ARTICLE

B1 oligomerization regulates PML nuclear body biogenesis and leukemogenesis

Yuwen Li¹,², Xiaodan Ma¹,², Zhiming Chen¹,², Haiyan Wu¹,², Pengran Wang¹,², Wenyu Wu¹, Nuo Cheng¹, Longhui Zeng¹, Hao Zhang¹, Xun Cai¹, Sai-Juan Chen¹, Zhu Chen¹ & Guoyu Meng¹

ProMyelocyticLeukemia (PML) protein can polymerize into a mega-Dalton nuclear assembly of 0.1–2 μm in diameter. The mechanism of PML nuclear body biogenesis remains elusive. Here, PMLRBCC is successfully purified. The gel filtration and ultracentrifugation analysis suggest a previously unrecognized sequential oligomerization mechanism via PML monomer, dimer, tetramer and N-mer. Consistently, PML B1-box structure (2.0 Å) and SAXS characterization reveal an unexpected networking by W157-, F158- and SD1-interfaces. Structure-based perturbations in these B1 interfaces not only impair oligomerization in vitro but also abolish PML sumoylation and nuclear body biogenesis in HeLaPml⁻/⁻ cell. More importantly, as demonstrated by in vivo study using transgenic mice, PML-RARα (PR) F158E precludes leukemogenesis. In addition, single cell RNA sequencing analysis shows that B1 oligomerization is an important regulator in PML-RARα-driven transactivation. Altogether, these results not only define a previously unrecognized B1-box oligomerization in PML, but also highlight oligomerization as an important factor in carcinogenesis.

¹State Key Laboratory of Medical Genomics, Shanghai Institute of Hematology, Rui Jin Hospital affiliated to Shanghai Jiao Tong University School of Medicine, 197 Ruijin Er Road, Shanghai 200025, China. ²These authors contributed equally: Yuwen Li, Xiaodan Ma, Zhiming Chen, Haiyan Wu, Pengran Wang. Correspondence and requests for materials should be addressed to G.M. (email: guoyumeng@shsmu.edu.cn)
he ProMyelocyticLeukemia (PML) protein, also known as TRIM19, is the core component of PML nuclear bodies (NBs) that are tightly associated with nuclear matrix. It has been reported that PML NB can play many instrumental roles in DNA-damage responses, apoptosis, cellular senescence, angiogenesis, etc. It has been reported that PML NB can make direct handshakes with >120 different proteins, hence is recently recognized as PML interactosome. Furthermore, via post-translational modifications like sumoylation, phosphorylation, acetylation, etc., PML can significantly influence its partners’ fate and function in cellular signaling pathways.

Unexpectedly, PML1 has been reported to be the central hubs of cellular signaling pathways. Although like most of other TRIM proteins, all of these are considered important for NB morphogenesis but not essential. Indeed, we have recently demonstrated that the PML oligomerization might occur prior to sumoylation. However, due to the lack of success in obtaining purified PML or its RBCC fragment, our knowledge of PML NB biogenesis is extremely sparse. In this study, we report the largest RBCC fragment, which is in good agreement with the crystallographic characterization of PML or its RBCC domain.

Results

RBCC oligomerization. Although early observation of PML NBs can date back to 1960s, the purification of PML protein remains a formidable challenge. This has significantly restricted the investigation and understanding of PML-driven polymerization and nuclear body biogenesis (middle and right panels, Fig. 1a). In this study, a modified pET vector containing an N-terminal human SUMO tag was engineered. The presence of SUMO moiety had significantly enhanced the expression and folding of PML1-256 that contained the RING, a B1-box (B1), a B2-box (B2), and a C-terminal coiled-coil (CC) (Fig. 1a). Previous studies have shown that the integrity of RBCC domain is essential for PML NB formation. Furthermore, the PML sumoylation and SUMO-interacting-motif are also considered important for NB morphogenesis but not essential. Indeed, we have recently demonstrated that the PML oligomerization might occur prior to sumoylation. However, due to the lack of success in obtaining purified PML or its RBCC fragment, our knowledge of PML NB biogenesis is extremely sparse. In this study, we report the largest RBCC fragment ever purified and the 2.0 Å crystal structure of B1 multimers. These results have prompted the in vitro and in vivo characterizations of PML and its oncogenic fusion PML-RARα (PR) in atomic detail.

B1-box multimers. In line with the GF/UC analysis above, the structure of B1-box was determined to 2.0 Å resolution, unveiling an unexpected oligomerization in PML (Figs. 2, 3, and Supplementary Figs. 1 and 2). In the B1-box monomer, the residues 120–127 fold into a loop-like subdomain 1 (SD1), which is ~12 Å away from the B1 core. In comparison, the residues 128–167, coordinated by two Zn ions, fold into a compact global domain, termed subdomain 2 (SD2). The relative orientation between SD1 and SD2 gives rise to a tea-cup-shape architecture, which is also observed in the NMR B1 monomer (Fig. 2a and Supplementary Fig. 2a). The subdomain 2 is a typical β-box-type Zn finger fold with a conserved β1–β2–α1–β3 arrangement and Zn-binding residues (Fig. 2b). The three strands, which are located in between the α1 helix and SD1, form an anti-parallel β-sheet. The α1 helix and β-sheet are coordinated with two Zn ions, residues C129/C132/C148/C151 (Zn1) and C140/C143/H155/H161 (Zn2). Above the Zn2 site is K160 (Fig. 2a), a critical sumoylation site for leukemia development. The long lysyl side-chain and the amine head are fully exposed to solvent. This spatial location is consistent with its role as a SUMO acceptor. Next to K160 are two bulky residues, W157 and F158. The side chains of W157 and F158 are forming a 90° angle to each other, which is important for B1-oligomerization (Fig. 2c, d).

It has been reported that TRIM B1/2-box could mediate multimerization (Fig. 2b). This is also the case for PML B1-box. The W157, F158, and SD1 could facilitate the formation of PML B1 dimers (Fig. 2c–e). In W157-dimer, the inter-molecular engagement is mainly mediated by two sets of hydrophobic interactions between W157 and F152/F158 (Fig. 2c). In F158-dimer, the bulky hydrophobic side-chains are located in the heart of the dimeric interface (Fig. 2d), reminiscent of L73–L73 handshake in PML RING dimerization. Interestingly, the Trp-Trp, Phe-Phe, and Leu-Leu interactions are often observed in TRIM multimerization (Supplementary Fig. 2b), highlighting the conserved α1 helices in B-boxes as an oligomeric hot spot. In addition to the F158–F158 interaction, the hydrogen bonds between W157–Q145 (2.7 Å) and R131–E153 (3.2 Å) can facilitate the B1–B1 interaction (Fig. 2d). In SD1-dimer, the N-terminal N122 and V123 residues could undergo a loop-to-strand transition to seal two B1s together via β-strand augmentation (Fig. 2e).

As revealed by crystal packing, these dimeric interfaces are the foundation of B1-driven higher order assembly (Fig. 3). The intermittent engagement between W157- and F158-interfaces enables the formation of a “brick”-like tetramer (Fig. 3a and Supplementary Fig. 2c). This, in turn, gives rise to two interesting positions of K160 and SD1s. In the α1-mediated W-shape coiled-coil, the K160 residues are consistently placed at the bottom of the B1 tetramer, giving rise to a highly concentrated Lys patch. Supportively, only this Lys position, but not K65 and K490, allows poly-SUMO reaction, consistent with the K160 linearizing in W157/F158-tetramer (Supplementary Fig. 2c). In addition, this tetramer is flanked with four SD1 subdomains that are ready for further β-augmentation (Fig. 3b). As a result, remarkable networking mediated by B1-box could be envisaged (Fig. 3c).

Functional characterizations of B1-driven oligomerization. In order to check PML B1-box multimerization in solution, small-angle X-ray scattering (SAXS) was used. In this experiment, the atomic structures of the PML B1-box monomer, dimers, tetramer, and N-mer reported here were subjected to crystal fitting using the OLIGOMER algorithm implemented in CRYOSOL. The SAXS characterization is in good agreement with the crystallographic observation (χ² = 1.21), suggestive of the B1 oligomerization in solution (Fig. 4a). Furthermore, when the W157- and F158-interfaces were perturbed, the B1 multimerization (but not the overall fold, Supplementary Fig. 3a, b) was severely impaired, precluding B1 tetramer and higher-order assemblies (Fig. 4b–d). For further cross-validation, an oligomerization-dependent crosslinking assay was performed (Fig. 4e, f). When an increased concentration of cross-linker GA were incubated with...
analysis of PML1 mutants displayed marked resistance to the GA cross-linking and western blots. In contrast, W157E, F158E, and I122P/V123P growing accumulation of B1 and RBCC polymers in sodium chloride elution, as monitored by SDS-PAGE. The theoretical molecular weight (MW) of PML1–256 is 28 kDa. c Gel filtration analysis of PML1–256. The elution peaks are designated as P1–P3, respectively. Source data are provided as a Source Data file. The Superdex 200 column used in this study was calibrated using the standard molecular marker kit (GE Healthcare). The reference molecular weights are indicated. d Analytical ultracentrifugation analysis of PML1–256, gel filtration eluants/peaks. The sedimentation coefficients of Peaks 1/2/3 are colored in black, magenta and blue, respectively. From left to right, the molecular masses of the UC peaks are 34, 54, 99, and 162 kDa, suggesting the assemblies of PML1–256 monomer, dimer, tetramer and much higher order polymers (termed N-mer in this report). Source data are provided as a Source Data file.

Besides, the B1 oligomerization was also validated by GF and analytical UC (Fig. 5a, b). As mentioned above, WT PML1–256 was eluted with three peaks containing 1/2/4/N-mers. In comparison, W157E and F158E significantly altered the PML1–256 elution profile (Fig. 5a). The monomeric peak (P4), which was not observed in WT, was now visible in mutants. The same results were obtained by analytical UC analysis (Fig. 5b). Similar disruptive impact, but at the later stage of PML oligomerization, could be observed for I122P/V123P (Fig. 5b). This was further supported by the mammalian two-hybrid assay using the full-length PML isoform IV protein (Fig. 5c). In this assay, we observed similar disruptive impacts by W157E, F158E, and I122P/V123P. The mutations all consistently impaired the PML self-interaction (Fig. 5c). More importantly, the PML isoform IV mutants W157E, F158E, and I122P/V123P, although expressed normally in HeLaPml+/− cells, precluded NB formation and PML sumoylation at basal level (Fig. 5d–f and Supplementary Fig. 4a). Consistently, similar results could be obtained in PML isoform I (Supplementary Fig. 5a, b). Complementary to these results, we also observed the same results in W157K, W157A, F158K, F158A, I122P, V123P, and Δ120–124 (Supplementary Fig. 4b, c). In marked contrast, the mutations of D137A, S135A/D137A, and A149E that targeted the NMR B1 dimeric interface had little impact on NB formation (Supplementary Fig. 4d, e), urging caution in the interpretation of NMR B1 dimer26.

It has been reported that partners recruitment is also important for NB formation and function16,19. We next want to know how the loss of RBCC oligomerization might affect the PML recruitment activity. Using immuno co-localization and mammalian two-hybrid assays, we examined the interactions between PML and DAXX, SUMO2/3 in HeLaPml+/− cells (Fig. 6a, b and Supplementary Fig. 6a, b). In the co-localization assay, the sumoylation mutation K160R was used as control. As expected, DAXX and SUMOs could colocalize with WT NBs, but not with the K160R mutant, reiterating the importance of sumoylation in PML partners recruitment19. Consistently, the B1 mutants, which also abolished PML auto-sumoylation at basal level (Fig. 5e), displayed reduced DAXX/SUMOs recruitment activity (Fig. 6a, b and Supplementary Fig. 6a, b). Of note, I122P/V123P, which targeted the possible later step in B1 polymerization, could still form NBs with the abnormal bigger size, and hence displayed NB-enhanced partners recruitment in the co-localization assay. In addition, when the B1 mutants were subjected to arsenic trioxide...
(ATO) treatment, we could observe NBs re-appearance, accompanied with NB-enhanced partners recruitment (Fig. 6c and Supplementary Fig. 6c). Similar partners recruitment and ATO rescue results could also be repeated in PML isoform I (Supplementary Fig. 5c, d). Altogether, these results highlight the idea that the RBCC oligomerization is likely to be the first step in PML NB biogenesis, preceding PML sumoylation and partners recruitment.

Interestingly, when various PML mutants were subjected to ATO treatment, we observed a striking rescue difference between B1 and RING oligomerizations. Unlike L73E in RING21, the damage caused by F158E or W157E/F158E in B1-box could be rescued by arsenic treatment (Fig. 7). After short exposure of As2O3, F158E and W157E/F158E could reform NBs (Fig. 7a, b). As a result, the B1 mutants recovered some levels of auto-sumoylation (Fig. 7c). In marked contrast, no rescue effect could be observed in RING mutant L73E (Fig. 7). The results not only supported our previous claim that PML–UBC9 interaction might occur in L73–RING dimer21, but also suggested a cooperative mechanism among RBCC, in which the RING tetramerization step might precede B1 polymerization in PML nuclear body biogenesis.
B1-box multimerization in PR. Concerning PR, we wanted to know whether the B1-box oligomerization is an important regulator in PR-driven transactivation and leukemogenesis. Firstly, the presence of B1 oligomerization in PR was confirmed by speckle formation assay in which W157E, F158E, and I122P/V123P consistently disrupted the PR nuclear dots (Supplementary Fig. 7a). Secondly, in vitro luciferase assay was used to monitor the impact of the multimerization upon PR transcription activity (Supplementary Fig. 7b). The results were also coherent, supportive of B1 multimerization in oncogenic fusion. Thirdly, using PR or PR F158E transgenic mice, we also observed a strict correlation between B1-oligomerization and leukemogenesis (Fig. 8a and Supplementary Fig. 7c–e). Over one and half year, most PR mice (n = 8) died from APL. In marked contrast, none of PR F158E mice developed leukemia (Fig. 8a). In order to get insight into the leukemogenic transactivation program triggered by oligomerization, the PR, PR F158E transgenic mice, and the normal FVB/N mice (termed WT) were subjected to single cell RNA sequencing analysis (Fig. 8b, c and Supplementary Table 1). According to the gene-expression profiling, the bone marrow cells from WT (3816), PR (4806), and PR F158E (3874) were classified into 15 different clusters that could be recognized as granulocyte, erythrocyte, monocyte, stem cell, B cell, and dendritic cell (Fig. 8b). The t-distributed stochastic neighbor embedding plots of the key markers used in the classification of different cell types are shown in the Supplementary Fig. 8a. The cell type distributions for PR were 50.4%, 39%, 7%, 1.9%, 1.2%, and 0.5% for granulocyte, erythrocyte, monocyte, stem cell, B cell, and dendritic cell, respectively. In comparison, PR F158E displayed a marked difference in cell differentiation (Fig. 8c). The cell type distributions for PR F158E were 26%, 63%, 3.9%, 4.5%, 0.9%, and 1.7% for granulocyte, erythrocyte, monocyte, stem cell, B cell, and dendritic cell, respectively (Supplementary Fig. 8b, c). In addition, the differential expression analysis of PR and PR F158E datasets with stringent threshold values of log2[fold change] > 1.5 and the adjusted p value/FDR < 0.5 helped to uncover a new set of PR-oligomerization target genes (Fig. 9a, b and Supplementary Table 2). Consistently, Trib3 and Atf5, which were upregulated by PR and recently implicated in APL, were clearly downregulated by PR F158E (Fig. 9a, b). Altogether, these results prompt the new thinking of PR, in which oligomerization might be a crucial regulator for leukemia development (Fig. 9c).

Discussion
The B-box domain has been widely observed in nature. Using the Dali server, an invariant PML B1-box fold could be observed, albeit with low sequence identity. They are ADP-ribosylation factor 1 (RMSD 1.5 Å and sequence identity 18%), ubiquitin carboxyl-terminal hydrolase 8 (RMSD 1.8 Å and sequence identity 26%), ubiquitin carboxyl-terminal hydrolase 5 (RMSD 1.8 Å and sequence identity 13%), ubiquitin E3 ligase Bre1a (RMSD 2.3 Å and sequence identity 18%), ubiquitin E3 ligase RNF123 (RMSD, 2.2 Å and sequence identity, 16%), DNA gyrase subunit A (RMSD 1.8 Å and sequence identity, 9%). This implies that...
B-box might serve as a power scaffold during protein evolution. The variations stemming from a conserved B-box fold might be important to protein diversities, including its enzymatic activities and assembly/oligomerization. Indeed, the B-box, which is often associated with RING and coiled-coil (CC) in TRIMs, has been reported to facilitate the assemblies of different oligomerization25–28. Different hydrophobic decorations in the conserved α1 helices lead to various oligomerizations (Supplementary Fig. 2b). Notably, the functional dimer-to-trimer switch regulated by Trp–Trp interaction (Supplementary Fig. 2b) is the key structural determinant for TRIM5α to polymerize into a mega-Dalton hexagonal network that is required for HIV-1 virus recognition28.

In this study, we also report a previously unrecognized B1-box polymerization (n > 4) that is essential for PML/TRIM19 nuclear body assembly. The oligomeric determinants such as W157, F158, and SD1-loop are uniquely observed in PMLs, but not in other TRIMs (Supplementary Figs. 1 and 2). Through W157E, F158E, and SD1-interfaces, PML B1-box can form homo-dimers, tetramer and much higher order polymer (Figs. 2 and 3). The crystallographic observation is further supported by in vitro characterizations (Figs. 4 and 5 and Supplementary Fig. 3). More importantly, in the context of RBCC and full-length PML, perturbations in the oligomeric interfaces significantly impaired PML self-interaction and precluded NB biogenesis, auto-sumoylation and partners recruitment (Figs. 5–7 and Supplementary Figs. 1–4).

**Fig. 4** SAXS and cross-linking assays of PML B1-box oligomerization. a–c SAXS analysis of WT B1 (a), B1 W157E (b), and B1 F158E (c). The experimental data are shown in black dots. The theoretical scattering pattern derived from PML B1 multimers is shown in red line. Bottom-left panels of each figure, the oligomeric distribution/dynamic in solution. Source data are provided as a Source Data file. d Statistical summary of SAXS experiments. e, f Oligomerization-dependent crosslinking assays. In e, PML B1 and mutants were cross-linked by an increased concentration of glutaraldehyde (GA). In f, PML1–256 and mutants were cross-linked by an increased concentration of glutaraldehyde (GA). The poly-B1 and poly-PML1–256 are highlighted with brackets and red boxes. Source data are provided as a Source Data file.
Supplementary Figs. 4–7). However, it is worthy to point out that the B1-mutations-alone cannot completely stop RBCC oligomerization (Fig. 5a, b), reiterating the importance/contributions of RING and CC oligomerization. In the previous studies, we know ATO can facilitate PML NB formation35. Here, as shown in Fig. 7, we observed different ATO responses in RING and B1 mutants. The lack of ATO rescue effect in L73E mutant supports the known ATO can facilitate PML NB formation35. Here, as shown in Fig. 7, we observed different ATO responses in RING and B1 mutants. The lack of ATO rescue effect in L73E mutant supports the knowledge that ATO can facilitate PML NB formation35. In the previous studies, we have shown that the L73E mutation, which inhibits PML RING tetramerization, is essential for leukemogenesis21. Using single-cell RNA sequencing technology, we can detect significantly altered PR-driven transactivation and differentiation in the PR F158E mouse. More consistently, the B1 single mutation that targets the PR oligomerization again precludes APL development in vivo (Figs. 8 and 9). This is further supported by the observation that PML coiled-coil interaction is also important for differentiation arrest and transformation in vivo36. All these results are coherent, highlighting RBCC oligomerization as an important regulator in leukemogenesis. Actually, the oligomerization theme might not be restricted to PR and APL. In oncogenic fusions NPM-RARα and Stat5b-RARα, oligomerization is thought to be instrumental to cell immortalization and diseases development37,38. In proto-oncogenic proteins erbB-2, MET and FGFR, the enhanced and gain-of-carcinogenic effects by multimerization are also observed in tumor growth and cancer progression39–42. Like the W157- and F158-interfaces observed in PML B1 oligomerization, the cancer-related multimerizations are often mediated/stabilized by the bulky hydrophobic residues such as Trp, Phe, Leu, Ile, and Val43–47. Indeed, leukemogenic oligomerizations are widely observed in BCR/ABL, AMLI/ETO, AMLI/MTG16, TEL/AMLI, CBFB/MMHC, MLL/ GAS7, MLL/AF1p, MLL/GEPRIN, and PAX5/PML, etc.48–52. In light of the single cell RNA sequencing results presented here, it might be reasonable to consider oligomerization as an important leukemogenic regulator in these drivers. One of the important research avenues in the future is to investigate whether/how the oncogenic oligomerization might influence the maintenance of the
leukemia-initiating cancer cells. If this is the case, the oncogenic oligomerization might represent a valuable target for future therapeutic exploration beyond APL.

Methods

DNA and protein preparation. PML B1-box consisting of residues 120–168 was constructed into a pET 32a vector with a cleavable His tag on the N-terminus. For PML1–256, a modified pET 15b was used. The cDNA encoding human SUMO moiety was engineered into pET 15b using restriction Ncol site via homologous recombination. The RBCC fragment encoding residues 1–256 was then ligated into the modified pET 15b vector using NdeI and Xhol sites. The resulting plasmid encoded a cleavable His tag in between SUMO and PML1–256. All the PML and PR mutants reported here were engineered using the Quick Change site-directed mutagenesis kit (QIAGEN). The primers used in this study were reported in Supplementary Tables 3 and 4.

The PML B1-box or PML1–256 was transformed into Escherichia coli BL21 (DE3) (Sangon) for protein expression. The cells were grown in LB Broth supplemented with 100 μg/ml ampicillin at 37 °C. When OD600 reached ~0.6, the cells were induced with 300 μM IPTG (Sangon) and 300 μM ZnCl2 (Sangon) for >12 h. The cells were harvested by centrifugation (4700 g, 1 h). The target proteins was pooled and further purified by the hydrophobic chromatography (Phenyl, GE Healthcare), in which the buffer A and B were 20 mM Tris, pH 8.0, 100 mM NaCl, and 20 mM Tris, pH 8.0, 1 M NaCl, respectively. The eluant containing the target proteins was pooled and further purified by the hydrophobic chromatography (Phenyl, GE Healthcare), in which the buffer A and B were 20 mM Tris, pH 8.0, 20 mM NaCl, 1 M (NH4)2SO4 and 20 mM Tris, pH 8.0, 20 mM NaCl, respectively. In the final GF step, the columns S100 and Superdex 200 (GE Healthcare) were selected for PML B1-box and PML1–256, respectively. The GF buffer was 20 mM Tris, pH 8.0, 100 mM NaCl.

Crytalization and structural determination. PML B1-box crystals were obtained by the vapor diffusion technique. The purified protein (38 mg/ml) was mixed at a 1:1 (v/v) ratio with the buffer containing 30% polyethylene glycol 6000, 1 M lithium chloride and 100 mM sodium acetate. The crystals were stabilized in a mixture of parafﬁn and paratone-N (Hampton Research), followed by X-ray examination at 100 K. Diﬀraction data were collected in Beamline station BL17U at Shanghai Synchrotron Radiation Facility (SSRF, Shanghai, China). Zn anomalous dispersion at 1.2824 Å was used for data collection and subsequent phasing. The diﬀraction data were processed, integrated and scaled using MOSFLM/SCALA. The statistics of the data collection are shown in Table 1.

Single wavelength dispersion algorithm in CRANK2, together with data in 20 and 2.0 Å, were used to obtain the first set of phases for PML B1-box. Eight Zn positions were successfully determined and the resulting electron density map was improved by solvent flattening program SOLOMON and auto-tracing program ARP/warp. This procured an excellent α-weighted 2Fo–Fc map for manual model building using COOT and structural refinement using REFMAC5 and Phenix.refine. The final model in ASU contains 196 residues, 8 Zn ions, and 82 water molecules. Ramachandran statistics by PROCHECK showed that 97.4% of the atoms are in the most favored region, and 2.6% are in the allowed regions. The detailed structure reﬁnement statistics are reported in Table 1. The PML B1-box structure has been deposited into the Protein Database Bank with the entry code of 6IMQ.
**Fig. 7** B1-oligomerization might occur after RING tetramerization. **a** Auto-sumoylation of PML RING and B1 mutants at basal level. The auto-sumoylation of HeLa Pml−/− expressing B1 and RING mutants was monitored by western blot using antibody against HA tag. Source data are provided as a Source Data file. **b** NB formation rescued by ATO. Left columns: HeLa Pml−/− cells without arsenic treatment. Right columns: One-hour exposure of 2 μM As2O3 before visualization. **c** PML sumoylation rescued by ATO. The WT and mutant cells, which had been treated with 2 μM As2O3, were disrupted by RIPA buffer. The PML sumoylation that precipitated the soluble PML into insoluble matrix was analyzed by western blot using antibody against HA tag. Source data are provided as a Source Data file.

**Fig. 8** B1-box oligomerization is essential for PML-RARα and APL. **a** Survival data of PML-RARα (PR) or PML-RARαF158E (PR F158E) transgenic mice. The APL phenotypes of these transgenic mice are shown in Supplementary Fig. 8c–e. **b, c** Single-cell RNA sequencing analysis of B1-oligomerization in PML-RARα. The tSNE projections were calculated based on the genetic profiling of ~12,500 bone marrow cells from WT, PR, and PR F158E mice. Information about the sequencing runs, represented as raw bases and average reads/cell per sample, is included in the Supplementary Table 1. To avoid batch differences, the Seurat alignment method canonical correlation analysis (CCA)62 was used for an integrated analysis of three datasets. For cell clustering, the graph-based method implemented in Seurat was used. In **b**, the cells/clusters were grouped into five different cell types, including granulocyte (clusters 0, 4, 5, 9, 12), erythrocyte (clusters 1, 2, 3, 6, 7,10), monocyte (cluster 8), stem cell (cluster 11), B cell (clusters13, 14), and dendritic cell (cluster 15). The tSNE plots of the key markers per cell type are shown in Supplementary Fig. 8a. In **c**, the WT, PR, and PR F158E cells are colored differently to highlight the altered PML-RARα-driven differentiation by B1 mutant. More information concerning the cell type and distribution percentages is shown in Supplementary Fig. 8bc.
Small angle X-ray scattering. Different concentrations (i.e., 2, 4, and 8 mg/ml) of WT PML B1 and mutants were used for SAXS characterization. The X-ray scattering was recorded for a total q range from 0.01 to 0.30 Å⁻¹ in beamline station BL19U2 (SSRF, China). The measurements were carried out with 1 s exposure time and repeated for 20 times to avoid radiation damage. Data subtraction and processing were performed using PRIMUS²⁹. Crystal data fitting was done using the OLIGOMER algorithm implemented in CRYSOL²⁹. In the analysis of crystal fitting, the crystallographic B1 structures including monomer, dimers, tetramer, and N-mer were used to estimate the oligomeric distribution in solutions. For simplicity, the WT and mutant B1 oligomerizations at 2 mg/ml are shown in Fig. 4a-d. The statistics of the data collection and scattering-derived parameters are shown in Supplementary Table 5.

Oligomerization-dependent crosslinking. For PML B1, the purified WT or mutant proteins (5 μg) were incubated with 0.05% (w/v), 0.1% (w/v), and 0.2% (w/v) glutaraldehyde (GA) on ice for 30 min, respectively. For HA-PML1–256, similar cross-linking experiments were repeated using 0.3 μg purified WT or mutant proteins (0.3 μg). The total reaction volume was 20 μl. The buffer used in the cross-linking assay was 20 mM Hepes, pH 7.4. The cross-linking reactions were stopped by adding...
#### Table 1 Data collection and structure refinement statistics of PML B1-box

| Protein | PML B1-box |
|---------|------------|
| **Data collection** | |
| Space group | P2₁ |
| Unit cell (Å) | a, b, c |
| | 36.7, 50.5, 51.4 |
| β | 101.4 |
| Molecule per ASU | 4 |
| Source/station | BL17U |
| Wavelength (Å) | 1.282 |
| Resolution range (Å) | 50.4-2.06 |
| **Observations** (<I/e(I)> > 0) | 45412 |
| α | 11287 (1650) |
| **High resolution shell (Å)** | |
| R<sub>sym</sub> (%) | 17.7 (69.4) |
| <I/e(I)> > 2.9 | 5.5 (2.9) |
| Completeness (%) | 98.0 (98.3) |
| Redundancy | 4.0 (4.1) |
| CC₁/₂ | 0.97 (0.61) |
| **Structure refinement** | |
| Resolution range (Å) | 25.2-2.06 |
| R-factor (%) | 18.6 |
| R-factor (high resolution shell) | 24.6 |
| R<sub>free</sub> (%) | 22.7 |
| R<sub>free</sub> (high resolution shell) | 27.7 |
| Total number of non-hydrogen atoms | 1700 |
| Protein atoms | 1609 |
| Water molecules | 82 |
| R.m.s. deviations | |
| Bond length (Å) | 0.006 |
| Bond angle (°) | 0.993 |
| Wilson B-factor (Å<sup>2</sup>) | 25.9 |
| Average B-factor protein atoms (Å<sup>2</sup>) | 40.3 |
| Average B-factor water atoms (Å<sup>2</sup>) | 45.5 |
| Ramachandran statistics (%) | |
| Most favored region | 97.4 |
| Allowed regions | 2.6 |
| Outliers | 0 |

Mammalian two-hybrid assay. This assay was performed using the CheckMate™ mammalian two-hybrid system (Promega) in 293T cells. For PML–PML interaction, the cDNAs of full-length WT PML isoform IV and mutants were engineered into pBIND and pACT vectors, respectively. For PML–partners interaction, the cDNAs of SUMOs/DAXX were cloned into pBIND vectors. The cDNAs of the WT and mutant PML isoform IV were engineered into pACT vectors. The 293T cells were then co-transfected with pG5-luc, pBIND-PML<sub>WT</sub>/pBIND-SUMOs/pBIND-DAXX and pACT-PML<sub>WT</sub>/PML<sub>mutants</sub> mixtures at 1:1 molar ratio using Lipofectamine 2000 transfection protocol (Invitrogen). Twenty-four hours after transfection, the 293T cells were harvested. The relative luciferase activities were monitored by the Dual-Luciferase Reporter Assay System (Promega).

PML NB biogenesis and PR speckle forming assays. The WT PML isoform I (UniProt code, P29590–1), WT PML isoform IV (UniProt code, P29590–5), and their mutants were engineered into pGFP-C1 vectors containing an N-terminal GFP tag, respectively. The GFP-tagged DAXX, SUMO2, and SUMO3 were cotransfected into the HeLa<sub>PML<sup>-/-</sup></sub> cells in which the endogenous pml gene was knocked out by the CRISPR-Cas9 strategy. The cells were grown in Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum before they were subjected to immunofluorescence examination. The cells were fixed with 4% paraformaldehyde. The slides were examined using a Leica TCS SP8 or Zeiss LSM870 confocal fluorescence microscope. It has been reported that PR can oligomerize and form small speckles in nucleoli<sup>39</sup>. This was used in this report to check the presence of B1 oligomerization in PR. Similar to the NB formation assay described above, the WT PR and mutants were engineered into pGFP-C1 vectors, followed by transfection into the HeLa<sub>PML<sup>-/-</sup></sub> cells. The expression and speckle formation activities were visualized by immunofluorescence.

ImmuNo co-localization assay. The PML partner proteins DAXX, SUMO2, and SUMO3 were engineered into pFLAG-CMV4 vectors, respectively. The FLAG-tagged DAXX, SUMO2, and SUMO3 were cotransfected into the HeLa<sub>PML<sup>-/-</sup></sub> cells with WT GFP-PML isoform I, WT GFP-PML isoform IV or their mutants at 1:1 molar ratio. Concerning arsenic treatment, the cells harboring the co-expression of PML and its partner protein were incubated with 2 μM As<sub>2</sub>O<sub>3</sub> for 1 h. For visualization, the cells were fixed with 4% paraformaldehyde and incubated for 10 min in a solution of 0.25% Triton X-100 in PBS. Then incubated with primary antibody anti-Flag (Sigma-Aldrich) overnight at 4 °C. Alexa 568-labeled anti-mouse secondary antibody was used for detection. The slides were examined with a Leica TCS SP8 or Zeiss LSM870 confocal fluorescence microscope.

PML sumoylation assay. HA-PML isoform IV, HA-PML isoform I, and derivatives derived from pLXV-RES-Puro vectors were co-transfected into HeLa<sub>PML<sup>-/-</sup></sub> cells with His-SUMO1/2/3. The cells were grown in Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum before western blot analysis. The sumoylation levels were detected using anti-HA antibody. Concerning arsenic response, the cells harboring WT PML and mutants were incubated with 2 μM As<sub>2</sub>O<sub>3</sub> for 1 h before lysis using RIPA buffer containing 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 1% PMSF, 1% protease inhibitor cocktail (Roche). The soluble and insoluble PMLs derived from sumoylation were fractionated by centrifugation (16,000g, 15 min). The pellets (i.e., the insoluble fraction) were solubilized by RIPA and 1% sodium dodecyl sulfate (SDS). Both the soluble and insoluble fractions were resolved in a 4–20% gradient SDS-PAGE and analyzed by immunoblot using anti-HA (Abcam ab9110, 1:5000 dilution) and anti-actin (Santa sc7778, 1:500 dilution) antibodies.

Luciferase reporter assay. The promoter regions harboring P11 motifs were cloned into the pGL3-basic luciferase reporter vector (Promega). The expression plasmids contained pLXV-RES-PR, pLXV-RES-PRL157E, pLXV-RES-PRLF158E and pCMV4-PU.1, respectively. The renilla luciferase plasmid pRL-SV40 (Promega) was used as control for transfection efficiency. The plasmids pGL-3-based pLXV-RES-PR/mutants, pCMV4-PU.1 and pRL-SV40 were all co-transfected to 293T cells using Lipofectamine 2000 (Invitrogen). Twenty-four hours after transfection, cells were harvested for the determination of luciferase activities using the Dual-luciferase reporter assay kit (Promega).

Transgenic mice. PR and PML-RARα1/158E (PR F158E) were expressed by human MRPS promoter. All animal experiments and care were carried out as approved by the Animal Care and Use Committee at Experimental Animal Centre in Shanghai Jiao Tong University School of Medicine, China. Standard checks such as HE staining, flow cytometry analysis of c-Kit, Gr-1 and Mac-1 in spleen and bone marrow cells were used to confirm APL development.

Single-cell RNA sequencing. The normal FVB/N mice (termet WT) and the PR and PR F158E transgenic mice at the same age (i.e., 78 weeks) were used for single-cell RNA sequencing (scRNA-Seq) analysis. All mouse experiments were conducted following the standard operating procedures approved by the Animal Care and Use Committee at Experimental Animal Centre in Shanghai Jiao Tong University School of Medicine, China. In brief, the bone marrow cells were firstly acquired.
from PR and PR F158E mice, the former of which has displayed an early stage of hematological disorders as monitored/confirmed by fluorescence-activated cell sorting analysis and enlarged spleen size. For single cell suspension, the phosphate-buffered saline (PBS) buffer with 0.04% bovine serum albumin was used. The cell counts and vitality were then verified by TC20® Automated Cell Counter (Thermo Fisher). Single cells were captured and barcoded in 10× Chromium Controller (10× Genomics). We used one mouse for each genotype. The single cell suspensions of PR and PR F158E were captured in one chip with two separated channels. The single cell suspension of WT was captured in another chip with one channel because of different dates for sample collection. Single cell RNA sequencing libraries were prepared using Chromium Single Cell 3’ V2 Reagent Kit (10× Genomics). Sequencing libraries were loaded on an Illumina NovaSeq with 2 × 150 paired-end kits. The FASTQ files were analyzed with the Cell Ranger Software Suite (version 2.0; 10× Genomics). We compared the raw data with mouse genome reference (mm10) to generate the filtered gene-barcode matrix which contained valid cell barcodes and transcript UMI counts with Cell Ranger Count. The single cell RNA sequencing statistics are listed in Supplementary Table 1. The aggregated gene-barcode matrix of three mice model was normalized by Cell Ranger aggr based on depth. Then, we used the normalized aggregated gene-barcode matrix to perform preliminary analysis of scRNA-seq, in which we performed the K-means method to cluster cells. Clusters were visualized with t-distributed Stochastic Neighbor Embedding of the principal components (t-SNE) as implemented in Cell Ranger. The cell type of each cluster was manually determined by recognized marker genes.

To validate the accuracy of the results of Cell Ranger, we also utilized Seurat62 to perform cell clustering. The filtered gene-barcode matrix of each mouse identified by Cell Ranger Count was inputted to Seurat. To further filter cells and genes in Seurat, we removed cell outliers (retain cells with 200–2500 genes), cells with high mitochondrial transcript proportion (>5%), as well as genes that were detected in less than three cells. The information of filtered cells and genes are listed in Supplementary Table 1. Then, we normalized the filtered gene-barcode matrix by total expression of each mouse. To avoid batch differences, we used the genes in Seurat, we removed cell outliers (retain cells with 200 detected in less than three cells. The information of identi

1. Carbone, R., Pearson, M., Minucci, S. & Pelicci, P. G. PML NBs associate with the hMre11 complex and p53 at sites of irradiation induced DNA damage. Oncogene 21, 1633–1640 (2002).
2. Guo, A. et al. The function of PML in p53-dependent apoptosis. Nat. Cell Biol. 11, 740–746 (2009).
3. Salomoni, P. & Pandolfo, P. P. The role of PML in tumor suppression. Cell 108, 165–170 (2002).
4. Pearson, M. & Pelcici, P. G. PML interaction with p53 and its role in apoptosis and replicative senescence. Oncogene 20, 7250–7256 (2001).
5. Pearson, M. et al. PML regulates p53 acetylation and premature senescence induced by oncogenic Ras. Nature 406, 207–210 (2000).
6. Bernardi, R. et al. PML inhibits HIF-1alpha translation and neoangiogenesis through repression of mTOR. Nature 442, 779–785 (2006).
7. Guan, D. & Kao, H. Y. The function, regulation and therapeutic implications of the tumor suppressor protein, PML. Cell Biosi. 5, 60 (2015).
8. Bernardi, R. & Pandolfo, P. P. Structure, dynamics and functions of promyelocytic leukemia nuclear bodies. Nat. Rev. Mol. Cell Biol. 8, 1006–1016 (2007).
9. Duprez, E. et al. SUMO-1 modification of the acute promyelocytic leukemia protein PML: implications for nuclear localisation. J. Cell Sci. 112, 381–393 (1999).
10. Scaglioni, P. P. et al. A CK2-dependent mechanism for degradation of the PML tumor suppressor. Cell 126, 269–283 (2006).
11. Grignani, F. et al. Fusion proteins of the retinoic acid receptor-alpha recruit histone deacetylase in promyelocytic leukemia. Nature 391, 815–818 (1998).
12. Jensen, K., Shiels, C. & Freemont, P. S. PML protein isoforms and the RBCCc/alpha motif. Oncogene 20, 7233–7237 (2001).
13. Hatakeyama, S. TRIM proteins and cancer. Nat. Rev. Cancer 11, 792–804 (2011).
14. Bodey, M. N., Duprez, E., Borden, K. & Freemont, P. S. Surface residue mutations of the pml ring finger domain alter the formation of nuclear matrix-associated pml bodies. J. Cell Sci. 110, 2197–2205 (1997).
15. Borden, K. L. et al. In vivo and in vitro characterization of the B1 and B2 zinc-binding domains from the acute promyelocytic leukemia promoncoprotein PML. Proc. Natl Acad. Sci. USA 93, 1601–1606 (1996).
16. Shen, T. H., Lin, H. K., Scaglioni, P. P., Yung, T. M. & Pandolfo, P. P. The mechanisms of PML-nuclear body formation. Mol. Cell 24, 331–336 (2006).
17. Zhong, S. et al. Role of SUMO-1-modified PML in nuclear body formation. Blood 95, 2748–2752 (2000).
18. Muller, S., Matunis, M. J. & Dejean, A. Conjugation with the ubiquitin-related modifier SUMO-1 regulates the partitioning of PML within the nucleus. EMBO J. 17, 61–70 (1998).
19. Sahin, U. et al. Oxidative stress-induced assembly of PML nuclear bodies controls cell proliferation and growth of promyelocytic leukemias. J. Cell Biol. 204, 931–945 (2014).
20. Weidkamp-Peters, S. et al. Dynamics of component exchange at PML nuclear bodies. J. Cell Sci. 121, 2731–2743 (2008).
21. Wang, P. et al. RING tetramerization is required for nuclear body biogenesis and PML sumoylation. Nat. Commun. 9, 1277 (2018).
22. de Thé, G., Rivière, M. & Bernhard, W. Examen au microscope électronique de la tumeur Vx2 du lapin domestique dérivée du papillome de SOPPE. Bull. Cancer 47, 570–584 (1960).
23. Brasch, K. & Ochs, R. L. Nuclear bodies (NBs): a newly “rediscovered” organelle. Exp. Cell Res. 202, 211–223 (1992).
24. Zhu, J. et al. Sumoylation site in PML/RARalpha targets to nuclear bodies. J. Biol. Chem. 2752 (2000).
25. Keown, J. R. & Goldstone, D. C. Structural crystal of the Trimafalpa box2 domain from rhesus macaque describes a plastic oligomerisation interface. J. Struct. Biol. 195, 282–285 (2016).
26. Huang, S. Y. et al. The B-box 1 dimer of human promyelocytic leukemia protein. J. Biol. NMR 60, 275–281 (2014).
27. Mosek, M. et al. Structural analysis of B-Box 2 from MuRf1: identification of a novel self-association pattern in a RING-like fold. Biochemistry 47, 10722–10730 (2008).
28. Wagner, J. M. et al. Mechanism of B-box 2 domain-mediated higher-order assembly of the retroviral restriction factor TRIMafalpa. eLife 5, pii: e15609 (2016).
29. Petrovskhov, M. V. et al. New developments in the ATLAS program package for small-angle scattering data analysis. J. Appl. Crystallogr. 45, 342–350 (2012).
30. Dycyk, J. A., Warrell, R. P. Jr., Evans, R. M. & Miller, W. H. Jr. Rapid diagnosis of acute promyelocytic leukemia by immunohistochiometric localization of PML/RAR-alpha protein. Blood 86, 862–867 (1995).
31. Wang, K. et al. PML/RARalpha targets promoter regions containing PU.1 consensus and RARE half sites in acute promyelocytic leukemia. Cancer Cell 17, 186–197 (2010).
32. Li, K. et al. TRIB3 promotes APL proliferation through stabilization of the oncoprotein PML-RARalpha and inhibition of p53-mediated senescence. Cancer Cell 31, 697–710.e697 (2017).
33. Borden, K. L. et al. Novel topology of a zinc-binding domain from a protein involved in regulating early xenograft development. EMBO J. 14, 5947–5956 (1995).
34. Tao, H. et al. Structure of the MID1 tandem B-boxes reveals an interaction reminiscent of intermolecular ring homodimers. Biochemistry 47, 2450–2457 (2008).
35. Zhu, J. et al. Arsenic-induced PML targeting onto nuclear bodies: implications for the treatment of acute promyelocytic leukemia. Proc. Natl Acad. Sci. USA 94, 3978–3983 (1997).
36. Occhioniorelli, M. et al. The self-association coiled-coil domain of PML is sufficient for the oncogenic conversion of the retinoic acid receptor (RAR) beta2 promoter. Leukemia 25, 814–820 (2011).
37. Cheng, G. X. et al. Distinct leukemia phenotypes in transgenic mice and different corepressor interactions generated by promyelocytic leukemia
variant fusion genes PLZF-RARalpha and NPM-RARalpha. Proc. Natl Acad. Sci. USA 96, 6318–6323 (1999).
38. Morrigl, R. et al. Stat5 tetramer formation is associated with leukemogenesis. Cancer Cell 7, 87–99 (2005).
39. Weiner, D. B. et al. Linkage of tyrosine kinase activity with transforming ability of the p185neu oncoprotein. Oncogene 4, 1175–1183 (1989).
40. Siegel, P. M. & Muller, W. J. Mutations affecting conserved cysteine residues within the extracellular domain of Neu promote receptor dimerization and activation. Proc. Natl Acad. Sci. USA 93, 8878–8883 (1996).
41. Pal, K. et al. Structural basis of TPR-mediated oligomerization and activation of oncogenic fusion kinases. Structure 25, 867–877.e863 (2017).
42. Wu, Y. M. et al. Identification of targetable FGFR gene fusions in diverse cancers. Cancer Discov. 3, 636–647 (2013).
43. Madison, D. L., Wirz, J. A., Siess, D. & Lundblad, J. R. Nicotinamide adenine dinucleotide-induced multimerization of the co-repressor CtBP1 relies on a switching tryptophan. J. Biol. Chem. 288, 27836–27848 (2013).
44. Jeffrey, P. D., Gorina, S. & Pavletich, N. P. Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 angstroms. Science 267, 1498–1502 (1995).
45. Fischer, N. W., Prodeus, A., Malkin, D. & Gariepy, J. p53 oligomerization status modulates cell fate decisions between growth, arrest and apoptosis. Cell Cycle 15, 3210–3219 (2016).
46. Grabocka, E. & Wedegaertner, P. B. Disruption of oligomerization induces nucleocytoplasmic shuttling of leukemia-associated rho Guanine-nucleotide exchange factor. Mol. Pharmacol. 72, 993–1002 (2007).
47. Zhang, X., Guresko, J., Shen, K., Cole, P. A. & Kuriyan, J. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell 125, 1137–1149 (2006).
48. Kurahashi, S. et al. PAX5–PML acts as a dual dominant-negative form of both PAX5 and PML. Oncogene 30, 1822–1830 (2011).
49. Liu, Y. et al. The tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO’s activity. Cancer Cell 9, 249–260 (2006).
50. So, C. W. & Cleary, M. L. Dimerization: a versatile switch for oncogenesis. Blood 104, 919–922 (2004).
51. Epuchi, M., Epuchi-Ishimae, M. & Greaves, M. The small oligomerization domain of gephyrin converts MLL to an oncogene. Blood 103, 3876–3882 (2004).
52. So, C. W., Lin, M., Ayton, P. M., Chen, E. H. & Cleary, M. L. Dimerization contributes to oncogenic activation of MLL chimeras in acute leukemias. Cancer Cell 4, 99–110 (2003).
53. Pollock, J. L. et al. A bcr-3 isoform of RARalpha–PML potentiates the development of PML–RARalpha-driven acute promyelocytic leukemia. Proc. Natl Acad. Sci. USA 96, 15103–15108 (1999).
54. Zeisig, B. B. et al. Recruitment of RXR by homotetrameric RARalpha fusion proteins is essential for transformation. Cancer Cell 12, 36–51 (2007).
55. Kwok, C., Zeisig, B. B., Dong, S. & So, C. W. Forced homo-oligomerization of RARalpha leads to transformation of primary hematopoietic cells. Cancer Cell 9, 95–108 (2006).
56. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760–763 (1994).
57. Perrakis, A., Morris, R. & Lamzin, V. S. Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. 6, 458–463 (1999).
58. Winn, M. D., Isupov, M. N. & Murshudov, G. N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr. D Biol. Crystallogr. 57, 122–133 (2001).
59. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221 (2010).
60. WMW, R. A. Laskowski, Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26, 283–291 (1993).
61. Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. & Schubert, D. Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. Biophys. J. 82, 1096–1111 (2002).
62. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36, 411–420 (2018).
63. Engh, R. A. & Huber, R. Accurate bond and angle parameters for X-ray protein structure refinement. Acta Crystallogr. A47, 392–400 (1991).

Acknowledgements
This work was supported by research grants 81770142, 81370620, 81570120, 31070645, 81800144, and 31800642 from National Natural Science Foundation of China, a research grant 20152504 from “Shanghai Municipal Education Commission—Gaofeng Clinical Medicine Grant Support”, “The Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institute of Higher Learning”, a research grant 111C1407200 from SMSTC, a research grant 12ZZ019 from SME, Program for New Century Excellent Talents in University (NCET-10–9571), Samuel Waxman Cancer Research Foundation. We thank the staffs Shan Li, Wei Qi, Zhixin Wang from Novogene Bioinformatics Technology Co., Ltd., Beijing for assistance during single cell RNAseq analysis.

Author contribution
Conceived and designed the experiments: G.M. Performed the experiments: Y.L., X.M., Z.M.C., H.W., P.W., W.W., N.C., L.Z. and H.Z. Analyzed the data: Y.L., X.M., Z.M.C., H.W., P.W., W.W., N.C., L.Z., H.Z., X.C., S.-J.C., Z.C. and G.M. Preparation of the figures: Y.L. and G.M. Wrote the paper: Y.L. and G.M. Supervised this work: G.M.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-11746-0.

Competing interest The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Peer review information Nature Communications thanks LLucia Alberti-Servera, Nathan Cowieson and other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, and indicate if changes were made. The copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.