



American Society of Hematology  
 2021 L Street NW, Suite 900,  
 Washington, DC 20036  
 Phone: 202-776-0544 | Fax 202-776-0545  
 editorial@hematology.org

## Systems medicine dissection of chr1q-amp reveals a novel PBX1-FOXM1 axis for targeted therapy in multiple myeloma

Tracking no: BLD-2021-014391R1

Nikolaos Trasanidis (Imperial College London, United Kingdom) Alexia Katsarou (Imperial College London, United Kingdom) Kanagaraju Ponnusamy (Imperial College London, ) Yao-An Shen (Johns Hopkins University School of Medicine, United States) Ioannis Kostopoulos (Imperial College London, United Kingdom) Bien Bergonia (Imperial College London, United Kingdom) Keren Keren (Imperial College London, United Kingdom) Paudel Reema (Imperial Experimental Cancer Medicine Centre and Cancer Research UK Imperial Centre, United Kingdom) Xiaolin Xiao (Imperial College London, ) Richard Szydlo (Imperial College / Hammersmith Hospital, United Kingdom) Pierangela Sabbattini (Imperial College London, United Kingdom) Irene Roberts (Oxford University, United Kingdom) Holger Auner (Imperial College London, United Kingdom) Kikkeri Naresh (Imperial College Healthcare NHS Trust, United Kingdom) Aristeidis Chaidos (Imperial College London, United Kingdom) Tian-Li Wang (Johns Hopkins University School of Medicine, United States) Luca Magnani (Imperial College London, United Kingdom) Valentina Caputo (London South Bank University, United Kingdom) Anastasios Karadimitris (Imperial College London, Hammersmith Hospital, United Kingdom)

### Abstract:

Understanding the biological and clinical impact of copy number aberrations (CNA) for the development of precision therapies in cancer remains an unmet challenge. Genetic amplification of chromosome 1q (chr1q-amp) is a major CNA conferring adverse prognosis in several types of cancer, including in the blood cancer multiple myeloma (MM). Although several genes across chr1q portend high-risk MM disease, the underpinning molecular aetiology remains elusive. Here, with reference to the 3D chromatin structure, we integrate MM patient multi-omics datasets with genetic variables to obtain an associated clinical risk map across chr1q and to identify 103 adverse prognosis genes in chr1q-amp MM. Prominent amongst these genes, the transcription factor *PBX1* is ectopically expressed by genetic amplification and epigenetic activation of its own preserved 3D regulatory domain. By binding to reprogrammed super-enhancers, *PBX1* directly regulates critical oncogenic pathways and a *FOXM1*-dependent transcriptional programme. Together, *PBX1* and *FOXM1* activate a proliferative gene signature which predicts adverse prognosis across multiple types of cancer. Notably, pharmacological disruption of the *PBX1*-*FOXM1* axis with existing agents (thiostrepton) and a novel *PBX1* small-molecule inhibitor (T417) is selectively toxic against chr1q-amplified myeloma and solid tumour cells. Overall, our systems medicine approach successfully identifies CNA-driven oncogenic circuitries, links them to clinical phenotypes and proposes novel CNA-targeted therapy strategies in multiple myeloma and other types of cancer.

**Conflict of interest:** No COI declared

**COI notes:**

**Preprint server:** No;

**Author contributions and disclosures:** NT designed study, conceived and implemented computational pipelines, designed and performed experiments, wrote manuscript. AlK, KP, YS, IK, BB, and KK performed experiments. XX assisted in bioinformatics data analysis. PD and NK performed and interpreted IHC analysis. RMS assisted in clinical informatics data analysis. AC and HWA provided clinical samples. IAGR, TL and LM wrote manuscript. VSC and AnK: designed study, supervised experiments, generated draft manuscript. All authors contributed to the final draft of the manuscript.

**Non-author contributions and disclosures:** No;

**Agreement to Share Publication-Related Data and Data Sharing Statement:** High-throughput sequencing data generated during this study have been deposited to the Gene Expression Omnibus repository (GEO): MMCL ChIP-seq and RNA-seq files (GSE165060) and primary MM ATAC-seq files (GSE153381). Code used in this study can be accessed from the specified github page: [https://github.com/nikostrasan/PBX1\\_project](https://github.com/nikostrasan/PBX1_project) For all non-commercial reagents (e.g. plasmid constructs, etc.) used in our study or any other materials or methods, please contact the corresponding authors: Dr Nikolaos Trasanidis (ntl12@ic.ac.uk), Dr Valentina Caputo (v.caputo@imperial.ac.uk), Prof. Anastasios Karadimitris (a.karadimitris@imperial.ac.uk).

Clinical trial registration information (if any):

## Systems medicine dissection of chr1q-amp reveals a novel PBX1-FOXM1 axis for targeted therapy in multiple myeloma

Nikolaos Trasanidis<sup>1</sup>, Alexia Katsarou<sup>1,2</sup>, Kanagaraju Ponnusamy<sup>1</sup>, Yao-An Shen<sup>3,4,5</sup>, Ioannis V Kostopoulos<sup>1,6</sup>, Bien Bergonia<sup>1</sup>, Keren Keren<sup>1</sup>, Paudel Reema<sup>7</sup>, Xiaolin Xiao<sup>1</sup>, Richard M Szydlo<sup>1</sup>, Pierangela MR Sabbattini<sup>1</sup>, Irene AG Roberts<sup>8</sup>, Holger W Auner<sup>1,2</sup>, Kikkeri N Naresh<sup>2,7</sup>, Aristeidis Chaidos<sup>1,2</sup>, Tian-Li Wang<sup>3,4</sup>, Luca Magnani<sup>9</sup>, Valentina S Caputo<sup>1,10</sup>, Anastasios Karadimitris<sup>1,2</sup>

1. Hugh & Josseline Langmuir Centre for Myeloma Research, Centre for Haematology, Department of Immunology and Inflammation, Imperial College London, London, United Kingdom
2. Department of Haematology, Hammersmith Hospital, Imperial College Healthcare NHS Foundation Trust, London, United Kingdom.
3. Departments of Pathology, Oncology and Gynecology and Obstetrics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
4. Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
5. Department of Pathology, School of Medicine, College of Medicine, Taipei Medical University, Taipei 110, Taiwan
6. Department of Biology, School of Science, National and Kapodistrian University of Athens, Athens, Greece
7. Imperial Experimental Cancer Medicine Centre and Cancer Research UK Imperial Centre, London, United Kingdom
8. Department of Paediatrics and MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, Oxford University and BRC Blood Theme, NIHR Oxford Biomedical Centre, Oxford, UK
9. Department of Surgery and Cancer, Imperial College London, London, United Kingdom
10. Cancer Biology and Therapy laboratory, School of Applied Science, London South Bank University, London, UK

### Correspondence

Anastasios Karadimitris  
Room 4S10C, Commonwealth Building  
Hammersmith Campus, Imperial College London  
Tel: +44 (0)20 3313 8438  
Email: a.karadimitris@imperial.ac.uk

Nikolaos Trasanidis  
Room 4S10, Commonwealth Building  
Hammersmith Campus, Imperial College London  
Tel: +44 (0) 7541928837  
Email: nikolaos.trasanidis12@imperial.ac.uk

Valentina Caputo  
Room 4S10, Commonwealth Building  
Hammersmith Campus, Imperial College London  
Tel: +44 (0)78 9148 2156  
Email: v.caputo@imperial.ac.uk

## Key points

- Clinical multi-omic analysis unveils a core PBX1-FOXM1 regulatory axis underlying high-risk proliferative phenotypes in chr1q-amp myeloma
- Preclinical profiling of a novel PBX1 inhibitor (T417) shows selective potency and supports its use against chr1q-amp myeloma

## Abstract

Understanding the biological and clinical impact of copy number aberrations (CNA) for the development of precision therapies in cancer remains an unmet challenge. Genetic amplification of chromosome 1q (chr1q-amp) is a major CNA conferring adverse prognosis in several types of cancer, including in the blood cancer multiple myeloma (MM). Although several genes across chr1q portend high-risk MM disease, the underpinning molecular aetiology remains elusive. Here, with reference to the 3D chromatin structure, we integrate MM patient multi-omics datasets with genetic variables to obtain an associated clinical risk map across chr1q and to identify 103 adverse prognosis genes in chr1q-amp MM. Prominent amongst these genes, the transcription factor *PBX1* is ectopically expressed by genetic amplification and epigenetic activation of its own preserved 3D regulatory domain. By binding to reprogrammed super-enhancers, PBX1 directly regulates critical oncogenic pathways and a FOXM1-dependent transcriptional programme. Together, PBX1 and FOXM1 activate a proliferative gene signature which predicts adverse prognosis across multiple types of cancer. Notably, pharmacological disruption of the PBX1-FOXM1 axis with existing agents (thiostrepton) and a novel PBX1 small-molecule inhibitor (T417) is selectively toxic against chr1q-amplified myeloma and solid tumour cells. Overall, our systems medicine approach successfully identifies CNA-driven oncogenic circuitries, links them to clinical phenotypes and proposes novel CNA-targeted therapy strategies in multiple myeloma and other types of cancer.

**Keywords:** Copy number aberrations, chr1q amplification, systems medicine, multiple myeloma, regulatory networks, PBX1, FOXM1, PBX1 inhibitor

Text word count: 4,332

Abstract word count: 214

No of figures: 6

No of references: 53

## Introduction

Genetic amplification of chr1q (chr1q-amp), one of the most frequent copy number aberrations (CNA), confers adverse prognosis in cancer<sup>1-3</sup>. In multiple myeloma (MM), an incurable cancer of the B lineage plasma cells (PC), chr1q-amp is a secondary genetic event present in 30-40% of patients at diagnosis and is associated with adverse prognosis, high-burden proliferative disease and drug resistance<sup>4-9</sup>. In addition, along with t(4;14) and del(17p), chr1q-amp is one of the top three genetic markers conferring adverse overall and progression free survival in MM<sup>10,11</sup>.

Previous studies, often guided by low resolution methodologies (e.g., FISH against 1q21 locus<sup>12</sup>), identified several chr1q21 genes associated with adverse prognosis in MM, including the *CKS1B*, *PDKZ1*, *ILF2*, *ARNT*, *ADAR1*, *IL6R*, *MCL1*, *BCL9* and *MDM2* genes<sup>13-20</sup>. However, genetic amplification that extends beyond chr1q21 has been previously reported in a small cohort of MM patients<sup>21</sup>, and non-chr1q21 genes (e.g., *CD1D*, *FCER1G*) have been linked to the biology and prognosis of MM<sup>22-24</sup>. These observations raise the prospect that several areas across chr1q may contribute to the biological profile and clinical impact of chr1q-amp.

Further, how genetic amplification affects the 3D chromatin architecture of chr1q and influences biological processes that promote high risk disease is not known. Understanding these processes could inform novel anti-cancer therapeutic approaches targeted to chr1q-amp that are currently lacking.

Here we employed a comprehensive systems medicine approach to resolve the 3D genome landscape of chr1q-amp and to integrate it with multi-omics MM patient datasets. This approach led to the identification of adverse prognosis genes across the whole chr1q arm, and particularly in the 1q22 and 1q23.3 bands. Amongst 1q23.3-associated genes, we identified the transcription factor PBX1, which, in co-operation with FOXM1, regulates myeloma PC proliferation and generates a selective therapeutic vulnerability in chr1q-amp MM that can be targeted by a novel PBX1 inhibitor.

## Methods

### Cell cultures

All cell lines and culturing conditions used in this study are detailed in the **Supplemental methods** section.

### Primary samples

Bone marrow aspirate samples from MM patients and peripheral blood samples from normal donor were obtained upon written informed consent and under research ethics committee approval (Research Ethics Committee Reference: 11/H0308/9). Bone marrow aspirates were subjected to red cell lysis. MM plasma cells were purified after two rounds of CD138 immunomagnetic selection (Miltenyi Biotec) following the manufacturer's instructions, as previously described<sup>25</sup>. Pre- and post-selection purity was assessed by flow-cytometric analysis (BD LSR-Fortessa) using a panel of fluorochrome-labelled anti-CD138, -CD45, -CD19, -CD56 and -CD38 monoclonal antibodies. Purified cells were immediately processed for ATAC-seq and RNA-seq analysis or stored in FBS + 10% DMSO at -150°C for later use.

Mononuclear cells from normal donor peripheral blood were isolated by Ficoll-Hypaque (Sigma-Aldrich) density centrifugation following the manufacturer's instructions, as described before<sup>26</sup>. The mononuclear cell interphase layer was aspirated, washed with 1ml PBS, centrifuged at 300g for 5min and resuspended in 100µl PBS. Peripheral blood B cells (PBBC) were isolated using the human Total B cell isolation kit II (Miltenyi Biotec) as per manufacturer's instructions.

## Molecular cloning

A modified pLKO.1 lentiviral vector (Addgene plasmid #27994) was used, in which, the puromycin marker gene was replaced by *eGFP* (for knock-down experiments) or *eBFP* (for rescue experiments). All shRNA oligos were cloned, as previously described<sup>27</sup>: scrambled (scrbl) control, 5'-CCTAAGGTTAAGTCGCCCTCG-3'; P11 (anti-*PBX1*), 5'-CGAAGCAATCAGCAAACACAA-3'; P31 (anti-*PBX1*), 5'-ATGATCCTGCGTTCCCGATT-3'; O1 (anti-*FOXM1*), 5'-CTCTTCTCCCTCAGATATAGA-3'; O4 (anti-*FOXM1*), 5'-GCCAATCGTTCTCTGACAGAA-3'. Successful cloning of recombined vectors was initially confirmed via diagnostic PCR, using the DreamTaq Green PCR Master Mix (2X) (Thermo Scientific) protocol and the 5'-TGGACTATCATATGCTTACCGTAAC-3' (F) and 5'-GTATGTCTGTTGCTATTATGTCTA-3' (R) primers, followed by 1% agarose gel electrophoresis. The DNA sequence of positive clones was further confirmed via Sanger Sequencing (outsourced to GeneWiz Ltd), using the same primers set.

Molecular cloning of the overexpression plasmid vector is described in the **Supplemental methods** section.

## RNA-seq

Total RNA was extracted from FACS-sorted myeloma plasma cells using the Nucleospin RNA kit (Macherey-Nagel). The Qubit RNA Assay kit (Life Technologies) was used to determine the RNA quantity. Quality of RNA extracts was assessed on the Bioanalyser using the RNA pico kit (Agilent). Samples with RIN value higher than 8 were processed using the NEBNext poly(A) mRNA Magnetic Isolation kit and the NEBNext Ultra II RNA Library Prep kit for Illumina (New England Biolabs), following manufacturer's instructions. The Qubit High Sensitivity DNA kit (Life Technologies) was used for library quantification; library size was evaluated using the Bioanalyser High Sensitivity DNA kit (Agilent). Libraries from the same experiment were diluted to 5nM, pooled together and sequenced at the BRC Genomics Facility (Imperial College London) using the Illumina HiSeq 4000 platform to obtain paired-end 75bp reads.

## ATAC-seq

ATAC-seq was performed as previously described<sup>28</sup>. Briefly, 50,000 purified myeloma plasma cells or myeloma cell lines, were washed with cold PBS (Sigma) at 500g at 4°C for 5 min. The cells were resuspended in 50 µL of cold Lysis Buffer (10 mM Tris-HCl, pH 7.4, 10mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% IGEPAL CA-630) and washed at 500g at 4°C for 10min. The nuclei were subjected to transposase reaction for 30min at 37°C; termination of the reaction and DNA purification was performed using a MiniElute Kit (Qiagen) and eluted twice with 10 µL. The purified DNA was amplified as described before with NEBNext High-Fidelity 2x PCR Master Mix (New England Biolabs). The PCR amplified product was cleaned twice with (0.9X) AMPure beads (Beckman). The quality of the libraries was assessed with the Bioanalyzer High Sensitivity DNA kit (Agilent). The libraries were quantified using the NEBNext Library Quant Kit for Illumina (New England Biolabs) on a StepOne Plus Real-Time PCR (Applied Biosystems). The libraries were sequenced at the Genomics Facility at ICL using the Illumina HiSeq 4000 platform to obtain paired-end 75bp reads.

## Bioinformatics and clinical informatics analysis

All methods used for statistical, bioinformatics and clinical informatics analysis are described in the **Supplemental methods** section.

## Data and code availability

High-throughput sequencing data generated during this study have been deposited to the Gene Expression Omnibus repository (GEO): MMCL ChIP-seq and RNA-seq files (GSE165060) and primary MM ATAC-seq and RNA-seq files (GSE153381).

Code used in this study can be accessed from the specified github page: <https://github.com/nikostrasan/PBX1-project>

Additional materials and methods information are described in the **Supplemental methods** section.

## Results

### Distinct patterns of amplification within chr1q shape its 3D chromatin architecture

We first explored whether and how genomic structural changes might impact the 3D chromatin structure of chr1q-amp myeloma cells. For this purpose, we constructed a correlation matrix of copy number scores across the chr1q arm (2D genome co-amplification map) using whole genome sequencing (WGS) data from MM patients (MMRF CoMMpass database<sup>29</sup>, n=896) and compared it with the 3D genome Hi-C contact maps of two chr1q-amp MM cell lines (MMCL; U266, RPMI8226<sup>30</sup>; **Figure 1A**). By applying to the 2D genome map the same computational method used for topologically associated domain (TAD) discovery<sup>30</sup>, we found four main blocks (B1-B4) of genomic co-amplification (termed topologically co-amplified domains; TCDs), which define distinct amplification patterns across MM patients (**Figure 1A and Supplemental Figure S1A**). Next, to understand the relationship between chr1q-amp and 3D chromatin structure, we compared the myeloma TCD and TAD maps, along with the Hi-C map of the reference GM12828 B cells<sup>31</sup> (**Supplemental Figure S1B**). First, we found that TAD organization was highly similar between the U266 and RPMI8226 myeloma cells but highly dissimilar amongst them and the GM12828 B cells (**Supplemental Figure S1B**), suggesting extensive re-organization of the chr1q-amp 3D genome in myeloma cells. Second, we identified weak overall similarity between the WGS co-amplification map and the 3D genome maps of U266, RPMI8226 and GM12828 cells, when we compared their insulation score profiles (**Supplemental Figure S1B**). In a complementary analysis that aimed to identify the genetic amplification breakpoints in all patients, we found that approximately 65% of the 135 GM12828 reference TADs are disrupted by chr1q-amp breakpoints (**Supplemental Figure S1C**) supporting further the notion that genetic amplification extensively disrupts the 3D chromatin architecture of chr1q. In both analyses, we also identified a pericentromeric cluster of multiple breaks as previously described<sup>32</sup>.

Finally, despite the extensive 3D chromatin re-organization within the TCDs, we observed that the B1-B4 borders coincide with TAD borders in chr1q-amp MM, suggesting that organization of these four hyper-domains evolves in parallel with some level of 3D chromatin structure retention and preservation. (**Supplemental Figure S1A, S1D**)

### Systems medicine analysis identifies adverse prognosis drivers beyond 1q21

Next, to identify all genes across chr1q that could potentially drive high-risk phenotype in MM and with reference to the 3D chromatin structure, we combined genomic (WGS; WES; whole exome sequencing), epigenomic (H3K27ac-seq), and transcriptomic (RNA-seq, DNA microarray) data with genetic variables from three studies: MMRF (n= 896); Arkansas (n=414); and *Jin2018* (n=12)<sup>29,30,33,34</sup> (**Figure 1B**). Of the 2,215 chr1q genes, we considered as candidate drivers of adverse prognosis only those fulfilling each of the following criteria: (1) their genetic amplification predicts adverse prognosis, independent of the prognostic impact of 73 other molecular markers (MMRF dataset; **Supplemental Figure S1E**); (2) their genetic amplification is significantly associated with their transcriptional overexpression (MMRF dataset); (3) their overexpression is significantly correlated with adverse prognosis (MMRF and Arkansas datasets); (4) their genetic amplification is accompanied by epigenetic activation (i.e., H3K27ac signal gain compared to non-amplified MM; *Jin2018* dataset); (**Figure 1C and Supplemental Table S1**).

This stepwise analysis identified 103 candidate genes residing exclusively in B1 and B4 hyper-domains, including the previously known *MCL1*, *CKS1B*, *ILF2* and *ARNT* genes located in chr1q21.3<sup>13-15</sup> (**Figure 1C**). Pathway analysis of all 103 genes showed significant enrichment for cell cycle-related processes, suggesting their involvement in the proliferative phenotype that is associated with chr1q-amp MM<sup>17</sup> (**Supplemental Figure S1F**). Interestingly, we identified 1q22 and 1q23.3 as the two cytogenetic bands that, relative to their gene density, contain the highest number of candidate adverse prognosis genes (**Supplemental Figure S1G**), with 1q23.3 displaying the highest association with adverse prognosis (**Figure 1C and 1D**). Therefore, there are additional regions, other than 1q21, which contribute to the high-risk, proliferative phenotype linked to chr1-amp in MM.

### **PBX1 is a novel biomarker of chr1q genetic amplification**

Amongst 1q23.3 genes, the transcription factor *PBX1* previously reported to promote cancer cell survival, metastasis and drug resistance<sup>35-37</sup> was notable for the highest H3K27ac signal gain across its own preserved TAD (**Figure 1E and Supplemental Figure S1D and S1H**). These features comprise a unique case of amplification of an entire regulatory domain linked to epigenetic activation, gene overexpression and adverse prognosis. Further analysis using the MMRF dataset confirmed *PBX1* as a marker of high-risk MM disease, with its amplification significantly correlating with its overexpression (**Supplemental Figure S2A and S2B**), while *PBX1* overexpression was associated with high-risk clinical features, high myeloma plasma cell proliferative index, progressive/relapsed disease and worse overall survival (**Supplemental Figure S2C-S2J**).

### **The pro-proliferative role of PBX1 in chr1q-amp MM**

We explored further the functional role of *PBX1* in chr1q-amplified MM cells, by assessing its mRNA and protein expression levels across healthy and tumour cells. Based on RNA-seq data, we found that in normal hematopoiesis, *PBX1* is expressed in bone marrow hematopoietic stem and progenitor cells as well as megakaryocytes, but not in B cells or plasma cells (**Supplemental Figure S3A**). In MM, we confirmed ectopic expression of *PBX1* in four chr1q-amp MMCL by RT-qPCR (**Figure 2A**) and in 9/11 patient myeloma PC samples with FISH-verified chr1q-amp by RT-qPCR and immunohistochemistry, using breast cancer as a positive control (**Figure 2B and Supplemental Figure S3B and S3C**). As expected, tonsillar germinal center B cells and submucosal plasma cells stained negative for *PBX1* expression (**Supplemental Figure S3C**).

Depletion of *PBX1* using two validated shRNAs (P31, P11) and assessed by GFP marker expression was toxic to MM.1S and U266 cells compared to scrambled shRNA control *in vitro* (**Figure 2C and Supplemental Figure S3D**) and impaired myeloma cell growth (MM.1S) in an *in vivo* subcutaneous MM model (**Figure 2D-2F and Supplemental Figure S3E-S3G**). To gain further insights, we performed RNA-seq analysis in both MMCL upon shRNA-mediated *PBX1* depletion (**Figure 2G-2H and Supplemental Table S2**). Transcriptome profiling of *PBX1*-depleted cells showed similar numbers of genes de-regulated in the two MMCL, while Gene Set Enrichment Analysis revealed significant enrichment for cell cycle-related pathways amongst down-regulated and interferon response pathways in up-regulated genes (**Figure 2H**). This is consistent with the reported enrichment of interferon response pathways in early-stage, non-proliferative MM and of cell cycle-related pathways in advanced disease and MMCL<sup>38,39</sup>. Accordingly, flow-cytometric analysis showed significant G1-phase cell cycle arrest in *PBX1*-depleted MMCL (**Figure 2I, Supplemental Figure S3H**).

### **Defining the epigenetic and regulatory programme of PBX1 in chr1q-amp cells**

ChIP-seq analysis against *PBX1* in MM.1S and U266 cells identified 30,000-40,000 binding sites (**Figure 3A and Supplemental Table S2**). Further annotation using chromHMM maps (built upon

ENCODE/Blueprint Consortium data) showed that 60-80% of PBX1 recruitment occurs in active-chromatin promoter and enhancer areas, while motif enrichment analysis identified the PBX1 motif among the top hits (**Figure 3A and Supplemental Figure S4A-S4D**). Additional analysis of H3K27ac-seq profiles from eight primary myeloma PC and nine MMCL<sup>34</sup> identified 2,400 super-enhancers (SEs), 70% of which are PBX1-bound (**Figure 3B**). Sample stratification based on chr1q-amp status showed significantly higher H3K27ac signal in PBX1-bound SEs in chr1q-amplified versus non-amplified cells, suggesting extensive epigenetic reprogramming associated with PBX1 binding in chr1q-amplified myeloma cells (**Figure 3C and Supplemental Figure S4E-S4F**). Interestingly, the PBX1-bound SEs in chr1q-amplified cells are predicted to regulate critical cellular pathways, including cell cycle (**Figure 3D**).

Next, we integrated the PBX1 cistrome with the PBX1-depleted transcriptomes to generate the gene regulatory network of PBX1 in chr1q-amplified cells (**Figure 3E and Supplemental Figure S4G-S4I and Supplemental Table S3**). We identified approximately 700 and 300 genes to be directly activated and repressed, respectively, by PBX1 in both MM.1S and U266 MMCL. Again, among other prominent oncogenic pathways, the former were primarily enriched in cell cycle-related biological processes and the latter in interferon response pathways (**Figure 3E**).

### **The PBX1-FOXM1 axis regulates cell proliferation in chr1q-amp MM**

Amongst the PBX1-dependent targets, we detected significant enrichment of the pro-proliferative *FOXM1* and *E2F* transcription factors and their corresponding targets (**Figure 3F**), such as the FOXM1-dependent *NEK2* that regulates drug resistance in MM<sup>40,41</sup>, (**Figure 4A**). Further, we identified PBX1 binding on active *PBX1*, *E2F1/2*, *NEK2* promoters and *PBX1*, *FOXM1*, *E2F2*, *NEK2* enhancers (**Figure 4B**), while FOXM1 was found to bind to the same *FOXM1* and *NEK2* regions as PBX1 (**Supplemental Figure Fig S5A**). To better explore the regulatory interplay among those factors (**Figure 4A**), we characterized further the role of FOXM1 in chr1q-amp cells. Knockdown of *FOXM1* using two validated shRNAs was toxic to MM.1S cells (**Figure 4C**), as previously shown<sup>40</sup>. In addition, depletion of *FOXM1* mRNA was associated with downregulation of *NEK2* but not of *PBX1* (**Figure 4D**), suggesting that FOXM1 acts downstream of PBX1 (**Figure 4A**). Moreover, RNA-seq analysis revealed approximately 800 differentially expressed genes after *FOXM1* knockdown in MM.1S cells (**Figure 4E**), with cell cycle-related pathways found to be significantly enriched amongst downregulated genes (**Figure 4F**). Cell cycle arrest at G2/M was corroborated by flow-cytometry, thus confirming the pro-proliferative role of FOXM1 in chr1q-amplified MMCL (**Supplemental Figure S5B**).

For further validation of the PBX1-FOXM1 axis, we forced expression of exogenous *PBX1* into MM.1S and NCU.MM1 chr1q-amplified MM cells (**Figure 4G**). This led to modest but significant increase in *FOXM1*, *NEK2* and *E2F2* mRNA levels (**Figure 4H**) and significantly reduced sensitivity of the MMCL to thiothrepton, an inhibitor of FOXM1 transcription<sup>40,42</sup> (**Figure 4I and Supplemental Figure S5C**). Rescue of *PBX1* depletion by shRNA-resistant *PBX1* cDNA resulted in a significantly lower MMCL toxicity, ameliorated cell cycle arrest and dampened downregulation of *FOXM1*, *NEK2* and *E2F2* (**Supplemental Figure S5D-5G**), thus validating the genetic and functional interactions in the PBX1-FOXM1 axis (**Figure 4A**) and its role in orchestrating an oncogenic, proliferative process in chr1q-amp MM cells.

### **The PBX1-FOXM1 regulatory axis generates a selective therapeutic vulnerability in primary chr1q-amp MM cells**

Next, we sought to validate activity of the PBX1-FOXM1 axis in primary myeloma plasma cells (**Figure 5A**). For this purpose, we combined RNA-seq with ATAC-seq profiling of highly purified chr1q-amplified (n=6) and non-amplified (n=6) primary myeloma PC, and explored differences in chromatin accessibility, gene expression and predicted TF connectivity (**Figure 5A and Supplemental Table S4**). In addition to previously established gene-markers (*CKS1B*, *IL6R*, *ARNT*, *PDKZ1*, *ADAR*), we also found overexpression of all main PBX1-FOXM1 module components (*PBX1*, *FOXM1*, *E2F1/2*, *NEK2*) in chr1-amp cells (**Figure 5B**). Moreover, there was significant enrichment of proliferative pathways and FOXM1-dependent targets in genes overexpressed in chr1q-amp cells (**Figure 5C**). Comparative ATAC-seq analysis revealed enhanced chromatin accessibility in the regulatory regions of genes over-expressed in the same cells (**Figure 5D**). Differential TF footprinting analysis revealed a higher number of TFs with increased connectivity (measured as differential regulatory potential,  $\Delta P$ ) in chr1q-amp versus non-amplified cells (**Figure 5E**). By combining transcriptional and regulation profiles, we identified 34 TFs with increased expression and connectivity in chr1q-amplified cells, including all four TFs involved in the PBX1-FOXM1 module (*PBX1*, *FOXM1*, *E2F1*, *E2F2*; **Figure 4A and 5F**). Notably, as compared to non-amplified cells (n=3), chr1q-amplified primary myeloma cells (n=3) were selectively sensitive to thiostrepton treatment, while expression of *FOXM1* and *NEK2*, but not *PBX1*, decreased in response to treatment (**Figure 5G and 5H**).

In addition, we validated functional activation of the PBX1 and shared PBX1-FOXM1 transcriptional programmes in a large cohort of patients (MMRF, n=813) and confirmed significant co-expression of *PBX1* and *FOXM1* with almost all of their gene targets across patients in two different cohorts (MMRF, Arkansas; **Supplemental Table S5 and Figure S6A**). Importantly, the majority of genes previously shown to comprise high-risk disease signatures in MM<sup>17,41,43,44</sup> were found to be directly regulated by PBX1 (**Supplemental Table S5 and Figure S6B**). Together, these findings strongly support the critical role of PBX1-FOXM1 axis in promoting proliferative regulatory circuitries determining adverse prognosis and high-risk disease in chr1q-amp MM patients.

### Targeted therapy against chr1q-amp in cancer using a novel, selective PBX1 inhibitor

As the PBX1-FOXM1 axis acts as a central regulatory hub for chr1q-amp MM cells, we next sought to explore the prognostic impact and therapeutic potential of selective PBX1 targeting in chr1-amp cells across several types of cancer. For this purpose, we first analysed transcriptomic data from multiple cancer patient cohorts and found that activation of the PBX1-dependent regulatory signature (n=320 genes) predicts adverse prognosis in multiple myeloma and 12 solid tumour patient cohorts, including breast, ovarian, lung and brain cancer, in which chr1q-amp is a frequent CAN<sup>2,3</sup> (**Figure 6A and Supplemental Figure S7A**). Next, we tested the impact of our novel, recently reported small-molecule compound T417, which specifically inhibits PBX1 binding to its cognate DNA motif<sup>45</sup> in chr1q-amp ovarian cancer cells. We screened four myeloma (MM.1S, U266, NCU.MM1, OPM2), two breast (MCF-7, LTED), two ovarian (OVCAR3, A2780), two lung (A549, H69AR) and one brain (SNB-75) cancer cell lines harboring at least one additional chr1q copy (**Supplemental Figure S7B**). Cell viability assays revealed sensitivity of all cell lines to T417 at low  $\mu\text{M}$  concentrations (4-28 $\mu\text{M}$ ), while no significant toxicity was detected upon treatment with the inactive analogue/pro-drug compound DHP52 in two myeloma and two ovarian cancer cell lines (**Figure 6B and Supplemental Figure S7C**). In addition, cell cycle analysis revealed significant depletion of the G2/M phase along with G0/1 phase arrest upon T417 treatment in all tested cell lines (**Figure 6C**). RT-qPCR-assessed mRNA levels of the PBX1-regulated *FOXM1*, *NEK2* and *E2F2* genes showed their significant decrease upon treatment with T417 in almost all 11 cell lines (**Figure 6D**). Interestingly, a significant decrease of *PBX1* mRNA itself was also detected in 8 out of 11 cell lines. This, in conjunction with the binding of PBX1 to its own promoter and putative enhancer, are consistent with a potential mechanism of *PBX1* transcriptional autoregulation (**Figure 4B**) which would potentiate activity of T417 in chr1q-amp cells. Next, using a subcutaneous xenograft myeloma model, we also validated the anti-myeloma activity of T417 *in vivo*. We observed significantly reduced tumour size and weight in the

T417-treated versus control mice, while in explanted myeloma cells we detected cell cycle arrest and mRNA depletion of the PBX1-regulated genes (**Figure 6E-6G** and **Supplemental Figure S7D-S7H**). In addition, selective cytotoxicity of T417 was detected against *PBX1*-expressing primary chr1q-amplified myeloma cells (X1-X3; n=3), but not against non-amplified MM (X4, X5; n=2) or normal donor peripheral blood B cells (PBBC; n=1) with undetectable *PBX1* mRNA levels (**Figure 6G** and **6H** and **Supplemental Figure S7I** and **S7J**).

Overall, these findings highlight the efficacy of T417 against chr1q-amplified cancer cells and its clinical potential as an adjuvant approach against chr1q-amp high-risk myeloma and other solid tumours.

## Discussion

Recurrent, high frequency CNA such as chr1q-amp are major oncogenic drivers shared across different types of cancer<sup>1-3</sup>. However, delineating the prognostic and functional role of hundreds to thousands of genes and downstream oncogenic pathways associated with specific CNA for development of targeted therapies remains an unmet challenge. In this study, we focused on chr1q-amp, the most frequent CNA linked to high-risk MM<sup>4-6,10</sup>.

First, by combining WGS and 3D genome data we found that genetic amplification disrupts a large proportion of the chromatin structure throughout the chr1q arm. This level of disruption likely reflects contribution of multiple mechanisms to structural changes in chr1q<sup>46</sup>, including isochromosome formation<sup>47</sup>, hypoxia-driven tandem duplications<sup>48</sup>, jumping translocations<sup>4</sup>, chromothripsis<sup>49</sup>, chromoplexy<sup>50</sup>, and combination of the above<sup>51</sup>. Nevertheless, we detected four main blocks of co-amplification (hyper-domains) which are the product of distinct amplification patterns and retain their overall chromatin structure across MM patients. Of those, only two hyper-domains (B1, B4) contribute to adverse prognosis, and therefore have potential implications for the chr1q-amp biology.

In contrast to previous studies which traditionally focused on 1q21 band alone<sup>13-17</sup>, here we employed a large-scale, integrative analysis of clinical and multi-omics datasets to identify adverse prognosis driver genes across the whole chr1q arm. This analysis validated previously reported high-risk markers in 1q21 locus<sup>13-16</sup>, but also linked novel genes to adverse prognosis and highlighted the biological and prognostic significance of two other new areas, namely the 1q22 and 1q23.3 bands. Collectively, the adverse prognosis genes identified across chr1q are predicted to promote cell cycle and proliferation, suggesting their direct involvement in the well-characterized proliferative phenotype associated with chr1q-amp in MM<sup>17,43</sup>.

Identification of *PBX1*, located in 1q23.3, as a prominent candidate driver of high-risk disease in chr1-amp MM, also exemplifies the potential of our approach for biological discovery. Indeed, the role of *PBX1* in promoting cancer cell survival, metastasis and drug resistance has been previously reported in other types of cancer<sup>35-37,52</sup>. In addition, previous in silico work predicted a role of *PBX1* in the biology of t(4;14) myeloma<sup>53</sup> which, interestingly, is highly associated with chr1q-amp<sup>6</sup>. Here we found ectopic expression of *PBX1* in chr1q-amp myeloma cells associated with genetic amplification and strong epigenetic activation of its entire TAD (including proximal and distal DNA elements), suggesting a selective process acting on a whole regulatory domain rather than the gene alone, as previously suggested in oncogenesis<sup>54</sup>. Moreover, our combined genetic, epigenetic and pharmacological approaches establish the mechanisms and regulatory networks through which *PBX1* regulates the activity of *FOXM1*, a master TF promoting cell cycle progression<sup>40,42</sup>. The proliferative circuitries regulated by *PBX1* and the *PBX1*-*FOXM1* axis are of wider importance in cancer, as they exert a powerful prognostic impact in several cancers. Pertinently, chr1-amp is one of the most frequent CNA not only in MM but also other cancers, including breast and ovarian cancer<sup>35-37</sup>.

The finding that pharmacological abrogation of the PBX1-FOXM1 axis selectively impacts survival of chr1q-amp myeloma cells is one of the most notable findings of this work. As well as providing proof-of-principle for developing CNA-specific therapeutic approaches, our data strongly support the central role of PBX1 and FOXM1 in regulation of the transcriptional programme driving the proliferative phenotype and adverse prognosis in chr1q-amp MM. In addition, these findings support our recent efforts for development of T417, a small molecule inhibitor of PBX1 binding to its cognate DNA motif<sup>45</sup>, and suggest the potential benefit of its use in MM and other cancers with chr1-amp and PBX1 over-expression. Indeed, along with our previously reported pre-clinical activity of T417 against ovarian cancer<sup>45</sup>, the data herein demonstrate selective targeting against MM, breast, lung, liver and brain cancer cells with chr1-amp. These findings not only validate the presence of a common, PBX1-FOXM1 axis underlying chr1q-amp that is active in many types of cancer, but also provide the basis for clinical development of T417 as a chr1q-amp-targeting therapy.

In summary, we showed that our systems medicine dissection of CNA in cancer, which includes integration of genetic, epigenetic, transcriptional and 3D-chromatin profiles, is a powerful strategy for discovery of genes and cellular oncogenic pathways of biological significance and clinical impact. Through this process, we show that the ectopically expressed PBX1, in co-operation with FOXM1, is a critical driver of the proliferative phenotype in chr1q-amp MM and several other cancers, and we provide proof-of-principle for selective therapeutic targeting of chr1q-amp, the most prevalent CNA in MM and in cancer in general.

### Acknowledgements

NT, VC, XX and KP were supported by Bloodwise (Blood Cancer UK), AIK was supported by Kay Kendall Leukaemia Fund and Imperial NIHR Biomedical Research Centre. We also acknowledge support from LMS/NIHR Flow-cytometry Facility, Imperial NIHR Biomedical Research Centre Genomics Facility, Imperial NIHR Biomedical Research Centre, Cancer Research UK Experimental Cancer Medicine Centre. The NCU.MM1 myeloma cell line was kindly provided by Dr. Ichiro Hanamura (Aichi Medical University, Japan). The OVCAR3 and A2780 cell lines were kindly provided by Dr. Sadaf Ghaem-Maghami (Imperial College London, UK). We would like to thank Dr. Udayakumar Achandira (Hammersmith hospital) for his kind advice and assistance with the MMCL FISH analysis.

### Author Contributions

NT designed study, conceived and implemented computational pipelines, designed and performed experiments, wrote manuscript. AIK, KP, YS, IK, BB, and KK performed experiments. XX assisted in bioinformatics data analysis. PD and NK performed and interpreted IHC analysis. RMS assisted in clinical informatics data analysis. AC and HWA provided clinical samples. IAGR, TL and LM wrote manuscript. VSC and AnK: designed and supervised the study, generated draft manuscript. All authors contributed to the final draft of the manuscript.

### Author Disclosures

The authors declare no relevant conflict of interest.

### References

1. Watkins TBK, Lim EL, Petkovic M, et al. Pervasive chromosomal instability and karyotype order in tumour evolution. *Nature*. Nov 2020;587(7832):126-132. doi:10.1038/s41586-020-2698-6
2. Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature*. Nov 2019;575(7781):210-216. doi:10.1038/s41586-019-1689-y
3. Beroukhi R, Mermel CH, Porter D, et al. The landscape of somatic copy-number alteration across human cancers. *Nature*. Feb 18 2010;463(7283):899-905. doi:10.1038/nature08822
4. Pawlyn C, Morgan GJ. Evolutionary biology of high-risk multiple myeloma. *Nat Rev Cancer*. Aug 24 2017;17(9):543-556. doi:10.1038/nrc.2017.63

5. Avet-Loiseau H, Li C, Magrangeas F, et al. Prognostic significance of copy-number alterations in multiple myeloma. *J Clin Oncol*. Sep 20 2009;27(27):4585-90. doi:10.1200/JCO.2008.20.6136
6. Shah GL, Landau H, Londono D, et al. Gain of chromosome 1q portends worse prognosis in multiple myeloma despite novel agent-based induction regimens and autologous transplantation. *Leuk Lymphoma*. Aug 2017;58(8):1823-1831. doi:10.1080/10428194.2016.1260126
7. Croft J, Ellis S, Sherborne AL, et al. Copy number evolution and its relationship with patient outcome-an analysis of 178 matched presentation-relapse tumor pairs from the Myeloma XI trial. *Leukemia*. Jul 2021;35(7):2043-2053. doi:10.1038/s41375-020-01096-y
8. Walker BA, Leone PE, Chiecchio L, et al. A compendium of myeloma-associated chromosomal copy number abnormalities and their prognostic value. *Blood*. Oct 14 2010;116(15):e56-65. doi:10.1182/blood-2010-04-279596
9. Schmidt TM, Barwick BG, Joseph N, et al. Gain of Chromosome 1q is associated with early progression in multiple myeloma patients treated with lenalidomide, bortezomib, and dexamethasone. *Blood Cancer J*. Nov 25 2019;9(12):94. doi:10.1038/s41408-019-0254-0
10. Shah V, Sherborne AL, Walker BA, et al. Prediction of outcome in newly diagnosed myeloma: a meta-analysis of the molecular profiles of 1905 trial patients. *Leukemia*. Jan 2018;32(1):102-110. doi:10.1038/leu.2017.179
11. Avet-Loiseau H, Attal M, Campion L, et al. Long-term analysis of the IFM 99 trials for myeloma: cytogenetic abnormalities [t(4;14), del(17p), 1q gains] play a major role in defining long-term survival. *J Clin Oncol*. Jun 1 2012;30(16):1949-52. doi:10.1200/JCO.2011.36.5726
12. Hanamura I, Stewart JP, Huang Y, et al. Frequent gain of chromosome band 1q21 in plasma-cell dyscrasias detected by fluorescence in situ hybridization: incidence increases from MGUS to relapsed myeloma and is related to prognosis and disease progression following tandem stem-cell transplantation. *Blood*. Sep 1 2006;108(5):1724-32. doi:10.1182/blood-2006-03-009910
13. Shaughnessy J. Amplification and overexpression of CKS1B at chromosome band 1q21 is associated with reduced levels of p27Kip1 and an aggressive clinical course in multiple myeloma. *Hematology*. 2005;10 Suppl 1:117-26. doi:10.1080/10245330512331390140
14. Shi L, Wang S, Zangari M, et al. Over-expression of CKS1B activates both MEK/ERK and JAK/STAT3 signaling pathways and promotes myeloma cell drug-resistance. *Oncotarget*. May 2010;1(1):22-33. doi:10.18632/oncotarget.105
15. Marchesini M, Ogoti Y, Fiorini E, et al. ILF2 Is a Regulator of RNA Splicing and DNA Damage Response in 1q21-Amplified Multiple Myeloma. *Cancer Cell*. Jul 10 2017;32(1):88-100 e6. doi:10.1016/j.ccell.2017.05.011
16. Teoh PJ, Chung TH, Chng PYZ, Toh SHM, Chng WJ. IL6R-STAT3-ADAR1 (P150) interplay promotes oncogenicity in multiple myeloma with 1q21 amplification. *Haematologica*. May 2020;105(5):1391-1404. doi:10.3324/haematol.2019.221176
17. Shaughnessy JD, Jr., Zhan F, Burington BE, et al. A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood*. Mar 15 2007;109(6):2276-84. doi:10.1182/blood-2006-07-038430
18. Mani M, Carrasco DE, Zhang Y, et al. BCL9 promotes tumor progression by conferring enhanced proliferative, metastatic, and angiogenic properties to cancer cells. *Cancer Res*. Oct 1 2009;69(19):7577-86. doi:10.1158/0008-5472.CAN-09-0773
19. Zhang B, Gojo I, Fenton RG. Myeloid cell factor-1 is a critical survival factor for multiple myeloma. *Blood*. Mar 15 2002;99(6):1885-93. doi:10.1182/blood.v99.6.1885
20. Teoh G, Urashima M, Ogata A, et al. MDM2 protein overexpression promotes proliferation and survival of multiple myeloma cells. *Blood*. Sep 1 1997;90(5):1982-92.
21. Zang M, Zou D, Yu Z, et al. Detection of recurrent cytogenetic aberrations in multiple myeloma: a comparison between MLPA and iFISH. *Oncotarget*. Oct 27 2015;6(33):34276-87. doi:10.18632/oncotarget.5371
22. Fu L, Cheng Z, Dong F, et al. Enhanced expression of FCER1G predicts positive prognosis in multiple myeloma. *J Cancer*. 2020;11(5):1182-1194. doi:10.7150/jca.37313
23. Spanoudakis E, Hu M, Naresh K, et al. Regulation of multiple myeloma survival and progression by CD1d. *Blood*. Mar 12 2009;113(11):2498-507. doi:10.1182/blood-2008-06-161281
24. Kassambara A, Hose D, Moreaux J, et al. Genes with a spike expression are clustered in chromosome (sub)bands and spike (sub)bands have a powerful prognostic value in patients with multiple myeloma. *Haematologica*. Apr 2012;97(4):622-30. doi:10.3324/haematol.2011.046821

25. Alvarez-Benayas J, Trasanidis N, Katsarou A, et al. Chromatin-based, in cis and in trans regulatory rewiring underpins distinct oncogenic transcriptomes in multiple myeloma. *Nat Commun.* Sep 14 2021;12(1):5450. doi:10.1038/s41467-021-25704-2
26. Iskander D, Wang G, Heuston EF, et al. Single-cell profiling of human bone marrow progenitors reveals mechanisms of failing erythropoiesis in Diamond-Blackfan anemia. *Sci Transl Med.* Sep 8 2021;13(610):eabf0113. doi:10.1126/scitranslmed.abf0113
27. Ponnusamy K, Tzioni MM, Begum M, et al. The innate sensor ZBP1-IRF3 axis regulates cell proliferation in multiple myeloma. *Haematologica.* Feb 18 2021;doi:10.3324/haematol.2020.274480
28. Caputo VS, Trasanidis N, Xiao X, et al. Brd2/4 and Myc regulate alternative cell lineage programmes during early osteoclast differentiation in vitro. *iScience.* Jan 22 2021;24(1):101989. doi:10.1016/j.isci.2020.101989
29. MMRF. Multiple Myeloma Research Foundation Researcher Gateway: Explore IA15, <https://research.themmr.org/rp/download>. 2020;
30. Wu P, Li T, Li R, et al. 3D genome of multiple myeloma reveals spatial genome disorganization associated with copy number variations. *Nat Commun.* Dec 5 2017;8(1):1937. doi:10.1038/s41467-017-01793-w
31. Rao SS, Huntley MH, Durand NC, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell.* Dec 18 2014;159(7):1665-80. doi:10.1016/j.cell.2014.11.021
32. Sawyer JR, Tian E, Thomas E, et al. Evidence for a novel mechanism for gene amplification in multiple myeloma: 1q12 pericentromeric heterochromatin mediates breakage-fusion-bridge cycles of a 1q12 approximately 23 amplicon. *Br J Haematol.* Nov 2009;147(4):484-94. doi:10.1111/j.1365-2141.2009.07869.x
33. Zhan F, Huang Y, Colla S, et al. The molecular classification of multiple myeloma. *Blood.* Sep 15 2006;108(6):2020-8. doi:10.1182/blood-2005-11-013458
34. Jin Y, Chen K, De Paepe A, et al. Active enhancer and chromatin accessibility landscapes chart the regulatory network of primary multiple myeloma. *Blood.* May 10 2018;131(19):2138-2150. doi:10.1182/blood-2017-09-808063
35. Magnani L, Ballantyne EB, Zhang X, Lupien M. PBX1 genomic pioneer function drives ERalpha signaling underlying progression in breast cancer. *PLoS Genet.* Nov 2011;7(11):e1002368. doi:10.1371/journal.pgen.1002368
36. Magnani L, Patten DK, Nguyen VT, et al. The pioneer factor PBX1 is a novel driver of metastatic progression in ERalpha-positive breast cancer. *Oncotarget.* Sep 8 2015;6(26):21878-91. doi:10.18632/oncotarget.4243
37. Wang J, Shidfar A, Ivancic D, et al. Overexpression of lipid metabolism genes and PBX1 in the contralateral breasts of women with estrogen receptor-negative breast cancer. *Int J Cancer.* Jun 1 2017;140(11):2484-2497. doi:10.1002/ijc.30680
38. Sarin V, Yu K, Ferguson ID, et al. Evaluating the efficacy of multiple myeloma cell lines as models for patient tumors via transcriptomic correlation analysis. *Leukemia.* Oct 2020;34(10):2754-2765. doi:10.1038/s41375-020-0785-1
39. Mahdipour-Shirayeh A, Erdmann N, Leung-Hagesteijn C, Tiedemann RE. sciCNV: High-throughput paired profiling of transcriptomes and DNA copy number variations at single cell resolution. *bioRxiv.* 2020:2020.02.10.942607. doi:10.1101/2020.02.10.942607
40. Gu C, Jing X, Holman C, et al. Upregulation of FOXM1 leads to diminished drug sensitivity in myeloma. *BMC Cancer.* Nov 21 2018;18(1):1152. doi:10.1186/s12885-018-5015-0
41. Zhou W, Yang Y, Xia J, et al. NEK2 induces drug resistance mainly through activation of efflux drug pumps and is associated with poor prognosis in myeloma and other cancers. *Cancer Cell.* Jan 14 2013;23(1):48-62. doi:10.1016/j.ccr.2012.12.001
42. Hegde NS, Sanders DA, Rodriguez R, Balasubramanian S. The transcription factor FOXM1 is a cellular target of the natural product thiostrepton. *Nat Chem.* Aug 21 2011;3(9):725-31. doi:10.1038/nchem.1114
43. Hose D, Reme T, Hielscher T, et al. Proliferation is a central independent prognostic factor and target for personalized and risk-adapted treatment in multiple myeloma. *Haematologica.* Jan 2011;96(1):87-95. doi:10.3324/haematol.2010.030296
44. Bergsagel PL, Kuehl WM, Zhan F, Sawyer J, Barlogie B, Shaughnessy J, Jr. Cyclin D dysregulation: an early and unifying pathogenic event in multiple myeloma. *Blood.* Jul 1 2005;106(1):296-303. doi:10.1182/blood-2005-01-0034
45. Shen Y-A, Jung J, Shimberg GD, et al. Development of Small Molecule Inhibitors Targeting PBX1 Transcription Signaling as a Novel Cancer Therapeutic Strategy. *bioRxiv.* 2021:2021.04.19.440460. doi:10.1101/2021.04.19.440460

46. Li Y, Roberts ND, Wala JA, et al. Patterns of somatic structural variation in human cancer genomes. *Nature*. Feb 2020;578(7793):112-121. doi:10.1038/s41586-019-1913-9
47. Sawyer JR, Tricot G, Lukacs JL, et al. Genomic instability in multiple myeloma: evidence for jumping segmental duplications of chromosome arm 1q. *Genes Chromosomes Cancer*. Jan 2005;42(1):95-106. doi:10.1002/gcc.20109
48. Black JC, Atabakhsh E, Kim J, et al. Hypoxia drives transient site-specific copy gain and drug-resistant gene expression. *Genes Dev*. May 15 2015;29(10):1018-31. doi:10.1101/gad.259796.115
49. Maura F, Boyle EM, Rustad EH, et al. Chromothripsis as a pathogenic driver of multiple myeloma. *Semin Cell Dev Biol*. May 3 2021;doi:10.1016/j.semcdb.2021.04.014
50. Maclachlan KH, Rustad EH, Derkach A, et al. Copy number signatures predict chromothripsis and clinical outcomes in newly diagnosed multiple myeloma. *Nat Commun*. Aug 27 2021;12(1):5172. doi:10.1038/s41467-021-25469-8
51. Hanamura I. Gain/Amplification of Chromosome Arm 1q21 in Multiple Myeloma. *Cancers (Basel)*. Jan 12 2021;13(2)doi:10.3390/cancers13020256
52. Liu Y, Xu X, Lin P, et al. Inhibition of the deubiquitinase USP9x induces pre-B cell homeobox 1 (PBX1) degradation and thereby stimulates prostate cancer cell apoptosis. *J Biol Chem*. Mar 22 2019;294(12):4572-4582. doi:10.1074/jbc.RA118.006057
53. Calura E, Bisognin A, Manzoni M, et al. Disentangling the microRNA regulatory milieu in multiple myeloma: integrative genomics analysis outlines mixed miRNA-TF circuits and pathway-derived networks modulated in t(4;14) patients. *Oncotarget*. Jan 19 2016;7(3):2367-78. doi:10.18632/oncotarget.6151
54. Morton AR, Dogan-Artun N, Faber ZJ, et al. Functional Enhancers Shape Extrachromosomal Oncogene Amplifications. *Cell*. Nov 27 2019;179(6):1330-1341 e13. doi:10.1016/j.cell.2019.10.039

## Figure legends

### Figure 1. Multi-layer, systems medicine analysis of chr1q amplification in multiple myeloma.

**(A)** Two-dimensional co-amplification (cyan) and three-dimensional Hi-C contact (red) maps of chr1q locus in MM cells used to identify topologically co-amplified domains (TCDs) and topologically associated domains (TADs), respectively. Map overlay identified four major co-amplified domains that retain a preserved 3D structure (B1-4 hyper-domains).

**(B)** Schematic overview of the analysis strategy used to detect candidate gene drivers of high-risk phenotypes in chr1q-amp MM. Scanning across the 2,215 genes in chr1q, those fulfilling all the following criteria were considered as candidate drivers: (1) their genetic amplification is significantly associated with poor prognosis (MMRF dataset, n=896); (2) their genetic amplification is significantly associated with overexpression (MMRF dataset, n=896); (3) overexpression is associated with poor prognosis in the MMRF (top panel, n=896) and Arkansas datasets (bottom panel, n=413); (4) significant epigenetic activation (i.e., H3K27ac gain) is detected in chr1q-amplified versus non-amplified samples (Jin et al., n=10). The B1-B4 hyper-domains were also used as a reference here (5). Overall, our analysis identified 103 genes across chr1q arm as candidate drivers of high-risk MM prognosis.

**(C)** Analysis overview, from top to bottom: chr1q cytogenetic map; copy-number profiles of chr1q genes across MMRF patients detecting whole-arm amplification (~29%), partial amplification (~7%), no amplification (~63%) and deletions (~1%); survival analysis of genetic amplification of chr1q genes across MMRF patients (WGS, 73 genetic parameters; dark green bars, P-value; light green bars, Hazard Ratio; grey bars, % bootstrapping confidence levels); Pearson correlation analysis between copy-number ratios (WGS) and expression (RNA-seq; blue bars indicate Pearson correlation P-values); survival analysis of chr1q gene expression (RNA-seq) in MMRF (brown) and Arkansas (yellow) datasets (bars indicate analysis P-values); differential H3K27ac analysis between chr1q-amplified (n=5) versus non-amplified (n=5) MM cells (red bars indicate differential log<sub>2</sub> fold-change enrichment scores); four chr1q domains (B1-4) with conserved TAD/TCD structures; Candidate pathogenic driver genes (n=103, pink bars) identified by the current analysis (the previously known *MCL1*, *ARNT*, *ILF2* and *CKS1B* genes are shown here).

(D) Analysis overview of candidate driver genes (103) across chr1q bands. Distribution of WGS multivariate analysis scores ( $-\log_{10}P$ -value; top) and percentage (%) of candidate genes (relative to band gene density) per cytogenetic band. The highest candidate genes density was detected in 1q22 and 1q23.3 bands (highlighted here), with 1q23.3 also displaying the highest survival significance scores.

(E) The *PBX1* gene as a prominent candidate occupying alone a single TAD, displays strong epigenetic activation across *PBX1* body and putative enhancers in chr1amp myeloma PC.

### Figure 2. *PBX1*-dependent myeloma cell proliferation.

(A) mRNA expression of *PBX1* in four MM cell lines.

(B) Immunohistochemical (IHC) analysis of trephine bone marrow samples from 11 MM patients detects no (neg), medium (1) or high (2) *PBX1* expression at clonal or subclonal level (% of *PBX1*+ cells).

(C) Time-course, flow-cytometry based analysis of MM.1S (left) and U226 (right) myeloma cell viability *in vitro*, upon lentiviral transduction with scrambled control (scrbl) and anti-*PBX1* shRNAs (P11, P31). Data collected from three biological replicates represent the fraction of GFP+ live cells on the timepoints displayed, after normalization against Day 3. Statistical analysis was performed using a two-way ANOVA with post-hoc multiple comparisons test. Error bars represent SEM (n=3).

(D & E) Knock-down of *PBX1* in MM.1S cells using an *in vivo* plasmacytoma xenograft mouse model; tumour size photograph (D) and tumour weight (E) measured at termination date (Day 32). Statistical analysis performed using Kruskal-Wallis with Dunn's post-hoc multiple comparisons test.

(F) Relative fraction of transduced cells detected at start (Day 0, Live/GFP+ cells) and termination (Day 32, Live/HLA+GFP+ cells) dates.

(G) RNA-seq analysis of *PBX1*-depleted MM.1S and U266 cells 3 days after lentiviral transduction. Heatmaps indicate differentially expressed genes shared between P11- and P31-depleted cells for each cell line.

(H) Gene Set Enrichment Analysis (GSEA) of up- (top) or down-regulated (bottom) genes in MM.1S (left) and U266 (right) myeloma cells illustrating significantly enriched molecular pathways in each cell line. Enrichment plots for the prominent cell cycle regulation pathway (E2F targets), which was identified as a top hit, are also presented here.

(I) Flow-cytometric cell-cycle analysis of MM.1S and U266 cells 6 days after *PBX1* knockdown. Data present the summary of 3 biological experiments. Analysis was done using parametric one-way ANOVA with post-hoc multiple comparisons test. \*:  $P < 0.05$ ; \*\*:  $P < 0.01$ ; \*\*\*:  $P < 0.001$ ; \*\*\*\*:  $P < 0.0001$

### Figure 3. Genome-wide analysis of *PBX1* function in chr1q-amplified myeloma cells.

(A) Heatmap representation of *PBX1* cistrome in MM.1S and U266 cells, as identified by ChIP-seq analysis (n=2 per cell line). Genomic annotation (left) and epigenomic chromHMM states (right) of significantly enriched regions are also presented here.

(B) Super-enhancer (SEs) analysis across 9 MM cell lines and 8 MM primary samples using H3K27ac ChIP-seq (data obtained from Jin *et al.*, 2018). Number of total (dark red) and *PBX1*-bound (red) SEs (red) across 17 MM samples and aggregated profile in all samples (right) is shown.

(C) Boxplot representations of average normalized H3K27ac signal of chr1q-amplified and non-amplified samples across 1,655 *PBX1*-bound SEs. Analysis was performed using a paired t-test.

(D) Pathway analysis of genes predicted to be regulated by *PBX1*-bound SEs in chr1q-amplified (+) and -non-amplified (-) cells.

(E) Integrative cistrome-transcriptome analysis with BETA-plus displays the regulatory programme of *PBX1* in MM.1S cells. Biological annotation of genes was performed using the Molecular Signatures Database. Node colors represent average predicted activation (blue) or repression (red) for each gene. Transcriptional targets of interest are highlighted in red font.

(F) Overrepresentation analysis against the ChEA database and NCI-Nature pathways of the direct PBX1 target genes in MM.1S (top) and U266 (bottom) cells. Terms of interest are highlighted in red font.

**Figure 4. PBX1 regulates directly FOXM1- and E2F1/2-associated transcriptional programmes in chr1q-amplified MM cells.**

(A) Regulatory connections between PBX1 and its downstream targets FOXM1, E2F1/2 and NEK2 in chr1q-amplified MM cells as emerged from data shown in b-i.

(B) IGV snapshots display the epigenomic features of prominent genetic loci: *PBX1* promoter and enhancer, *E2F1* promoter, *E2F2* promoter and enhancer, *FOXM1* enhancer, *NEK2* promoter and enhancer. From top to bottom: PBX1 ChIP-seq in MM.1S and U266 cells; ChromHMM maps in MM.1S and U266 cells (colour code same as Fig 3A); Super-enhancers as identified in chr1q-amplified MMCL and primary samples.

(C) Flow cytometry-based analysis of MM.1S cell survival (n=3) upon transduction with anti-*FOXM1* shRNAs (O1, O4) and scrambled control (scrbl) lentiviral vectors. Statistical analysis was performed by a two-way ANOVA with post-hoc multiple comparisons test.

(D) Analysis of PBX1, *FOXM1* and *NEK2* expression levels by RT-qPCR after lentiviral transduction with anti-*FOXM1* and scrambled control shRNA in MM.1S cells (n=3). Statistical analysis was performed using a one-way ANOVA with post-hoc multiple comparisons test.

(E) Heatmap representation of differentially expressed genes after *FOXM1* depletion with O1 and O4 shRNAs in comparison to scrambled control (RNA-seq, n=2).

(F) Over-representation analysis of significantly upregulated (top) and downregulated (bottom) genes upon *FOXM1* knockdown in MM.1S cells.

(G) Intracellular staining followed by flow-cytometric analysis of MM.1S (top) and NCU.MM1 (bottom) cells transduced with control (MIGR-EV) or PBX1-overexpressing (MIGR-PBX1) vectors using anti-PBX1 or isotype control antibodies (mean fluorescence intensity ratio between antibodies is shown).

(H) RT-qPCR analysis of *NEK2*, *E2F2* and *FOXM1* mRNA expression in PBX1-overexpressing versus control MM.1S (top) and NCU.MM1 (bottom) cells (n=4). Data were analysed using a one-way ANOVA with post-hoc multiple comparisons test.

(I) Drug sensitivity assays in MIGR-EV and MIGR-PBX1 transduced MM.1S (top) and NCU.MM1 (bottom) cells 48h after treatment with the FOXM1 inhibitor, thiostrepton (n=3). IC<sub>50</sub> values were calculated for each cell line using a non-linear fitting model (fitting line represented here).

Error bars show standard errors of the mean \*:  $P < 0.05$ ; \*\*:  $P < 0.01$ ; \*\*\*:  $P < 0.001$ ; \*\*\*\*:  $P < 0.0001$ ; n/s: not significant.

**Figure 5. Differential regulome and thiostrepton cytotoxicity profiling of primary chr1q-amplified versus non-amplified MM cells.**

(A) Schematic representation of experimental strategy. Myeloma plasma cells were isolated via magnetic beads selection (CD138+) from bone marrow aspirates derived from 6 chr1q-amplified (chr1q-amp(+)) and 6 non-amplified (chr1q-amp(-)) MM patients. Differential regulome (TF expression and wiring) analysis was performed via parallel chromatin accessibility (ATAC-seq) and transcriptome (RNA-seq) profiling.

(B) Volcano plot displaying differentially expressed genes [chr1q-amp(+), green; chr1q-amp(-), orange]. Genes implicated in chr1q-amp pathogenesis in this study (pink) or previous studies (black) are indicated here.

(C) Enrichment analysis (NCI-Nature pathways) of differentially expressed genes in two patient subgroups.

(D) Differential ATAC-seq analysis between chr1q-amp(+) and chr1q-amp(-) myeloma plasma cells. Increased accessibility was found on genetic loci of genes of interest upon chr1q amplification (as indicated here).

- (E) Differential ATAC-seq footprinting analysis of expressed TFs in chr1q-amp(+) versus chr1q-amp(-) cells ( $\Delta P$ : differential regulatory potential). TFs of interest are indicated here.
- (F) Scatter plot representation of differential expression (x-axis) and differential regulatory potential (y-axis) of 63 TFs displaying significant differences in both dimensions. Green quartile: TFs with increased expression and  $\Delta P$  in chr1q-amp(+) cells; orange quartile: TFs with decreased expression and  $\Delta P$  in chr1q-amp(+) cells. Key transcription factors are also highlighted here.
- (G) Selective sensitivity of chr1q-amp(+) (n=3, green) versus chr1q-amp(-) (n=3, orange) primary myeloma plasma cells to thiothrepton at 48h after treatment.  $IC_{50}$  values were calculated for each patient sample using a non-linear fitting model (fitting line represented here). \*\*\*\*,  $P < 0.0001$ .
- (H) Transcriptional profiling (RT-qPCR) of *FOXM1* and *NEK2* mRNA levels in chr1q-amp(+) (green) and chr1q-amp(-) (orange) primary samples 24h upon thiothrepton (1 $\mu$ M) or mock (0nM) treatment. The (%) decrease in *FOXM1* and *NEK2* mRNA levels is also indicated here.

**Figure 6. Selective targeting of chr1q-amplified tumour cells with a selective PBX1 inhibitor.**

- (A) Survival analysis of multiple myeloma, breast, ovarian lung and brain cancer patient cohorts based on the PBX1 signature expression (red: high, black: low; n=320 genes, **Supplemental Table S5**). Kaplan–Meier plots and statistical analysis depict the significantly poorer survival of patients with active PBX1 signature
- (B) Cytotoxicity profiles (n=3) of multiple myeloma (MM.1S, OPM2, U266, NCU.MM1), breast (MCF-7, LTED), ovarian (OVCAR-3, A2780), lung (A549, H69AR) and brain (SNB-75) cancer cell lines 48h after treatment with the small-molecule PBX1 inhibitor T417. Three independent experiments were performed per cell line and  $IC_{50}$  values were calculated using a non-linear fitting model (fitting line represented here).
- (C) Cell cycle profiling of 11 cancer cell lines 48h after treatment with 1% DMSO (control) or T417 (20 $\mu$ M). Three independent experiments were performed per cell line. Asterisks indicate statistical comparisons performed using a two-way non-parametric ANOVA with post-hoc multiple comparisons test.
- (D) Assessment of *PBX1*, *FOXM1*, *NEK2* and *E2F2* mRNA levels in 11 cancer cell lines 16-20h after treatment with 1% DMSO (control) or T417 (20 $\mu$ M). Bar graphs illustrate transcriptional levels normalized to corresponding control samples (n=3 replicates). Analysis performed using paired t-test.
- (E) Tumour volumes (mm<sup>3</sup>) of the MM.1S xenografts measured in vehicle- (control) and T417-treated (10 mg/kg/injection) mice across experimental timepoints. Statistical analysis was performed using a two-way ANOVA with post-hoc multiple comparisons test.
- (F) Photograph of tumour sizes explanted at termination date (Day 23) from control- and T417-treated mice.
- (G) Heatmap representation of *PBX1*, *FOXM1*, *NEK2* and *E2F2* mRNA levels assessed by RT-qPCR. For the *in vivo* experiment in tumour explanted cells, values represent pairwise comparisons of T417 group (D1-D5) against vehicle-treated group (C1-C5). For *in vitro* primary myeloma plasma cell samples, values represent T417-treated (20 $\mu$ M) versus control-treated (1% DMSO) cells. Grey values correspond to non-applicable (NA) comparisons due to undetectable mRNA levels in control-treated cells.
- (H) Cell viability of primary chr1q-amplified MM (X1, X2, X3; green), non-amplified MM (X4, X5; orange) and normal donor peripheral blood B cells (PBBC; orange) at 48h after treatment with 1% DMSO (control) or T417 (20 $\mu$ M). Non-linear data fitting and  $IC_{50}$  calculations were performed as described in (B).

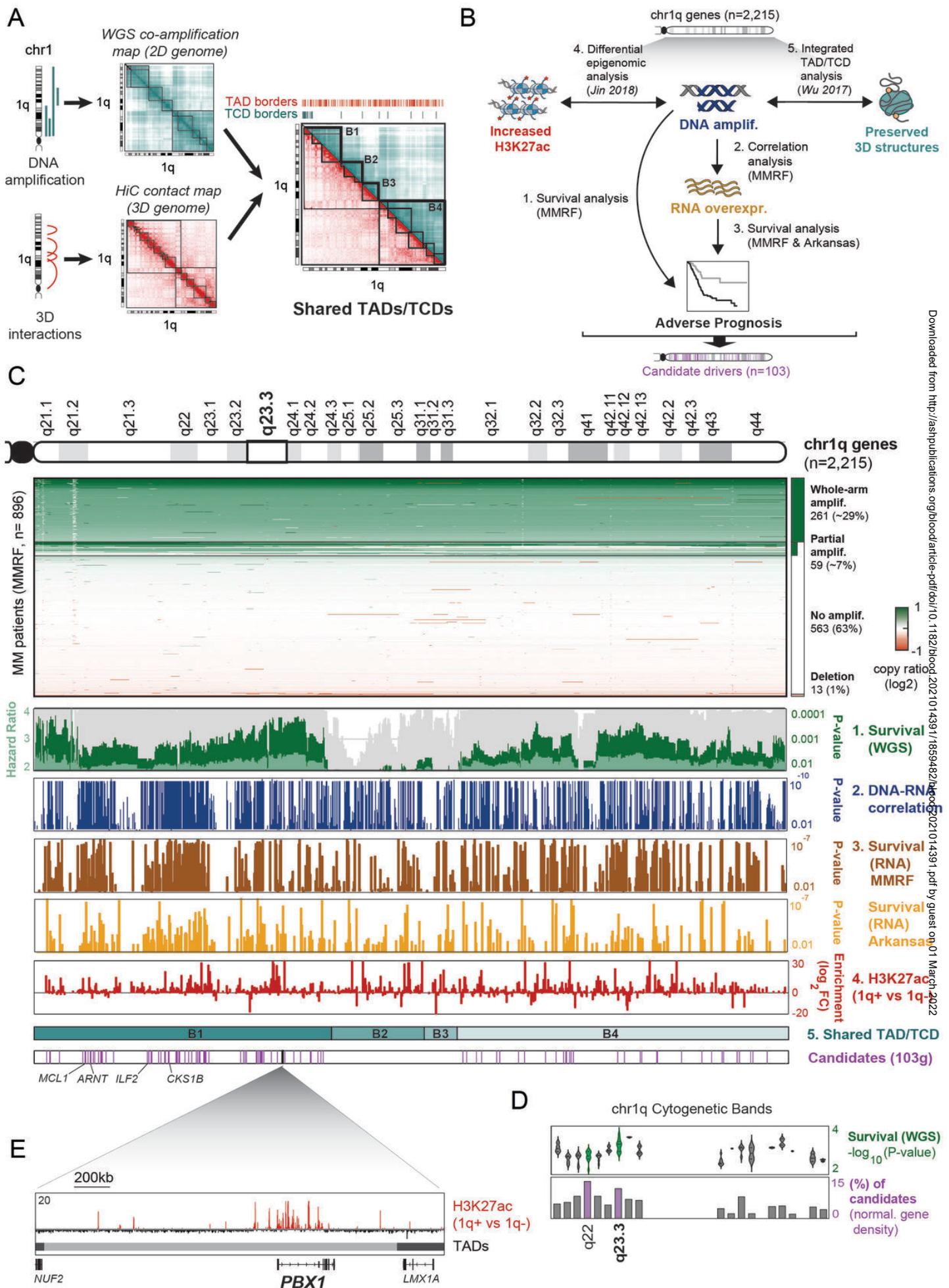


Figure 1

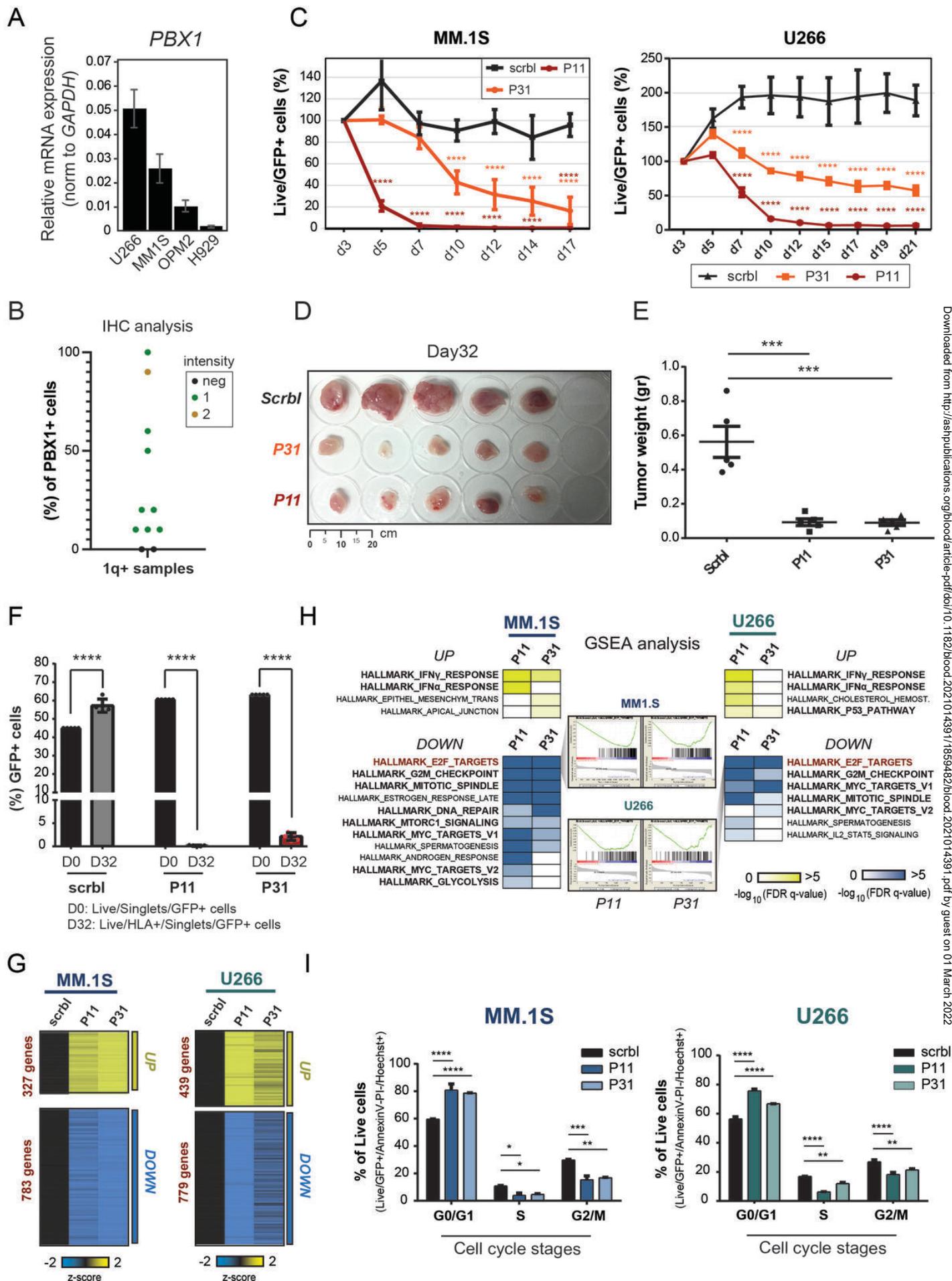


Figure 2

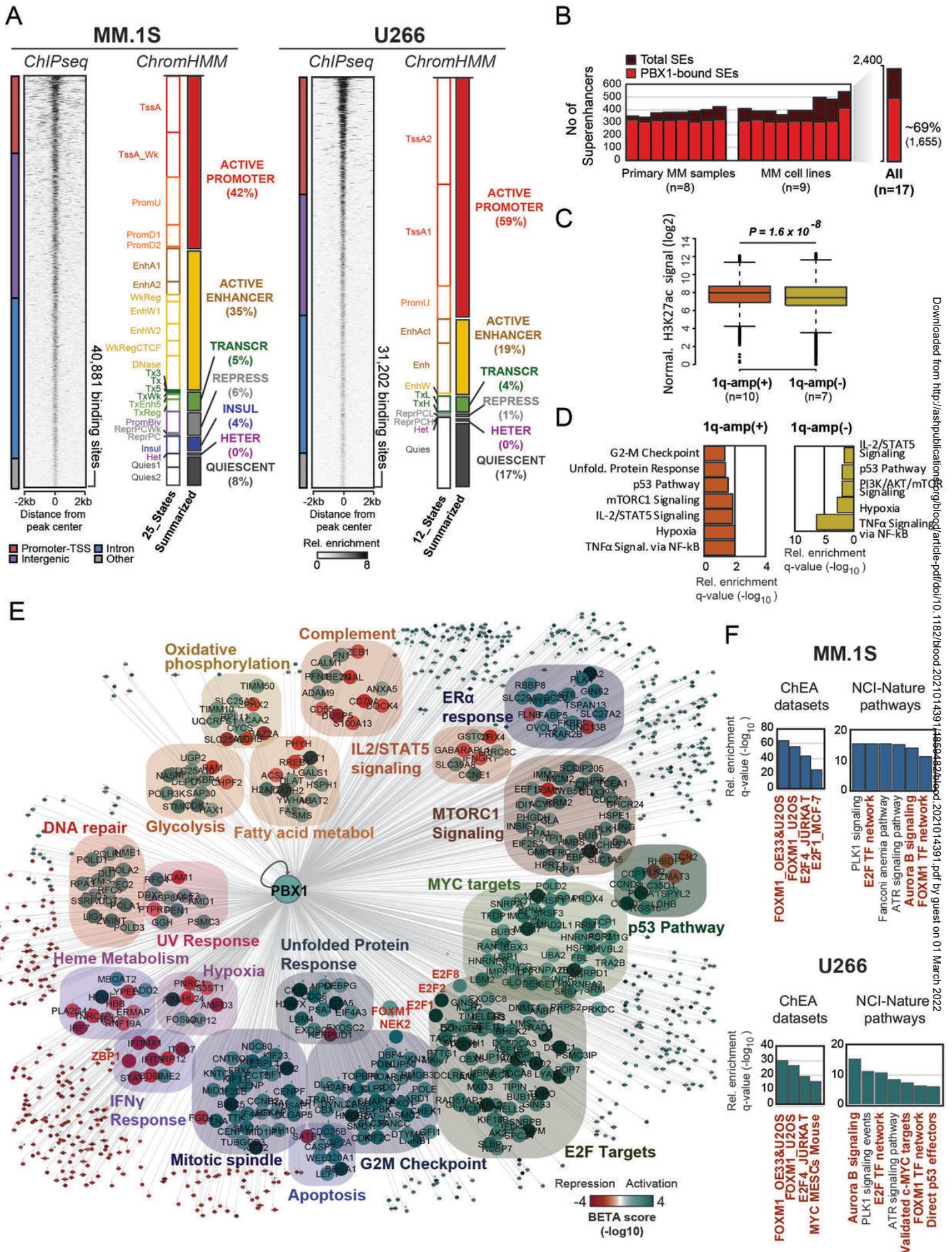
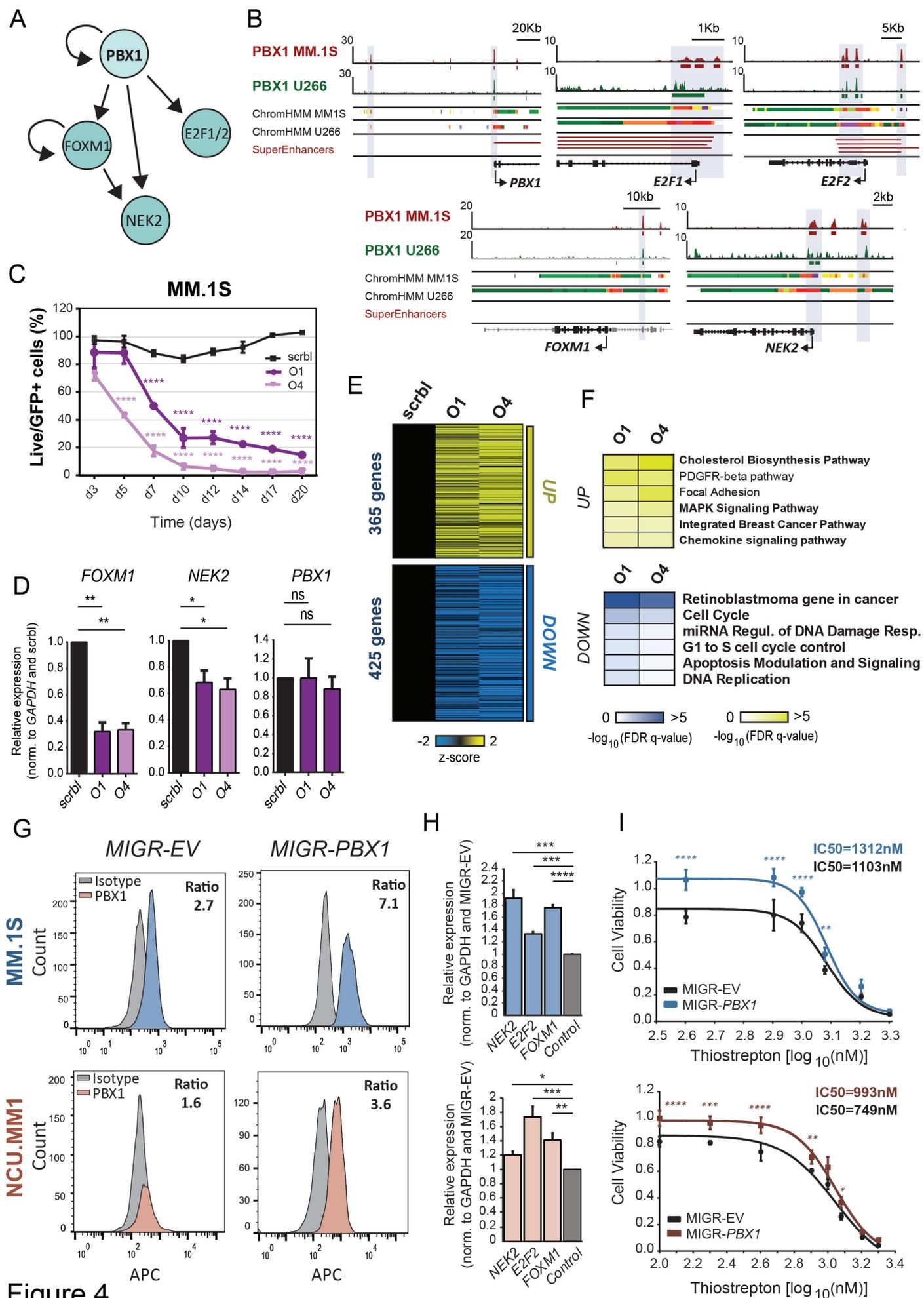


Figure 3



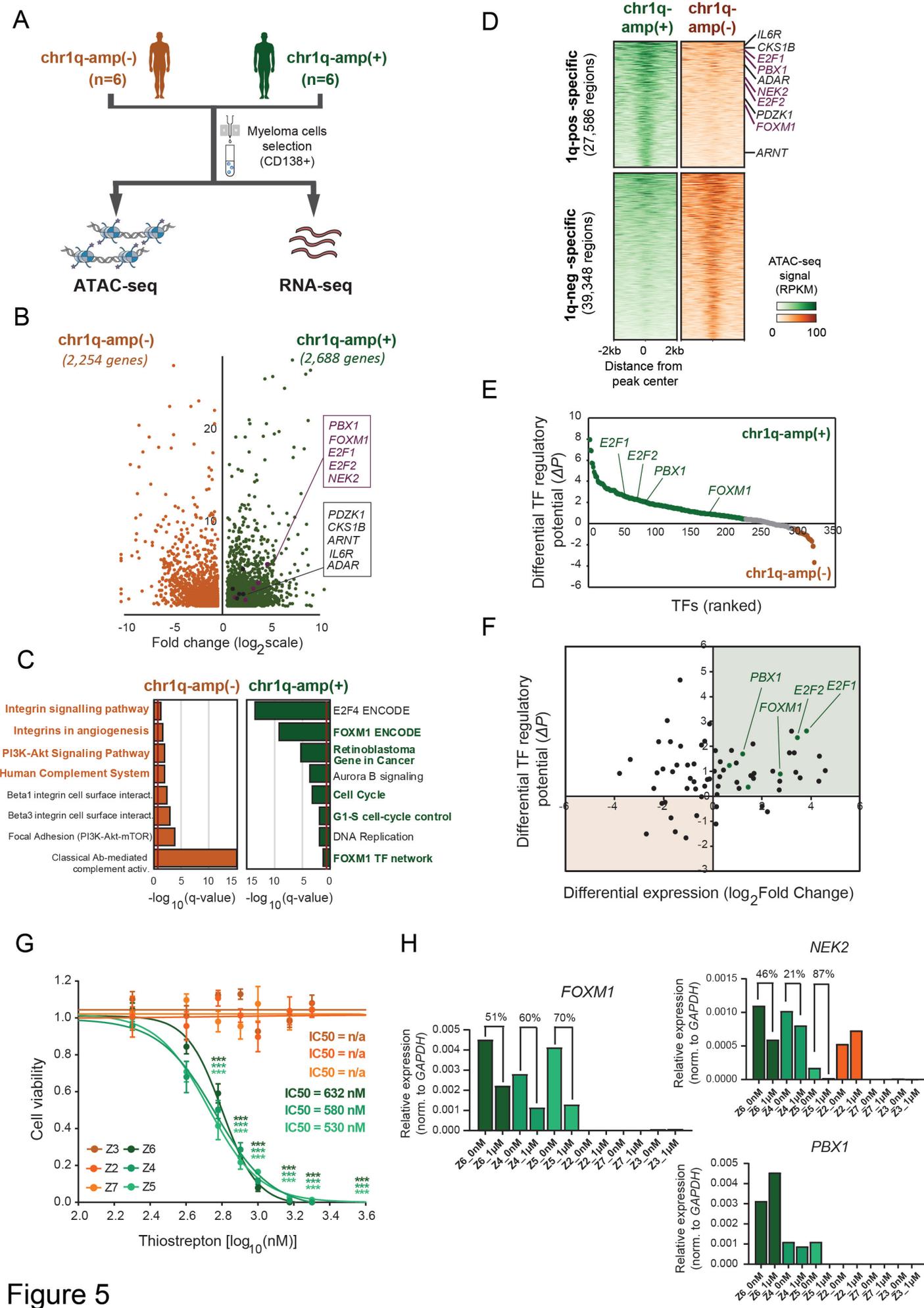


Figure 5

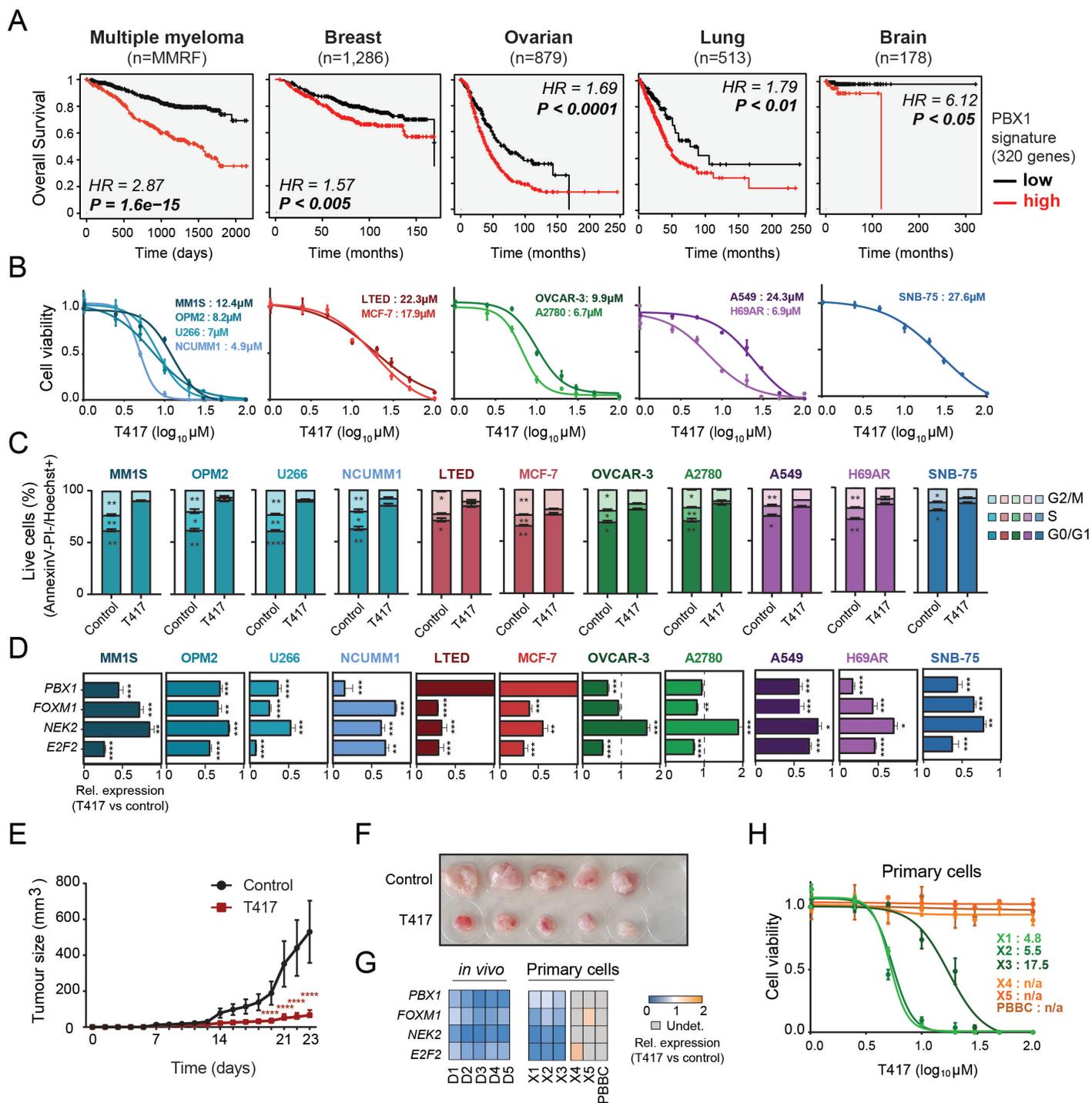


Figure 6