune evasion, migration, and metastasis.

We analyzed the cellular reactions of cancer cells with differential expression activities of genes in CRC samples via gene expression patterns (Fig. 6, S40-S41). The potentials of EMT, neurogenesis, angiogenesis, inflammatory response, immune evasion, migration, metastasis werer found to be highly similar to the ones of an HCC cancer cells. We verified our findings in published datasets of human B-acute lymphoblastic leukemia (Fig. S42). The results showed lymphocyte migration and immune response (including B cell activation and cytokine production) programs were not significantly altered in the populations of leukemia cells with gene expression activations enhanced. The measurements of other published datasets of human tumor cells confirmed the observations that a human tumor cell might carry the potential abilities to undergo EMT, neurogenesis, angiogenesis, metastasis and to stimulate inflammatory response and immune evasion (Figs. S43–48).



(caption on next page)

Fig. 4. Energy production in a human tumor cell with a given gene expressing pattern. The tSNE projections and stacked bar charts show 4 levels of glycolysis, TCA cycle, ATP production, and oxidative phosphorylation, while the subsequent box plots illustrate the differences in these characteristics under varying gene expression counts and levels. The tSNE projections were split by 10 categories of overall genes, formed by integrating 10 gene expression levels and 10 number levels, with lower levels consolidated into broader groups. The stacked bar charts show the 4 levels of energy production in the P1 sample of CRC, and from blue to red indicating increasing levels. Statistical differences are indicated by asterisks. *p*-values were calculated by Wilcoxon test. ns: not significant. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

2.7. The expression patterns of designed driven genes in tumor cells

We analyzed the expression patterns of the genes that were supposed to be driving genes in human cancers. We first measured published datasets of human non-small cell lung cancer cells with EGFR mutations [33] (Fig. 7A). The results showed that EGFR mRNAs were not detected in a significant fraction of cancer cells. A portion of cancer cells with any given levels of gene expression carried the high levels of EGFR mutant mRNA, suggesting that EGFR mutant gene expression is not dependent on the gene transcription activities in a non-small cell lung cancer cell. Next, we analyzed the published datasets of human acute lymphoblast leukemia cells with ETV6-RUNX1 fusing gene and revealed similar gene expressing patterns. A portion of leukemia cells with any given levels of gene expression carried high levels of ETV6 or RUNX1 mRNA (Fig. S49). To verify our results, we analyzed the expression of driving genes in mouse tumor induced by transgenes. The results obtained from Myc induced mouse prostatic carcinoma [34] cells showed that the expression of transgenic myc mRNA was detected in a fraction of cancer cells with any given gene expressing activation (Fig. 7C). Similar results of supposed driving gene expression were observed in tumor cells of MYCN-driven mouse neuroblastoma [35] (Fig. 7B). All the analyzing results are consistent with our previous observations [24] and show that the proposed driving gene expression in a cancer cell does not rely upon the transcriptional activities in a human tumor cell. The data also reveal that uncertainty of transcription activation occurs in a tumor cell.

3. Discussion

ScRNA-seq can generate very large datasets of thousands of gene transcripts in each individual cell. The scRNA-seq datasets can cluster cells according to their similarity and the expression of single or multiple genes and identify rare cell types, cell subtypes, disease-specific cell types and interactions among cells. Furthermore, computational analyses, such as pseudo-time diffusion mapping or RNA velocity, can provide cell linkages and trajectories between cell types or among cell subtypes [36]. To date, analytical pipelines for scRNA-seq datasets have matured and limitations of scRNA-seq assays were discovered. For instance, tissue dissociations and single cell captures are the potential to induce ectopic gene expression, especially the expression of stress response genes. The induced expression of genes can lead to mischaracterization of certain cell subpopulations [37]. However, overall analysis, not determination of specific gene involved in certain pathways, of scRNA-seq datasets can overcome the limitations of scRNA-seq assays. In this study, we analyze scRNA-seq datasets to provide fine map of gene expression networks (GENs) of a human individual tumor cell and to show that the number of gene expressed shapes the mRNA counts in a human tumor cell. The gene expression patterns in each human tumor cell display distinctive gene expressing networks. Designed orienting gene expression networks always exist in a human tumor cell.

There are two kinds of designed orienting gene expression networks existing in a human tumor cell. One is the gene expression networks for surviving. For instance, the gene expression networks of energy production are key programs for cell survival. We identify that there is steady gene expression

involved in tricarboxylic acid (TCA) cycle and ATP production, in spite of the alteration of the gene expressing activation and the gene expression networks responsible for glycolysis and oxidative phosphorylation. We also measure other gene expression networks (data not shown) involved in cell survival and housekeeping and obtain the similar phenotypes (data not shown). Another kind of oriented gene expression networks is required for differentiation programs no matter whether the gene transcription activations are high or low in a human tumor cell. Dedifferentiation, blocked differentiation, and transdifferentiation are features of human tumor cells. Human tumor cells always fail to conduct the conveyer belt of terminal differentiation to form functional mature cells [38]. Immature cells, undifferentiated cells, and cells with markers of other type cells can be detected in the human tumor tissues [39]. Suppression of differentiation factor expression are always detected in human tumor tissues and are shown to facilitate tumorigenesis by enabling more well-differentiated cells to dedifferentiate into incompletely differentiated cells [40]. Often-identified factors mediate alterations in differentiation programs during lineage development from stem cells and progenitors, thereby stopping the differentiating steps of an incipient cell that freeze an undifferentiated state in human tumors [41]. Our measurements show that there always existed gene expression networks involved in differentiation programs in a human tumor cell, regardless whether the cell has high transcriptional activation or low transcriptional activities. The human tumor cell seems to try to differentiate via the way its origin cell differentiates. However, the uncertain gene expression networks are always involved in multiple cellular behaviors that disturb and block its designated differentiation pathways in a human tumor cell. The disturbance might determine the features of a tumor cell to remain in the state of dedifferentiation, blocked differentiation, or transdifferentiation.

In a human tumor cell, mRNA levels are directly associated with the expression number of genes. Exclusive of the designed orienting gene expression networks, a given human tumor cell always carries a distinctive gene expression pattern despite gene transcriptional activations. We identify that the number of expressing genes and the mRNA levels are not correlated to the copy number variations in a human tumor cell. In addition, the patterns of chromatin accessibility are highly similar to the gene expression patterns in human tumor cells. The results suggest that distinctive gene expression pattern is not related to the unique copy number variations

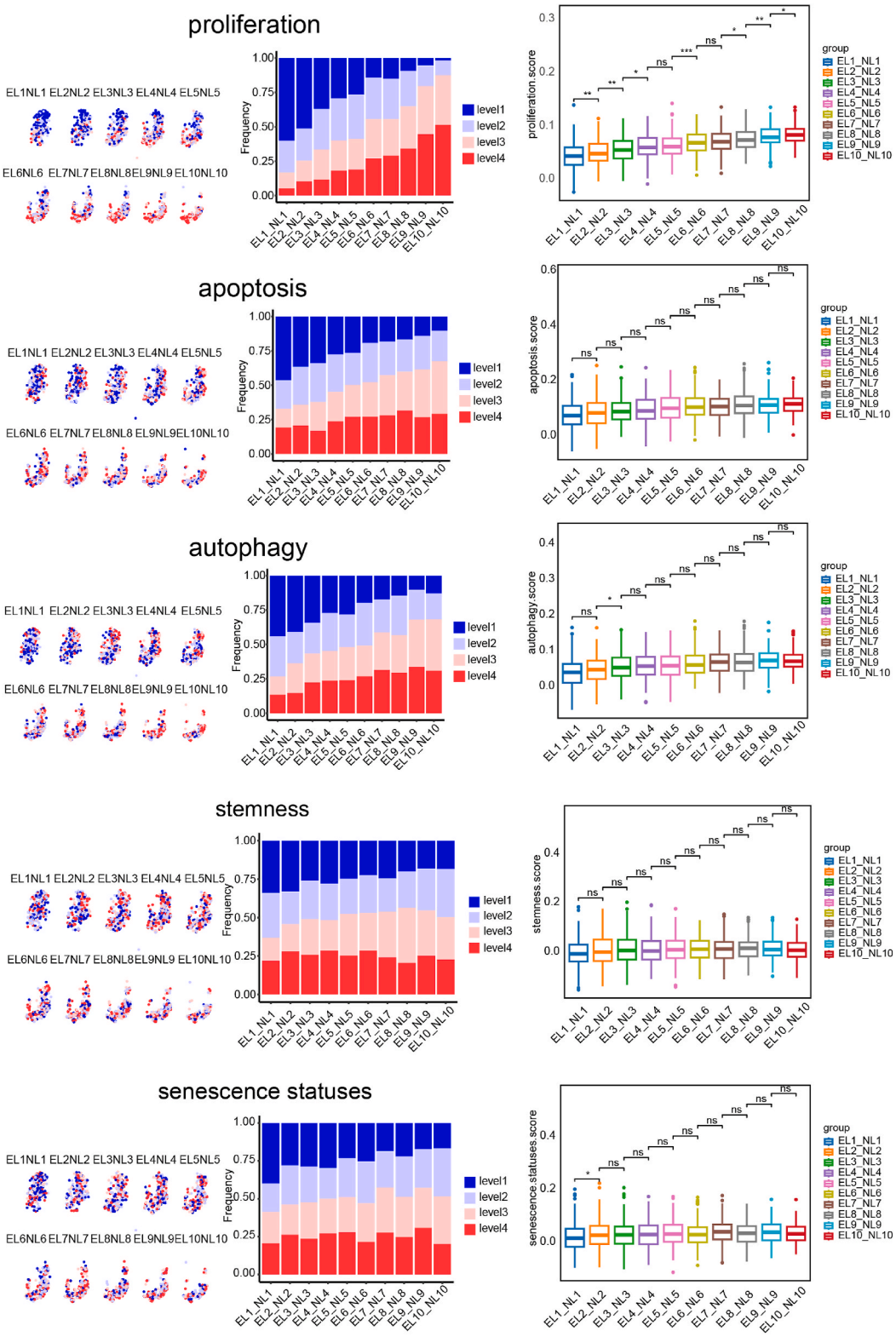


Fig. 5. Cellular programs of a human tumor cell with a given gene expressing pattern. The tSNE projections and stacked bar charts show 4 levels of proliferation, apoptosis, autophagy, stemness, and senescence statuses, while the subsequent box plots illustrate the differences in these characteristics under varying gene expression counts and levels. The tSNE projections were split by 10 categories of overall genes, formed by integrating 10 gene expression levels and 10 number levels, with lower levels consolidated into broader groups. The stacked bar charts show the 4

levels of ceullar programs in the P1 sample of CRC, and from blue to red indicating increasing levels. Statistical differences are indicated by asterisks. *p*-values were calculated by Wilcoxon test. ns: not significant. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

but might result from distinctive chromatin structures in an individual human tumor cell. Genomic and epigenetic mutations and alterations, including single-nucleotide variants (SNVs), small insertions or deletions, somatic mutations, gene copy number alterations, chromatin structural variants, chromatin regulator mutations, DNA methylation changes, and altered compositions of enhancers and repressors, and mutations in protein that modulate chromatin structure and activity of enhancers or repressors crossing genomes, are considered fundamental components of cancer genesis, formation, and progression [42,43]. It seems that genomic and epigenetic mutations and alterations organize, modulate, and maintain distinctive chromatin architectures in a human tumor cell. Thus, we propose that the chromatin architecture caused by genomic and epigenetic mutations and alterations in a human tumor cell are further modified by chromatin instabilities in a cancer cell, as well as signals from surrounding cells and elements include cancer cells, endothelial cells, neural cells, fibroblasts, and inflammatory cells and other elements in external microenvironment of a tumor cell. The modified chromatin architectures thereby globally control gene expression in a human tumor cell and are responsible for the distinctive gene expression of a given human tumor cell.

The discoveries here have profoundly influenced mechanisms of how cancer cells undergo its progresses and modify a conception we have previously proposed [24]. Rapid transformations in the expressions of genes are induced by incidental dynamics of chromatin architectures, which are caused by genomic and epigenetic mutations and alterations and microenvironment signals, in a human tumor cell [44,45]. The transformations of gene expression drives a human tumor cell to express balanced combinations of synergistic, agonistic, and antagonistic genes and undergoes a given biological process by a quick change in gene expression networks resulting from the simultaneous activation of a new gene expression program and the extinction of the old one. The finally outcoming gene expression networks control a human tumor cell to perform a specific cellular process in cancer tissues. Further experiments using single-cell mass cytometry and absolute quantification are required to capture the temporal dynamics of chromatin architectures and provide evidence to solve the mechanisms of how a cell selected becomes a cancer cell or how a cancer cell is progressing in cancer tissues.

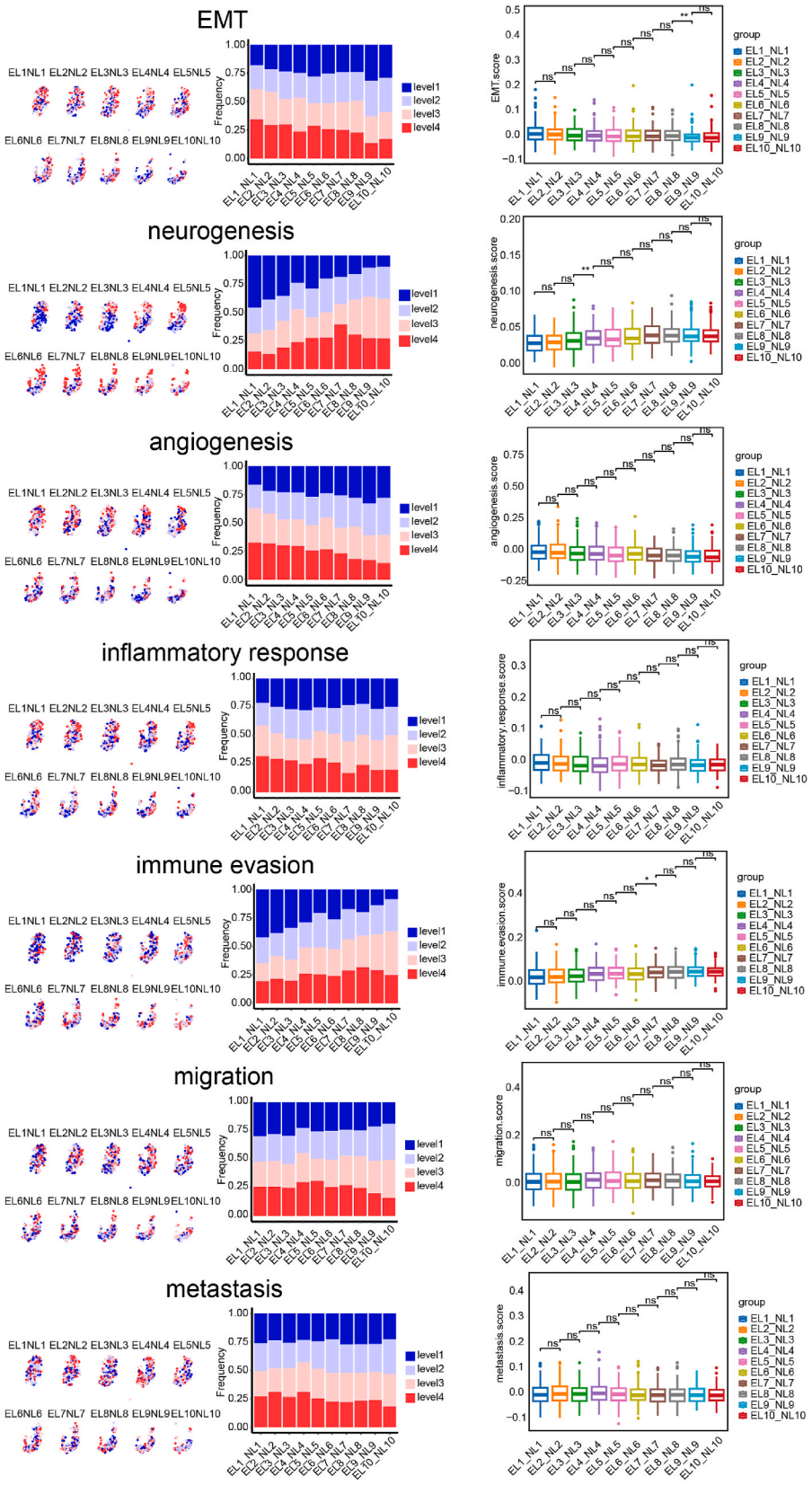
4. Methods

4.1. Data sources

We used hepatocellular cancer [24] and Colorectal cancer [46] single-cell RNA-seq data which we previously published. Specifically, HCC data can be found at the China National GeneBank DataBase under the accession number CNP0004283. CRC data can be accepted in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2021) in National Genomics Data Center (Nucleic Acids Res 2021), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: PRJCA007492) that are publicly accessible at https://ngdc.cncb.ac.cn/gsa-human. For the other tumor single-cell RNA-seq data, we utilized resources from the Gene Expression Omnibus (GEO). The respective accession numbers are as follows: acute lymphoblastic leukemia (ALL) [25] is GSE148218, uveal melanoma (UM) [26] is GSE139829, osteosarcoma (OS) [27] is GSE152048, prostate cancer (PCa) [28] is GSE193337, pancreatic ductal adenocarcinoma (PDAC) [29] is GSE155698, breast cancer (BRCA) [30] is GSE180286, and non-small cell lung cancer (NSCLC) [31] is GSE207422. Additionally, we incorporated known tumor-driven genes data from both humans and mice, including ETV6-RUNX1 in human ALL [25], EGFR mutant in human lung adenocarcinoma [33] (GSE171145), MYCN in mouse neuroblastoma [35] (GSE180101), and c-MYC in mouse prostate cancer [34] (GSE163316). As for the sc-ATAC data, we extracted chronic lymphocytic leukemia [32] data with the accession number GSE159417.

4.2. Single-cell RNA-seq data processing

All downstream analyses were implemented using R version 4.0.2. For the HCC and CRC single-cell RNA-seq data, we deal with these as same as previously described in methods [24,46]. The tumor single-cell data retrieved from GEO primarily comprises standard 10X Genomics format. These data were then read using the read10X function and subsequently processed to create a SeuratObject utilizing the Seurat package v4.4.0. Before analyses, we performed strict quality control (QC) to ensure cell activity, the QC standards are (i) genes expressed by < 3 cells were excluded; (ii) cells that had either fewer than 200 (low-quality cells) or over 8000 expressed genes (possible doublets), or over 20 % of unique molecular identifiers (UMIs) derived from the mitochondrial genome were removed. After QC, the NormalizeData and ScaleData functions from the Seurat package were employed to normalize and scale the gene expression data for all samples. Then 3000 highly variable genes (HVGs) of each cell were generated with the FindVariableFeatures function for Principal Component Analysis (PCA) analysis. After performing PCA, graph-based clustering was conducted using the top 30 principal components (PCs) with a resolution ranging from 0.1 to 0.8 to identify suitable cluster numbers. The clustering results were then visualized using t-distributed Stochastic Neighbor Embedding (tSNE) or uniform manifold approximation and projection (UMAP) for a comprehensive representation. Additionally, to address individual differences and batch effects, we employed the RunHarmony function especially when notable disparities among samples were observed in the tSNE or UMAP visualizations. Then we used the FindAllMarkers function (Wilcoxon rank sum test, logfc.threshold = 0.5) to identify cluster markers and used FindMarkers (Wilcoxon rank sum test, logfc.threshold = 0.5) to find differentially expressed genes between clusters. Moreover, we leveraged the



(caption on next page)

Fig. 6. Cellular reaction programs of a human tumor cell with a given gene expressing pattern. The tSNE projections and stacked bar charts show 4 levels of EMT, neurogenesis, angiogenesis, inflammatory response, immune evasion, migration, and metastasis, while the subsequent box plots illustrate the differences in these characteristics under varying gene expression counts and levels. The tSNE projections were split by 10 categories of overall genes, formed by integrating 10 gene expression levels and 10 number levels, with lower levels consolidated into broader groups. The stacked bar charts show the 4 levels of cellular reaction programs in the P1 sample of CRC, and from blue to red indicating increasing levels. Statistical differences are indicated by asterisks. *p*-values were calculated by Wilcoxon test. ns: not significant. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

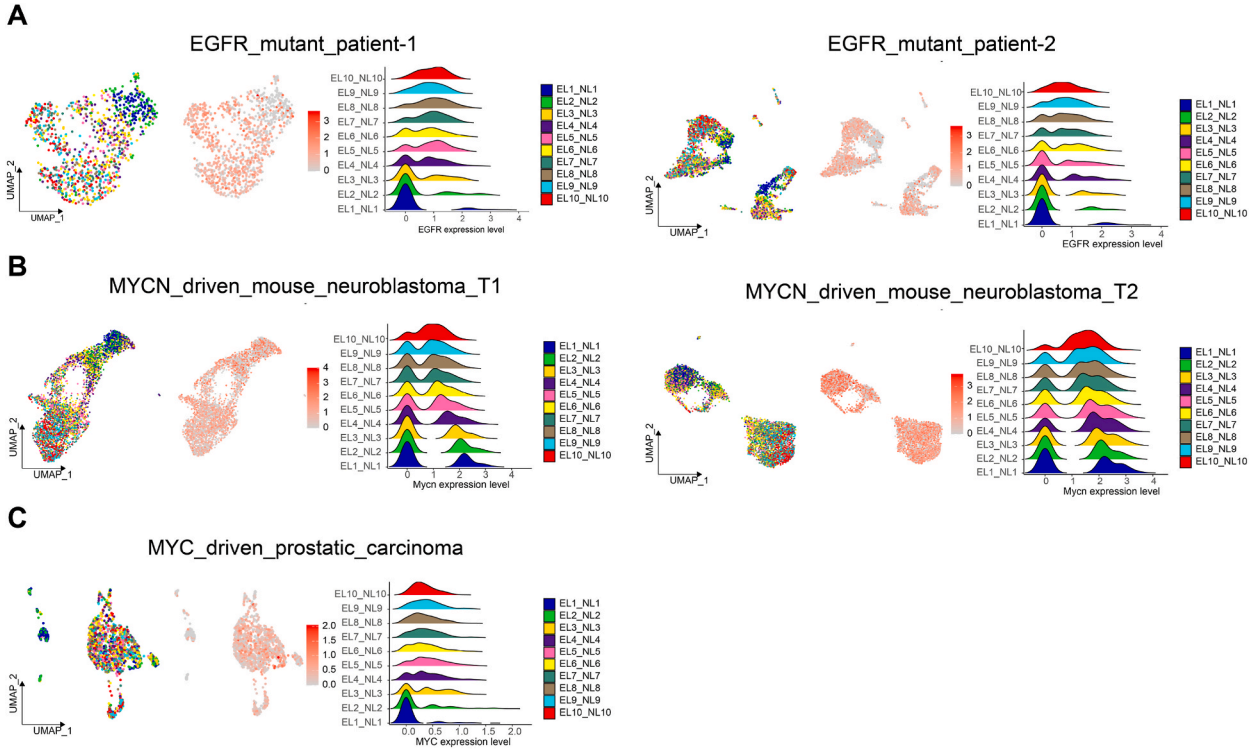


Fig. 7. The expression patterns of designed driven gene in tumor cells. A.B.C. The tSNE projection shows the distribution of 10 categories of overall genes, formed by integrating 10 gene expression levels and 10 number levels, with lower levels consolidated into broader groups. The feature plot shows the expression level of EGFR, and the ridge plot depicts the density of EGFR expression level. A. EGFR in human lung adenocarcinoma with EGFR mutations. B. MYCN in MYCN-driven mouse neuroblastoma. C. MYC in MYC-driven mouse prostate cancer.

SingleR package for automatic annotation and incorporated marker genes from relevant articles to enrich and enhance our annotation results. Then, the classic marker genes were used to annotate the cell types of each cluster.

4.3. Single-cell copy number variation (CNV) analysis

To identify malignant cells, the CNVs for each cell in each sample except ALL were estimated with the inferCNV package v1.2.1 [47]. Leveraging immune cells from the same sample as a reference, we inferred copy number variations in tumor cells based on gene expression profiles. A sliding window including 100 genes was used to average the expression of these genes. By aligning the genes to human chromosomes, we derived the overall inferred copy number variations for each cell. In the heatmap, each horizontal line depicts a single cancer cell. The red areas indicate higher copy numbers relative to the reference, while the blue areas signify lower copy numbers.

4.4. Classification defining of tumor cells based on expressed overall genes

The malignant cell clusters from the scRNA-seq data of every sampls were extracted at first. Based on the single-cell expression matrix, the “nFeature_RNA” (represents the measured total number of genes) and “nCount_RNA” (represents the measured total number of mRNA molecules) information of each cell were calculated after creating SeuratObject. Using the quantile function within each sample, we categorized the tumor cells into 4, 6, 8, and 10 classifications based on “nFeature_RNA”. Consequently, we obtained the number levels (NLs) for each quantile, where each NL represents the corresponding interval range within the respective quantile.

The classification of "nCount_RNA" was labeled using EL, following a similar approach as that for "nFeature_RNA". The final classification IDs of overall genes were obtained by combining the EL and NL labels of each cell, which were then utilized for subsequent analytical procedures.

4.5. Classification defining of tumor cells based on single-cell chromatin accessibility

To further investigate the extent of chromatin accessibility in tumor cells, we analyzed single-cell chromatin accessibility data from one CML sample. Analogous to single-cell RNA sequencing analysis, we employed the open number (ON) and open degree (OD) metrics to classify chromatin accessibility patterns.

4.6. Pearson correlation analysis

The Pearson correlation analysis was conducted to investigate the correlation between "nFeature_RNA", "nCount_RNA", and CNV.score. Initially, the CNV.score was computed based on the outcomes of inferCNV, which generates the file named infercnv.observations.txt. Subsequently, this score was appended to the Seurat object's metadata by employing the AddMetaData function. To visualize the correlation among these variables, the ggscatter and stat_cor functions were utilized.

4.7. Trajectory analysis

To explore the potential lineage differentiation within the tumor, we inferred its trajectory based on pseudotime analysis using Monocle v2.28.0. The newCellDataSet function was used to create a monocle object based on the normalized data and metadata of above Seurat object. To select genes suitable for trajectory inference, the dispersionTable function from Monocle 2 was employed to compute a smooth function capturing expression variance across cells alongside their mean expression levels. Genes with a mean expression of at least 0.1 were exclusively considered for analysis. Subsequently, the variable genes identified by Seurat underwent dimension reduction using the reduceDimension function, specifying reduction_method = "DDRTree" and max_components = 2. To visualize and order the single cells, the plot_cell_trajectory function from Monocle 2 was leveraged, enabling the inference of cell trajectories. By adjusting the color_by parameter, all classification labels were mapped onto the respective trajectories.

4.8. Evaluation of multiple cellular potentialities

To evaluate the potentialities linked to cell states and identities in individual cancer cells, scores of numerous cellular programs and behaviors were computed utilizing the AddModuleScore function within the Seurat R package. We specifically measured the potentialities related to HCC, CRC, and ALL. For HCC, we delved into potentialities encompassing hepatocytic signatures, toxic substances, proliferation, apoptosis, autophagy, stemness, senescence statuses, Epithelial-mesenchymal transition (EMT), neurogenesis, angiogenesis, inflammatory response, immune evasion, migration, metastasis, glycolysis, tricarboxylic acid (TCA) cycle, ATP production, and oxidative phosphorylation. For CRC, we explored similar potentialities to HCC, including proliferation, apoptosis, autophagy, stemness, EMT, neurogenesis, angiogenesis, inflammatory response, immune evasion, migration, metastasis, as well as goblet cells, enterocytes, Paneth cells, and Lgr5+ cell characteristics. In the case of ALL, our focus was on potentialities such as apoptosis, autophagy, proliferation, lymphocyte migration, and immune response (comprising B cell activation and cytokine production), along with $CD38^{-}$leukemia cell, hematopoietic cell, immature B cell, and mature B cell signatures. The corresponding gene sets were downloaded from the Molecular Signatures Database (MSigDB; https://www.gsea-msigdb.org/gsea/msigdb/). Based on the AddModuleScore function algorithm, all analyzed genes were grouped based on their average expression, and control genes were selected randomly from each expression bin, as per established protocols [48]. The score of each cell was determined by calculating the average expression level of all genes within a specific gene set. Subsequently, based on the quartiles of the scores across all samples, four distinct levels were assigned to each cellular program. Cells exhibiting scores below the Q1 (lower quartile value) were categorized as level 1. Cells with scores ranging from Q1 to Q2 (median value) were designated as level 2. Cells scoring between Q2 and Q3 (upper quartile value) were classified as level 3. Finally, cells with scores exceeding Q3 were grouped into level 4.

To visualize the distribution of these levels across samples, we utilized the Dimplot and ggplot functions. Specifically, we focused on the final 10 classifications, which were aggregated based on their lower-level groupings. This visualization approach allowed us to gain a comprehensive understanding of the expression patterns and potentialities associated with each cellular program across different cancer types.

4.9. Expression of driven genes in different types of tumors

To better understand the expression level of different types of tumors, we utilized data from humans and mice. We employed the Dimplot, ggplot, and geom_density_ridges functions to visualize the driving gene expression level of each sample.

4.10. Statistical analysis

All statistical analyses were implemented in R software. For the comparisons between two subpopulations of cancer cells, the ggpubr package v0.4.0 was used (Wilcoxon test). ns: not significant, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$.

Data availability statement

Data associated with this study has been deposited into publicly available repositories and is accessible in the China National GeneBank DataBase (CNGBdb, https://db.cngb.org/), the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2021) in the National Genomics Data Center (Nucleic Acids Res 2021), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: PRJCA007492) [https://ngdc.cncb.ac.cn/gsa-human], and the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). R codes for analyzing the data are available upon reasonable request.

CRediT authorship contribution statement

Yinchun Lv: Methodology, Formal analysis, Data curation. **Yulin Chen:** Data curation, Conceptualization. **Xue Li:** Methodology, Data curation. **Siying Li:** Formal analysis. **Qiaorong Huang:** Data curation. **Ran Lu:** Methodology, Data curation, Conceptualization. **Junman Ye:** Methodology, Data curation. **Wentong Meng:** Project administration. **Xiaolong Chen:** Project administration, Data curation. **Xianming Mo:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The Ethics Committee of West China Hospital, Sichuan University approved this study. We thank the participants and their families for their kind cooperation, generosity, and patience. The authors thank Yifan Mo, Sidney Kimmel medical college, for manuscript editing. This work was supported by the 1.3.5 project for disciplines of excellence of West China Hospital, Sichuan University (ZYJC21003, ZYGD23026).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e35529.

References

[1] O. Farc, V. Cristea, An overview of the tumor microenvironment, from cells to complex networks, Exp. Ther. Med. 21 (2021) 96, https://doi.org/10.3892/etm.2020.9528 (Review).
[2] J. Fang, et al., Exploring the crosstalk between endothelial cells, immune cells, and immune checkpoints in the tumor microenvironment: new insights and therapeutic implications, Cell Death Dis. 14 (2023) 586, https://doi.org/10.1038/s41419-023-06119-x.
[3] D. Roulin, Y. Cerantola, A. Dormond-Meuwly, N. Demartines, O. Dormond, Targeting mTORC2 inhibits colon cancer cell proliferation in vitro and tumor formation in vivo, Mol. Cancer 9 (2010) 57, https://doi.org/10.1186/1476-4598-9-57.
[4] X. Tong, et al., Targeting cell death pathways for cancer therapy: recent developments in necroptosis, pyroptosis, ferroptosis, and cuproptosis research, J. Hematol. Oncol. 15 (2022) 174, https://doi.org/10.1186/s13045-022-01392-3.
[5] W. Gao, X. Wang, Y. Zhou, X. Wang, Y. Yu, Autophagy, ferroptosis, pyroptosis, and necroptosis in tumor immunotherapy, Signal Transduct. Targeted Ther. 7 (2022) 196, https://doi.org/10.1038/s41392-022-01046-3.
[6] C.A. Schmitt, B. Wang, M. Demaria, Senescence and cancer - role and therapeutic opportunities, Nat. Rev. Clin. Oncol. 19 (2022) 619–636, https://doi.org/10.1038/s41571-022-00668-4.
[7] M. Mareel, M.J. Oliveira, I. Madani, Cancer invasion and metastasis: interacting ecosystems, Virchows Arch. 454 (2009) 599–622, https://doi.org/10.1007/s00428-009-0784-0.
[8] A. Dongre, R.A. Weinberg, New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer, Nat. Rev. Mol. Cell Biol. 20 (2019) 69–84, https://doi.org/10.1038/s41580-018-0080-4.
[9] H. Clevers, The cancer stem cell: premises, promises and challenges, Nat. Med. 17 (2011) 313–319, https://doi.org/10.1038/nm.2304.
[10] F.R. Greten, S.I. Grivennikov, Inflammation and cancer: triggers, mechanisms, and consequences, Immunity 51 (2019) 27–41, https://doi.org/10.1016/j.immuni.2019.06.025.
[11] N.N. Pavlova, J. Zhu, C.B. Thompson, The hallmarks of cancer metabolism: still emerging, Cell Metabol. 34 (2022) 355–377, https://doi.org/10.1016/j.cmet.2022.01.007.
[12] P. Carmeliet, R.K. Jain, Molecular mechanisms and clinical applications of angiogenesis, Nature 473 (2011) 298–307, https://doi.org/10.1038/nature10144.
[13] D.A. Silverman, et al., Cancer-associated neurogenesis and nerve-cancer cross-talk, Cancer Res. 81 (2021) 1431–1440, https://doi.org/10.1158/0008-5472.Can-20-2793.
[14] D. Hanahan, Hallmarks of cancer: new dimensions, Cancer Discov. 12 (2022) 31–46, https://doi.org/10.1158/2159-8290.Cd-21-1059.
[15] S. Yuan, R.J. Norgard, B.Z. Stanger, Cellular plasticity in cancer, Cancer Discov. 9 (2019) 837–851, https://doi.org/10.1158/2159-8290.Cd-19-0015.
[16] J. Fan, K. Slowikowski, F. Zhang, Single-cell transcriptomics in cancer: computational challenges and opportunities, Exp. Mol. Med. 52 (2020) 1452–1465, https://doi.org/10.1038/s12276-020-0422-0.
[17] D. Jovic, et al., Single-cell RNA sequencing technologies and applications: a brief overview, Clin. Transl. Med. 12 (2022) e694, https://doi.org/10.1002/ctm2.694.
[18] B. De Craene, G. Berx, Regulatory networks defining EMT during cancer initiation and progression, Nat. Rev. Cancer 13 (2013) 97–110, https://doi.org/10.1038/nrc3447.

[19] J.M. Sperger, et al., Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 13350–13355, https://doi.org/10.1073/pnas.2235735100.
[20] Z. Arshad, J.F. McDonald, Changes in gene-gene interactions associated with cancer onset and progression are largely independent of changes in gene expression, iScience 24 (2021) 103522, https://doi.org/10.1016/j.isci.2021.103522.
[21] Y. Wang, et al., MAPK1 promotes the metastasis and invasion of gastric cancer as a bidirectional transcription factor, BMC Cancer 23 (2023) 959, https://doi.org/10.1186/s12885-023-11480-3.
[22] M. Billmann, V. Chaudhary, M.F. ElMaghraby, B. Fischer, M. Boutros, Widespread rewiring of genetic networks upon cancer signaling pathway activation, Cell Syst 6 (2018) 52–64.e54, https://doi.org/10.1016/j.cels.2017.10.015.
[23] R.Q. Figueiredo, et al., Elucidating gene expression patterns across multiple biological contexts through a large-scale investigation of transcriptomic datasets, BMC Bioinf. 23 (2022) 231, https://doi.org/10.1186/s12859-022-04765-0.
[24] J. Zhao, et al., Gene expression networks involved in multiple cellular programs coexist in individual hepatocellular cancer cells, Heliyon 9 (2023) e18305, https://doi.org/10.1016/j.heliyon.2023.e18305.
[25] J. Mehtonen, et al., Single cell characterization of B-lymphoid differentiation and leukemic cell states during chemotherapy in ETV6-RUNX1-positive pediatric leukemia identifies drug-targetable transcription factor activities, Genome Med. 12 (2020) 99, https://doi.org/10.1186/s13073-020-00799-2.
[26] M.A. Durante, et al., Single-cell analysis reveals new evolutionary complexity in uveal melanoma, Nat. Commun. 11 (2020) 496, https://doi.org/10.1038/s41467-019-14256-1.
[27] Y. Zhou, et al., Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma, Nat. Commun. 11 (2020) 6322, https://doi.org/10.1038/s41467-020-20059-6.
[28] I. Heidegger, et al., Comprehensive characterization of the prostate tumor microenvironment identifies CXCR4/CXCL12 crosstalk as a novel antiangiogenic therapeutic target in prostate cancer, Mol. Cancer 21 (2022) 132, https://doi.org/10.1186/s12943-022-01597-7.
[29] C.J. Halbrook, et al., Differential integrated stress response and asparagine production drive symbiosis and therapy resistance of pancreatic adenocarcinoma cells, Nat. Can. (Ott.) 3 (2022) 1386–1403, https://doi.org/10.1038/s43018-022-00463-1.
[30] K. Xu, et al., Single-cell RNA sequencing reveals cell heterogeneity and transcriptome profile of breast cancer lymph node metastasis, Oncogenesis 10 (2021) 66, https://doi.org/10.1038/s41389-021-00355-6.
[31] J. Hu, et al., Tumor microenvironment remodeling after neoadjuvant immunotherapy in non-small cell lung cancer revealed by single-cell RNA sequencing, Genome Med. 15 (2023) 14, https://doi.org/10.1186/s13073-023-01164-9.
[32] Z. Duren, et al., Sc-compReg enables the comparison of gene regulatory networks between conditions using single-cell data, Nat. Commun. 12 (2021) 4763, https://doi.org/10.1038/s41467-021-25089-2.
[33] L. Yang, et al., Single-cell transcriptome analysis revealed a suppressive tumor immune microenvironment in EGFR mutant lung adenocarcinoma, J Immunother Cancer 10 (2022), https://doi.org/10.1136/jitc-2021-003534.
[34] X. Qiu, et al., MYC drives aggressive prostate cancer by disrupting transcriptional pause release at androgen receptor targets, Nat. Commun. 13 (2022) 2559, https://doi.org/10.1038/s41467-022-30257-z.
[35] A. Costa, et al., Single-cell transcriptomics reveals shared immunosuppressive landscapes of mouse and human neuroblastoma, J Immunother Cancer 10 (2022), https://doi.org/10.1136/jitc-2022-004807.
[36] K.J. Kobayashi-Kirschvink, et al., Prediction of single-cell RNA expression profiles in live cells by Raman microscopy with Raman2RNA, Nat. Biotechnol. (2024), https://doi.org/10.1038/s41587-023-02082-2.
[37] P.V. Kharchenko, The triumphs and limitations of computational methods for scRNA-seq, Nat. Methods 18 (2021) 723–732, https://doi.org/10.1038/s41592-021-01171-x.
[38] D. Friedmann-Morvinski, I.M. Verma, Dedifferentiation and reprogramming: origins of cancer stem cells, EMBO Rep. 15 (2014) 244–253, https://doi.org/10.1002/embr.201338254.
[39] A. Jögi, M. Vaapil, M. Johansson, S. Påhlman, Cancer cell differentiation heterogeneity and aggressive behavior in solid tumors, Ups. J. Med. Sci. 117 (2012) 217–224, https://doi.org/10.3109/03009734.2012.659294.
[40] X. Yang, et al., Suppression of cell tumorigenicity by non-neural pro-differentiation factors via inhibition of neural property in tumorigenic cells, Front. Cell Dev. Biol. 9 (2021) 714383, https://doi.org/10.3389/fcell.2021.714383.
[41] G. Peng, P.P.L. Tam, N. Jing, Lineage specification of early embryos and embryonic stem cells at the dawn of enabling technologies, Natl. Sci. Rev. 4 (2017) 533–542, https://doi.org/10.1093/nsr/nwx093.
[42] M.A. Dawson, T. Kouzarides, Cancer epigenetics: from mechanism to therapy, Cell 150 (2012) 12–27, https://doi.org/10.1016/j.cell.2012.06.013.
[43] M. Greaves, C.C. Maley, Clonal evolution in cancer, Nature 481 (2012) 306–313, https://doi.org/10.1038/nature10762.
[44] S. Deng, Y. Feng, S. Pauklin, 3D chromatin architecture and transcription regulation in cancer, J. Hematol. Oncol. 15 (2022) 49, https://doi.org/10.1186/s13045-022-01271-x.
[45] B.V. Chakravarthi, S. Nepal, S. Varambally, Genomic and epigenomic alterations in cancer, Am. J. Pathol. 186 (2016) 1724–1735, https://doi.org/10.1016/j.ajpath.2016.02.023.
[46] S. Li, et al., An integrated map of fibroblastic populations in human colon mucosa and cancer tissues, Commun. Biol. 5 (2022) 1326, https://doi.org/10.1038/s42003-022-04298-5.
[47] A.P. Patel, et al., Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma, Science 344 (2014) 1396–1401, https://doi.org/10.1126/science.1254257.
[48] I. Tirosh, et al., Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq, Science 352 (2016) 189–196, https://doi.org/10.1126/science.aad0501.