GENETICS

FOXA1 defines cancer cell specificity

2016 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).
10.1126/sciadv.1501473

Gaihua Zhang,* Yongbing Zhao,* Yi Liu,* Li-Pin Kao, Xiao Wang, Benjamin Skerry, Zhaoyu Li†

A transcription factor functions differentially and/or identically in multiple cell types. However, the mechanism for cell-specific regulation of a transcription factor remains to be elucidated. We address how a single transcription factor, forkhead box protein A1 (FOXA1), forms cell-specific genomic signatures and differentially regulates gene expression in four human cancer cell lines (HepG2, LNCaP, MCF7, and T47D). FOXA1 is a pioneer transcription factor in organogenesis and cancer progression. Genomewide mapping of FOXA1 by chromatin immunoprecipitation sequencing annotates that target genes associated with FOXA1 binding are mostly common to these cancer cells. However, most of the functional FOXA1 target genes are specific to each cancer cell type. Further investigations using CRISPR-Cas9 genome editing technology indicate that cell-specific FOXA1 regulation is attributable to unique FOXA1 binding, genetic variations, and/or potential epigenetic regulation. Thus, FOXA1 controls the specificity of cancer cell types. We raise a "flower-blooming" hypothesis for cell-specific transcriptional regulation based on these observations.

INTRODUCTION

Tissue-specific or cell-specific transcriptional regulation has been found in almost all species (*1*). Organogenesis and neoplastic transformation are tightly controlled by tissue-specific/cell-specific transcriptional regulation. Studies of cell-specific transcriptional regulation have focused on global patterns of transcriptional machinery in different tissues and cell types at certain developmental stages or during cancer initiation and progression (*2*). Whether a single transcription factor has common and/or differential regulation—or how a transcription factor regulates differentially—in multiple types of cells is a fundamental but less understood question in genetics. Forkhead box protein A1 (FOXA1) is a pioneer transcription factor that regulates organogenesis and cancer progression in the liver, breast, prostate, lung, and endometrium (*3*–*10*). Here, we address how a single transcription factor, FOXA1, forms cancer-specific genomic signatures and differentially regulates gene expression among four human cancer cell lines, including liver cancer cells (HepG2), prostate cancer cells (LNCaP), and breast cancer cells (MCF7 and T47D).

RESULTS

Functional and cell-specific FOXA1 targeting in human cancer cell lines

High-throughput sequencing technology coupled with chromatin immunoprecipitation (ChIP-Seq) allows us to decipher genomewide scenarios of common and cell-specific regulation of transcription factors. Thus, we applied a series of genomic approaches to investigate global FOXA1 regulation in these cancer cells, including ChIP-Seq (fig. S1A). We collected 20 sets of FOXA1 ChIP-Seq data on HepG2, LNCaP, MCF7, and T47D cells from the ENCODE and GEO databases (see Supplementary Materials and Methods) and reanalyzed these ChIP-Seq data using the algorithms Bowtie2 and HOMER for reads alignment, peak calling, and motif filtering (*11*, *12*). We found 67,753, 88,517, 70,010, and 71,021 total FOXA1 binding peaks in HepG2, LNCaP, MCF7, and T47D cells, respectively (fig. S1B). Not surprisingly, MCF7 and T47D cells had the most uniquely common peaks (10,259) because both of them are estrogen receptor α–positive breast cancer cell lines. About half (47 to 54%) of the FOXA1 binding peaks were unique to each cell line, given that the common sites between MCF7 and T47D were considered to be unique to estrogen receptor α–positive breast cancer cells compared to liver and prostate cancer cells (Fig. 1A and fig. S1B). However, we found that most (68 to 75%) of the FOXA1 target genes associated with these FOXA1 binding peaks were common to the four cancer cell lines (Fig. 1B), suggesting that most of these unique FOXA1 binding peaks target common genes.

Further analysis showed that each FOXA1 target gene was associated with multiple FOXA1 binding peaks (from 1 to 99 peaks), more than 90% of the genes had fewer than 10 FOXA1 binding peaks, and about 30 to 40% of the target genes were only associated with one FOXA1 binding peak in all four cancer cell lines (fig. S1C). In addition, we identified five types of gene-peak pairs of FOXA1 targeting among the four cancer cell lines, including unique, common, and mixed peaks (fig. S2): (i) only one unique FOXA1 binding peak targeting a single gene in one of the four cell lines (Unique, $n = 1$); (ii) multiple unique FOXA1 binding peaks targeting a single gene in one of the four cell lines (Unique, $n > 1$); (iii) a single common FOXA1 binding peak targeting a single gene in two to four cell lines (Common, $n = 1$); (iv) multiple common FOXA1 binding peaks targeting a single gene in two to four cell lines (Common, $n > 1$); and (v) mixed unique and common FOXA1 binding peaks targeting a single gene among the four cell lines (Mixed). The majority of FOXA1 targeting among the four cancer cell lines was regulated by mixed FOXA1 binding of both unique and common peaks (fig. S1D). These data suggest that the majority of cell-specific FOXA1 regulation results from differential FOXA1 binding at the regulatory region of the same target gene among the four cell lines. Why about 90% of the human genes were bound by a single factor, FOXA1, in the four human cancer cell lines cannot be concluded yet from ChIP-Seq data. Thus, the identification of functional binding and targeting from multiple binding peaks is critical for the elucidation of functional FOXA1 regulation.

Department of Cancer Biology, Mayo Clinic, 4500 San Pablo Road, Jacksonville, FL 32224, USA.
*These authors contributed equally to this work.
†Corresponding author. E-mail: li.zhaoyu@mayo.edu

Downloaded from http://advances.sciencemag.org/ on January 23, 2019

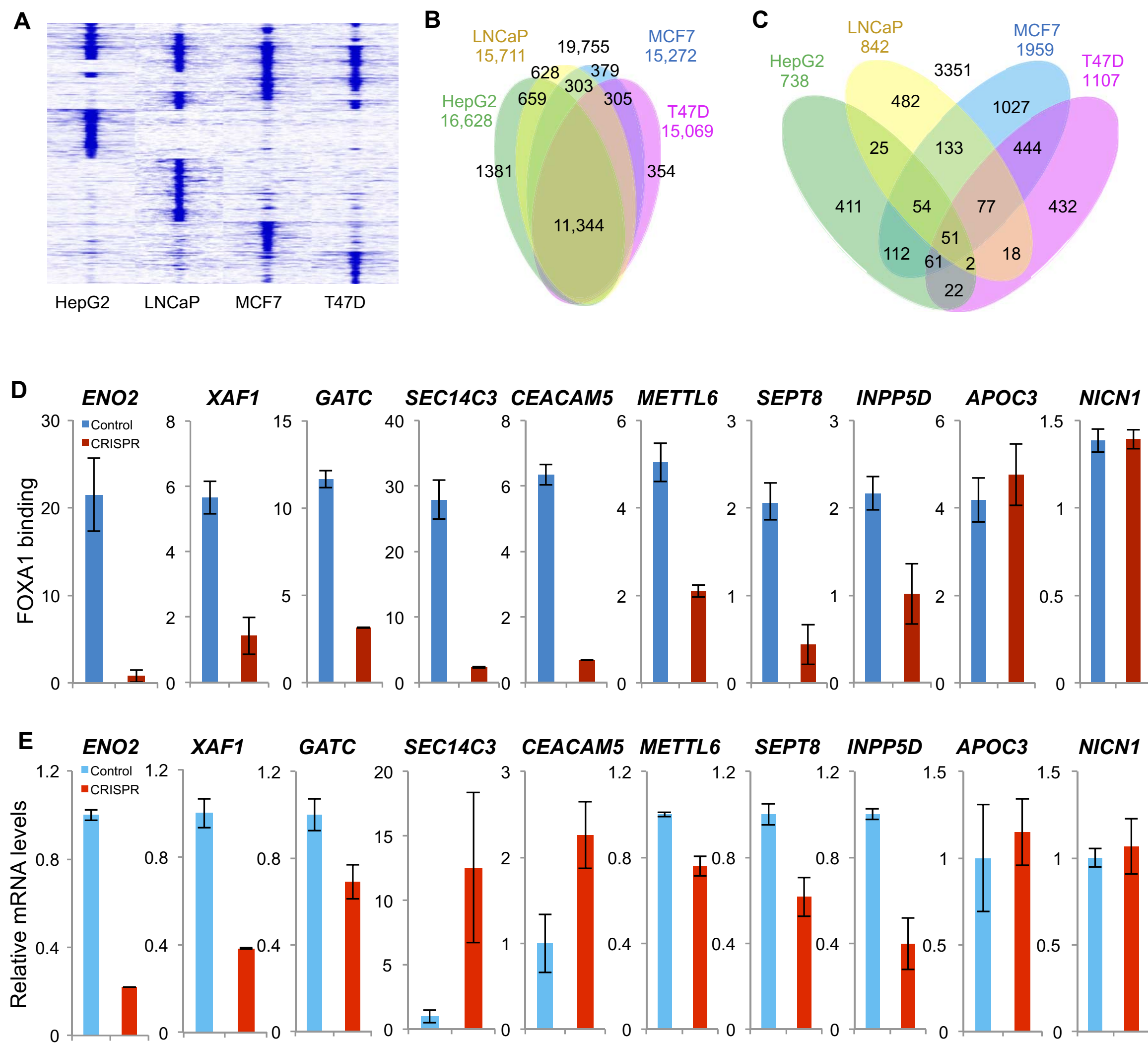

Fig. 1. Cell-specific FOXA1 targeting in human cancers. (**A**) Heatmap analysis of FOXA1 binding in human cancer cells. (**B**) FOXA1-associated genes from ChIP-Seq data are mostly common to human cancer cells. (**C**) Functional and direct FOXA1 target genes are mostly unique to each cancer cell line. (**D** and **E**) FOXA1 binding is disrupted (D) and expression of FOXA1 target genes is reversed (E) by CRISPR-mediated genome editing at selected FOXA1 binding sites. $P < 0.05$ for all assays between the control group and the CRISPR group. Data are means ± SEM.

"Loss-of-function" analysis is generally used to identify the functional regulation of a transcription factor. We analyzed FOXA1-regulated genes from differential gene expression data on these four cell lines with and without FOXA1 knockdown (*6*, *13*–*15*). We found 1005, 1061, 2482, and 1359 FOXA1-regulated genes in HepG2, LNCaP, MCF7, and T47D cells, respectively (fig. S1E). Surprisingly, 40 to 62% of the FOXA1-regulated genes were unique to each cancer cell line, with 40 and 54% of these genes being unique to two similar breast cancer cell lines (MCF7 and T47D, respectively) (fig. S1E). These differentially expressed genes could be regulated directly or indirectly by FOXA1. By intersecting these genes with FOXA1-associated target genes from ChIP-Seq data (Fig. 1B), we found that most (73 to 81%) of the differentially expressed genes were functionally and directly regulated target genes of FOXA1 (Fig. 1C and fig. S1E). About half of these functional and direct FOXA1 target genes were unique to each cancer cell line (Fig. 1C). These data suggest that functional FOXA1 regulation is exhibited in a cell-specific manner in different types of human cancer cells.

Given that FOXA1 defines cancer cell specificity, suppression of FOXA1 in these cancer cells should abolish this unique feature of FOXA1 regulation. First, a heatmap analysis showed clear differential expression patterns of functional FOXA1 target genes in the four human cancer cell lines (fig. S1F). Indeed, these differential expression patterns were greatly attenuated after FOXA1 expression was suppressed (fig. S1F). Further pathway analysis showed that these functional FOXA1 target genes were integral to the processes of cancer, cell growth and proliferation, cell death and survival, organismal injury and abnormality, and cellular development (fig. S3, I to L, and table S1). Collectively, these data strongly support our notion that FOXA1 mostly regulates differential target genes among these four human cancer cell lines.

Unique FOXA1 binding and targeting account for cell-specific FOXA1 regulation

To decipher the mechanisms that underlie this unique feature of cell-specific FOXA1 regulation, we applied three approaches to examining

Downloaded from http://advances.sciencemag.org/ on January 23, 2019

the uniqueness of FOXA1 binding, genetic variations at the core binding elements of FOXA1, and histone markers near FOXA1 binding sites. First, we analyzed the FOXA1 binding peaks associated with functional and direct FOXA1 target genes. Upon intersection with FOXA1 ChIP-Seq data, 7248, 11,689, 13,248, and 11,042 FOXA1 binding peaks were found to be associated with these functional and direct FOXA1 target genes (fig. S1G). In contrast to the uniqueness of the functional FOXA1 target genes (Fig. 1C), only about 18 to 24% of these potentially functional FOXA1 binding peaks were unique to each cell line (fig. S1G). Most of these potentially functional FOXA1 binding peaks were common and mixed peaks for each gene among the four cancer cell lines (fig. S1H), which were different from the mostly mixed peaks found in all of the FOXA1 binding peaks (fig. S1D), although both the number of peaks associated with each gene and the genomic locations (potentially gene regulatory regions) of these peaks were similar between these two groups (fig. S1, C and I, and table S2). When we only analyzed the unique functional FOXA1 target genes in the four cancer cell lines, we found that only 14 to 22% of the FOXA1 binding peaks were unique to each cancer cell line (fig. S1J), and these genes were mainly regulated by mixed FOXA1 binding peaks among the four cell lines (fig. S1K). These unique and functional FOXA1 target genes associated with only unique peaks were calculated in fig. S1L, including both a single FOXA1 peak and multiple unique peaks targeting one unique gene (Unique, n = 1 or n > 1; fig. S2) among the four cancer cell lines. These genes account for about 18% of the total unique and functional FOXA1 target genes (Fig. 1C and fig. S1L).

Collectively, these data indicate that the uniqueness of functional FOXA1 targeting in each cell line is mostly determined by either the unique FOXA1 binding peaks or the activated common FOXA1 binding peaks that could turn on and turn off the transcription of FOXA1 target genes in a cell-specific manner. For unique FOXA1 target genes, except for those genes having only one unique peak, determining which of the multiple FOXA1 binding peaks associated with each target gene is functional remains uncertain. Nevertheless, we could still identify a number of FOXA1 binding peaks of these unique FOXA1 target genes that were solely unique to each cancer cell line (fig. S1L). To validate the functions of the unique FOXA1 binding peaks for these unique FOXA1 target genes in the four cancer cell lines, we applied a novel genome editing approach, CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) (*16*, *17*), to disrupt the core binding element (TG/ATTT) of FOXA1 binding with guide RNA (gRNA)–directed Cas9 nuclease cleavage. We designed eight sets of gRNA fragments to target eight unique FOXA1 binding sites, which were associated with unique FOXA1 target genes, including *ENO2*, *XAF1*, *SEC14C3*, *GATC*, *CEACAM5*, *METTL6*, *INPP5D*, and *SEPT8*, in HepG2, LNCaP, MCF7, and T47D, respectively. After CRISPR treatment, FOXA1 binding was significantly impaired at eight selected sites (Fig. 1D), and the expression of the target genes was significantly reversed (Fig. 1E). In contrast, disruption of FOXA1 binding sites associated with FOXA1-independent genes (*APOC3* in MCF7 cells and *NICN1* in HepG2 cells) by CRISPR treatment did not affect the expression of *APOC3* or *NICN1*, respectively (Fig. 1, D and E). Thus, unique FOXA1 regulation among the four human cancer cell lines is partially attributable to unique FOXA1 binding and targeting. Unique FOXA1 binding in these cancer cells may lead to cell-specific gene expression and phenotypes.

Cell-specific genetic variants at FOXA1 binding sites

Genetic mutations are considered to be a direct cause of cancer. Our recent studies and those of others have shown that genetic variations at FOXA1 and FOXA2 binding sites could introduce a loss of function or a gain of function of FOXA binding and were highly correlated with the incidence and tumorigenesis of liver cancer, prostate cancer, and breast cancer in humans (*7*, *8*, *18*–*22*). Thus, we investigated whether the unique FOXA1 targeting in these four cancer cell lines could be attributable to genetic variants at FOXA1 binding sites, using deep sequencing of the entire genome for HepG2, LNCaP, MCF7, and T47D cells (table S3). We obtained more than 7.6 million high-quality single nucleotide variants (SNVs) from these four cancer cell lines, approximately 14% of which were novel SNVs compared to the human SNP138 database (Fig. 2A). More than 4 million SNVs were found in each human cancer cell line, and about 23% of these were common to all four cancer cell lines (fig. S3A). Further analysis showed that these SNVs were possibly responsible for about half a million gains or losses of the consensus FOXA1 binding element (motif) TG/ATTT in the entire human cancer genome (fig. S3B). An updated analysis of FOXA1 binding peaks and associated genes, with SNV-related FOXA1 motif gain or loss considered, is shown in fig. S3 (C and D). By intersecting the SNV data with FOXA1 ChIP-Seq data, we found 2436 SNVs at the FOXA1 binding element (TG/ATTT) that could lead to a gain or loss of FOXA1 binding peaks (Fig. 2B), accounting for only less than 0.5% of the total SNV-related gain or loss of the TG/ATTT motif in the genome. In addition, 31 to 43% of these SNV-related FOXA1 peak gains or losses were unique to each cancer cell line (Fig. 2B and fig. S3E). We used Sanger sequencing and chromatin immunoprecipitation–quantitative polymerase chain reaction (ChIP-qPCR) to validate 46 of 48 selected FOXA1 binding sites for which SNVs could cause a gain or loss of FOXA1 binding in the four cancer cell lines (table S4). We also found that most of the SNV-related motif gains or losses were located at potential gene regulatory regions, including promoter, intron, and intergenic regions (table S2). Similar to target genes associated with all FOXA1 binding (Fig. 1B), target genes associated with SNV-related FOXA1 peak gains and losses were mostly common to the four cancer cell lines (fig. S3F). The close numbers between FOXA1 peaks from SNV-related gains and losses and their target genes (Fig. 2B and fig. S3F) indicate that these SNV-introduced gains or losses of FOXA1 binding mostly target different genes in human cancer cells.

Next, we analyzed functional FOXA1 target genes associated with these gains or losses of FOXA1 binding by intersecting the functional FOXA1 direct target genes in Fig. 1C with the target genes associated with SNV-introduced gains or losses of FOXA1 binding in fig. S3F. We found that 256 functional FOXA1 target genes could be affected by these SNV-containing FOXA1 binding peaks and that most of them were unique to each cell line (Fig. 2C). However, FOXA1 binding peaks associated with these genes showed much less cell specificity compared to their target genes (fig. S3, G and H), suggesting that most of these unique target genes may be regulated by common or mixed (unique + common) FOXA1 binding peaks. These FOXA1 direct target genes associated with SNV-introduced gains or losses of FOXA1 binding account for about 7 to 14% of all unique FOXA1 direct target genes in the four human cancer cell lines. Next, we performed a pathway analysis of functional FOXA1 target genes associated with those SNV-containing FOXA1 binding peaks and found that most of these genes played key roles in cancer, cell proliferation

Downloaded from http://advances.sciencemag.org/ on January 23, 2019

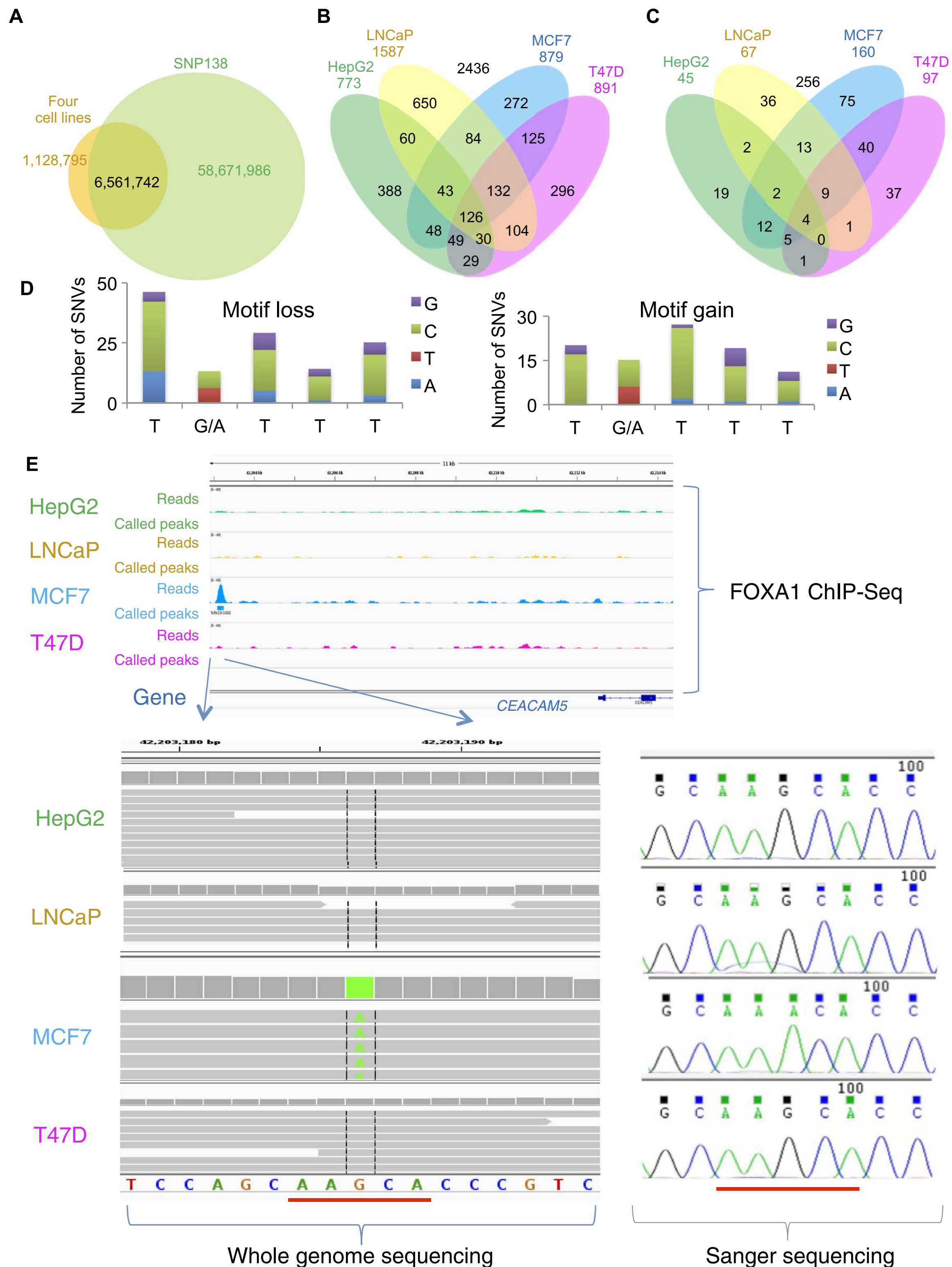

Fig. 2. Unique FOXA1 targeting mediated by genetic variants. (**A**) Comparison of SNVs from the four cancer cell lines (HepG2, LNCaP, MCF7, and T47D) with SNP138. (**B**) SNV-introduced gain and loss of FOXA1 binding peaks. (**C**) Functional FOXA1 target genes associated with SNV-introduced gain and loss of FOXA1 binding sites. (**D**) Frequency of SNVs in the FOXA1 binding motifs for SNV-led motif loss or gain from (C) in all of the four cancer cell lines. (**E**) Example of SNV-introduced gain of FOXA1 binding sites targeting *CEACAM5* in MCF7 cells.

Downloaded from http://advances.sciencemag.org/ on January 23, 2019

and growth, and cell death and survival (fig. S3, I to L, and table S1). We also analyzed the frequency of these variants at each nucleotide of the FOXA1 binding motif for those functional FOXA1 binding peaks in Fig. 2C. Overall, the changes between T and C were mostly enriched in the entire motif for both motif loss and gain (Fig. 2D). In SNV-led motif loss, the first nucleotide T had the most frequently changed variants, whereas the third nucleotide T was most frequently changed in SNV-led motif gain (Fig. 2D).

To further validate the gain of the TG/ATTT motif in the functional and unique FOXA1 binding peaks, we used the CRISPR technology to disrupt a gain mutation in a FOXA1 binding site associated with the *CEACAM5* gene in MCF7 cells and found that CRISPR led to impaired FOXA1 binding and reversed gene expression of *CEACAM5* (Figs. 1, D and E, and 2E). These data suggest that cell-specific FOXA1 targeting in human cancer cells could also result from genetic variants at FOXA1 binding sites.

Epigenetic regulation in the functioning of cell-specific FOXA1 targeting

Recent studies showed that H3K4me1/H3K4me2 and H3K27ac marked active enhancer regions (*23–25*), whereas FOXA1 mostly bound to intragenic enhancer regions in the four human cancer cell lines (table S2), consistent with our previous studies (*26*, *27*). Studies have shown that epigenetic factors of histone variants and modification (such as H2A.Z, H3K4me1, and H3K4me2) play critical roles in the functioning of FOXA factors (*10*, *27*). Thus, we speculated that the unique feature of cell-specific FOXA1 regulation might also result from differential epigenetic regulation near FOXA1 binding sites, especially for those common FOXA1 binding peaks that are found in multiple cell lines but are only functional in certain cell lines. We sequenced and collected 32 ChIP-Seq data sets for the peak calling of eight histone markers (including H2A.Z, H3K4me1, H3K4me2, H3K4me3, H3K9me3, H3K27me3, H3K27ac, and H3K36me3) in these four cancer cell lines from our own libraries and from the ENCODE and GEO databases (see Supplementary Materials and Methods). To examine the potential impact of histone modification on the functioning of FOXA1 regulation, we analyzed the distributions of these eight histone markers near FOXA1 binding sites and found that H2A.Z, H3K4me1, H3K4me2, H3K4me3, H3K27ac, and H3K36me3 were enriched near FOXA1 binding sites associated with functional FOXA1 target genes (fig. S4A), but that only H2A.Z, H3K4me1, H3K4me2, and H3K27ac showed cell-specific enrichment near FOXA1 binding sites that were only associated with those functional and direct FOXA1 target genes (Fig. 3, A and B, and fig. S4, B to K); for example, for each cell line, H3K27ac was enriched near functional FOXA1 binding but not near nonfunctional FOXA1 binding (control) with regard to their target genes (Fig. 3A); for all four cell lines, H3K27ac was only enriched near cell-specific functional FOXA1 binding sites (Fig. 3B). However, cell-specific enrichment of H3K4me3 near FOXA1 binding sites was only observed in HepG2, LNCaP, and MCF7 cells, but not

Downloaded from http://advances.sciencemag.org/ on January 23, 2019

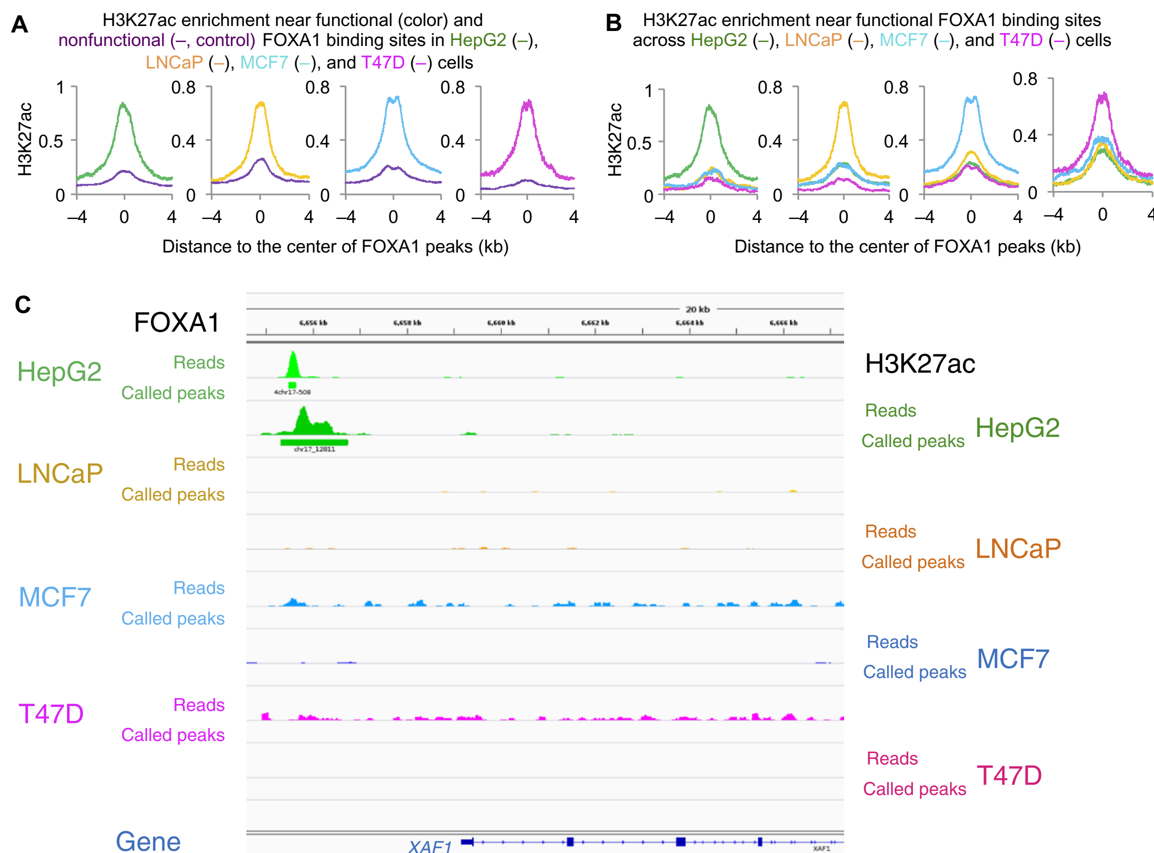

Fig. 3. Correlations of histone modification with functional FOXA1 targeting. (**A** and **B**) Cell-specific enrichment of H3K27ac near functional (A) and cell-specific (B) FOXA1 binding peaks. (**C**) Example of H3K27ac enrichment near FOXA1 binding sites targeting *XAF1* in HepG2 cells.

in T47D cells (fig. S4, H and I); cell-specific enrichment of H3K36me3 near FOXA1 binding sites was barely observed in any of the four cancer cell lines (fig. S4, J and K). Further analysis showed that no different enrichment of these histone markers was observed near unique and common FOXA1 binding peaks associated with functional FOXA1 target genes, indicating that histone modification might be essential for all functional FOXA1 binding.

To further investigate the functionality of FOXA1 binding associated with these histone markers, we applied the CRISPR technology to disrupt a selected FOXA1 binding site associated with H3K27ac for the *XAF1* gene and found that CRISPR led to impaired FOXA1 binding and reversed gene expression of *XAF1* (Figs. 1, D and E, and 3C). These data suggest that the functioning of cell-specific FOXA1 targeting in human cancer cells could require certain histone modification and that H2A.Z, H3K4me1, H3K4me2, and H3K27ac could mark functional FOXA1 binding and targeting in the human cancer genome.

DISCUSSION

We discovered a novel feature of FOXA1 regulation in liver, prostate, and breast cancer cells in humans: there is unique FOXA1 targeting in each cancer cell type and even between two breast cancer cell lines. We also found that this unique regulation of FOXA1 was determined by the unique binding of FOXA1 in the four human cancer cell lines, by a gain or loss of FOXA1 binding from genetic variants at the cis-regulatory elements of FOXA1, and/or by potential epigenetic regulation near functional FOXA1 binding sites. Future investigation on how these epigenetic factors participate in FOXA1 regulation is guaranteed. Transcriptional regulation is determined by both the binding and the functioning of transcription factors. FOXA1 binds nearly 90% of the genes in the human cancer genome, but only less than 17% of this FOXA1 targeting is functional, and these functional FOXA1 targets are mostly specific to cancer cell type (Fig. 1, B and C). These findings allow us to raise a "flower-blooming" hypothesis for cell-specific transcriptional regulation in human cancer cells (Fig. 4). The "blooming" (functioning) of a transcription factor in a cell-specific manner could be controlled by its unique binding, genetic variations, epigenetic regulation, and other factors, including coregulators and chromatin remodeling (Fig. 4). The next question is, "How can this hypothesis be generalized to other transcription factors?" In addition, what are the functions of the rest of the "nonfunctional" FOXA1 binding in the genome? Our studies and those of others have shown that FOXA1 binding might be involved in chromatin remodeling, might act as a scaffold for other transcription factors such as steroid hormone receptors, and/or might mark enhancer regions in the genome (*10*, *19*, *27*). In addition, CRISPR-mediated genome editing at FOXA1 binding sites or other transcription factor binding sites could be a powerful tool for answering these questions by manipulating targeted gene transcription at selected binding sites of transcription factors. Nevertheless, our findings of cell-specific FOXA1 regulation among

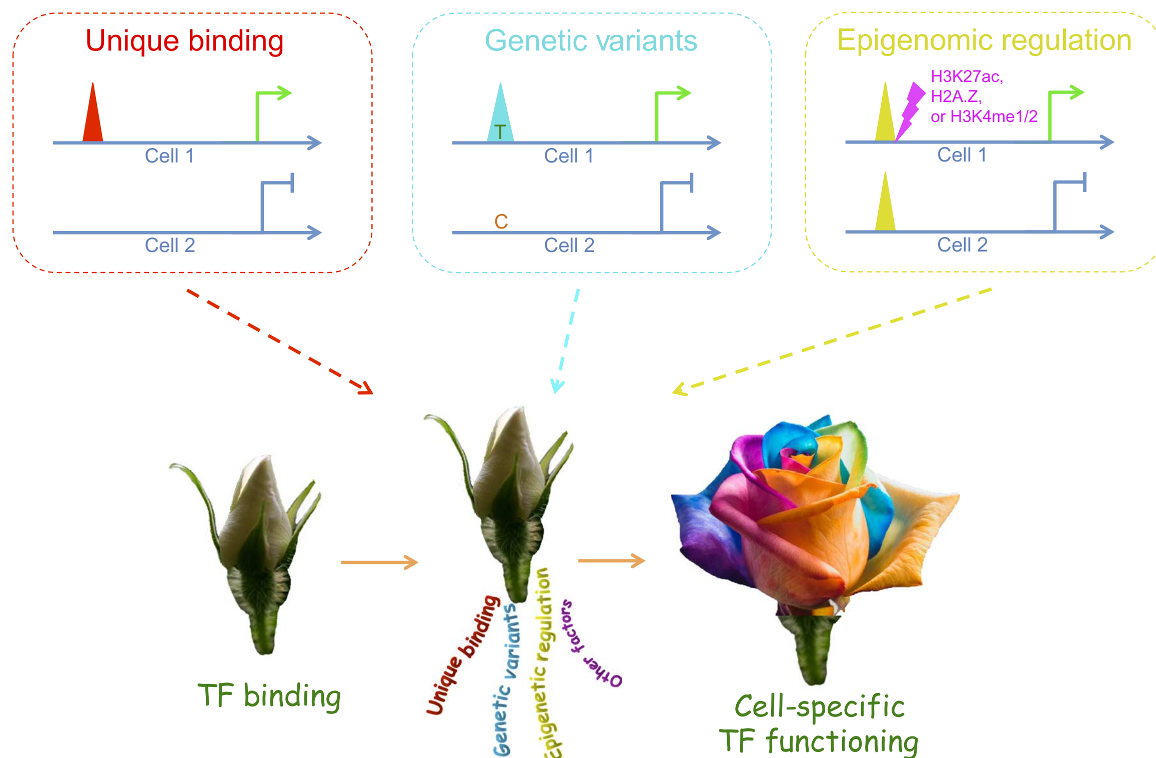

Fig. 4. Flower-blooming hypothesis for cell-specific transcriptional regulation. Cell-specific functioning of a transcription factor (TF) could be determined by unique binding of TF, genetic variants at TF binding sites, potential epigenomic regulation near TF binding sites, and/or other factors. TF regulation from highly common TF binding across cell types to cell-specific functional TF binding resembles the flower-blooming process. Triangle peaks represent TF binding.

Downloaded from http://advances.sciencemag.org/ on January 23, 2019

the four human cancer cell lines not only provide mechanistic explanations for cell-specific transcriptional regulation of a single transcription factor but also shed light on the identification of novel cell-specific biomarkers for liver, prostate, and breast cancers in humans. In addition, unique FOXA1 regulation between two breast cancer cell lines indicates that this unique feature of FOXA1 regulation may also be utilized as an individual biomarker and therapeutic target for each breast cancer patient. Related studies of patient-derived tumor samples will provide better answers regarding this aspect and will be worth undertaking in the future.

MATERIALS AND METHODS

Genomic location analysis of FOXA1 and histones

We generated ChIP-Seq data for H2A.Z, H3K4me2, H3K9me3, H3K27me3, H3K27ac, and H3K36me3 in T47D cells. FOXA1 ChIP-Seq and other ChIP-Seq data for histone markers in HepG2, LNCaP, MCF7, and T47D cells are provided with a detailed protocol in Supplementary Materials and Methods.

Whole genome sequencing by a PCR-free approach

Genomic DNA from the four human cancer cell lines HepG2, LNCaP, MCF7, and T47D was sequenced with a PCR-free approach using HiSeq 2000 (Illumina). Sequencing reads were analyzed using a Mayo Clinic–developed pipeline for analysis of genetic variants. SNVs were used for the analysis of motif gain or loss at FOXA1 binding sites. The validation of SNVs was performed using regular Sanger sequencing (table S4). A detailed protocol is provided in Supplementary Materials and Methods.

CRISPR-mediated disruption of FOXA1 binding

A novel genome editing technology, CRISPR, was applied to disrupt FOXA1 binding sites with gRNA-directed Cas9 nuclease cleavage (*16*, *17*, *28*–*31*). A detailed protocol is provided in Supplementary Materials and Methods.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/2/3/e1501473/DC1
Materials and Methods
Fig. S1. Cell-specific FOXA1 binding and targeting.
Fig. S2. Five types of FOXA1 targeting in the four human cancer cell lines.
Fig. S3. SNV-led gain or loss of FOXA1 binding and targeting.
Fig. S4. Histone modification in cell-specific FOXA1 regulation.
Table S1. Pathway analysis of functional FOXA1 target genes in human cancer cells by Ingenuity.
Table S2. Genomic annotation of FOXA1 binding peaks in human cancer cells.
Table S3. Whole genome sequencing data for human cancer cells.
Table S4. Validation of SNV-introduced gain and loss of FOXA1 binding in human cancer cells.
References (*32*–*38*)

REFERENCES AND NOTES

1. D. T. Odom, R. D. Dowell, E. S. Jacobsen, W. Gordon, T. W. Danford, K. D. MacIsaac, P. A. Rolfe, C. M. Conboy, D. K. Gifford, E. Fraenkel, Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat. Genet.* **39**, 730–732 (2007).
2. J. A. D'Alessio, K. J. Wright, R. Tjian, Shifting players and paradigms in cell-specific transcription. *Mol. Cell* **36**, 924–931 (2009).
3. K. H. Kaestner, The FoxA factors in organogenesis and differentiation. *Curr. Opin. Genet. Dev.* **20**, 527–532 (2010).
4. K. M. Jozwik, J. S. Carroll, Pioneer factors in hormone-dependent cancers. *Nat. Rev. Cancer* **12**, 381–385 (2012).
5. G. M. Bernardo, R. A. Keri, FOXA1: A transcription factor with parallel functions in development and cancer. *Biosci. Rep.* **32**, 113–130 (2012).
6. A. Hurtado, K. A. Holmes, C. S. Ross-Innes, D. Schmidt, J. S. Carroll, FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat. Genet.* **43**, 27–33 (2011).
7. R. Cowper-Sal lari, X. Zhang, J. B Wright, S. D. Bailey, M. D. Cole, J. Eeckhoute, J. H. Moore, M. Lupien, Breast cancer risk–associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. *Nat. Genet.* **44**, 1191–1198 (2012).
8. Z. Li, G. Tuteja, J. Schug, K. H. Kaestner, Foxa1 and Foxa2 are essential for sexual dimorphism in liver cancer. *Cell* **148**, 72–83 (2012).
9. J. S. Carroll, X. S. Liu, A. S. Brodsky, W. Li, C. A. Meyer, A. J. Szary, J. Eeckhoute, W. Shao, E. V. Hestermann, T. R. Geistlinger, E. A. Fox, P. A. Silver, M. Brown, Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**, 33–43 (2005).
10. M. Lupien, J. Eeckhoute, C. A. Meyer, Q. Wang, Y. Zhang, W. Li, J. S. Carroll, X. S. Liu, M. Brown, FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**, 958–970 (2008).
11. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
12. S. Heinz, C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh, C. K. Glass, Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
13. H.-J. Jin, J. C. Zhao, I. Ogden, R. C. Bergan, J. Yu, Androgen receptor–independent function of FoxA1 in prostate cancer metastasis. *Cancer Res.* **73**, 3725–3736 (2013).
14. J. Wang, J.-S. Park, Y. Wei, M. Rajurkar, J. L. Cotton, Q. Fan, B. C. Lewis, H. Ji, J. Mao, TRIB2 acts downstream of Wnt/TCF in liver cancer cells to regulate YAP and C/EBPα function. *Mol. Cell* **51**, 211–225 (2013).
15. G. M. Bernardo, G. Bebek, C. L. Ginther, S. T. Sizemore, K. L. Lozada, J. D. Miedler, L. A. Anderson, A. K. Godwin, F. W. Abdul-Karim, D. J. Slamon, R. A. Keri, FOXA1 represses the molecular phenotype of basal breast cancer cells. *Oncogene* **32**, 554–563 (2013).
16. O. Shalem, N. E. Sanjana, E. Hartenian, X. Shi, D. A. Scott, T. S. Mikkelsen, D. Heckl, B. L. Ebert, D. E. Root, J. G. Doench, F. Zhang, Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
17. M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
18. X. Zhang, R. Cowper-Sal lari, S. D. Bailey, J. H. Moore, M. Lupien, Integrative functional genomics identifies an enhancer looping to the *SOX9* gene disrupted by the 17q24.3 prostate cancer risk locus. *Genome Res.* **22**, 1437–1446 (2012).
19. C. S. Ross-Innes, R. Stark, A. E. Teschendorff, K. A. Holmes, H. R. Ali, M. J. Dunning, G. D. Brown, O. Gojis, I. O. Ellis, A. R. Green, S. Ali, S.-F. Chin, C. Palmieri, C. Caldas, J. S. Carroll, Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
20. K. B. Meyer, M. O'Reilly, K. Michailidou, S. Carlebur, S. L. Edwards, J. D. French, R. Prathalingham, J. Dennis, M. K. Bolla, Q. Wang, I. de Santiago, J. L. Hopper, H. Tsimiklis, C. Apicella, M. C. Southey, M. K. Schmidt, A. Broeks, L. J. Van't Veer, F. B. Hogervorst, K. Muir, A. Lophatananon, S. Stewart-Brown, P. Siriwanarangsan, P. A. Fasching, M. P. Lux, A. B. Ekici, M. W. Beckmann, J. Peto, I. dos Santos Silva, O. Fletcher, N. Johnson, E. J. Sawyer, I. Tomlinson, M. J. Kerin, N. Miller, F. Marme, A. Schneeweiss, C. Sohn, B. Burwinkel, P. Guénel, T. Truong, P. Laurent-Puig, F. Menegaux, S. E. Bojesen, B. G. Nordestgaard, S. F. Nielsen, H. Flyger, R. L. Milne, M. P. Zamora, J. I. Arias, J. Benitez, S. Neuhausen, H. Anton-Culver, A. Ziogas, C. C. Dur, H. Brenner, H. Müller, V. Arndt, C. Stegmaier, A. Meindl, R. K. Schmutzler, C. Engel, N. Ditsch, H. Brauch, T. Brüning, Y.-D. Ko; The GENICA Network, H. Nevanlinna, T. A. Muranen, K. Aittomäki, C. Blomqvist, K. Matsuo, H. Ito, H. Iwata, Y. Yatabe, T. Dörk, S. Helbig, N. V. Bogdanova, A. Lindblom, S. Margolin, A. Mannermaa, V. Kataja, V.-M. Kosma, J. M. Hartikainen, G. Chenevix-Trench; kConFab Investigators, Australian Ovarian Cancer Study Group, A. H. Wu, C.-C. Tseng, D. Berg, D. O. Stram, D. Lambrechts, B. Thienpont, M.-R. Christiaens, A. Smeets, J. Chang-Claude, A. Rudolph, P. Seibold, D. Flesch-Jayns, P. Radice, P. Peterlongo, B. Bonanni, L. Bernard, F. J. Couch, J. E. Olson, X. Wang, K. Purrington, G. G. Giles, G. Severi, L. Baglietto, C. Mclean, C. A. Haiman, B. E. Henderson, F. Schumacher, L. Le Marchand, J. Simard, M. S. Goldberg, F. Labrèche, M. Dumont, S.-H. Teo, C.-H. Yip, S.-Y. Phuah, V. Kristensen, G. G. Alnæs, A.-L. Børresen-Dale, W. Zheng, S. Deming-Halverson, M. Shrubsole, J. Long, R. Winqvist, K. Pylkäs, A. Jukkola-Vuorinen, S. Kauppila, I. L. Andrulis, J. A. Knight, G. Glendon, S. Tchatchou, P. Devilee, R. A. Tollenaar, C. M. Seynaeve, M. García-Closas, J. Figueroa, S. J. Chanock, J. Lissowska, K. Czene, H. Darabi, K. Eriksson, M. J. Hooning, J. W. M. Martens,

Downloaded from http://advances.sciencemag.org/ on January 23, 2019

A. M. W. van den Ouweland, C. H. M. van Deurzen, P. Hall, J. Li, J. Liu, K. Humphreys, X.-O. Shu, W. Lu, Y.-T. Gao, H. Cai, A. Cox, M. W. R. Reed, W. Blot, L. B. Signorello, Q. Cai, P. D. P. Pharoah, M. Ghoussaini, P. Harrington, J. Tyrer, D. Kang, J.-Y. Choi, S. K. Park, D.-Y. Noh, M. Hartman, M. Hui, W.-Y. Lim, S. A. Buhari, U. Hamann, A. Försti, T. Rüdiger, H.-U. Ulmer, A. Jakubowska, J. Lubinski, K. Jaworska, K. Durda, S. Sangrajrang, V. Gaborieau, P. Brennan, J. McKay, C. Vachon, S. Slager, F. Fostira, R. Pilarski, C.-Y. Shen, C.-N. Hsiung, P.-E. Wu, M.-F. Hou, A. Swerdlow, A. Ashworth, N. Orr, M. J. Schoemaker, B. A. J. Ponder, A. M. Dunning, D. F. Easton, Fine-scale mapping of the FGFR2 breast cancer risk locus: Putative functional variants differentially bind FOXA1 and E2F1. *Am. J. Hum. Genet.* **93**, 1046–1060 (2013).

21. C. S. Grasso, Y.-M. Wu, D. R. Robinson, X. Cao, S. M. Dhanasekaran, A. P. Khan, M. J. Quist, X. Jing, R. J. Lonigro, J. C. Brenner, I. A. Asangani, B. Ateeq, S. Y. Chun, J. Siddiqui, L. Sam, M. Anstett, R. Mehra, J. R. Prensner, N. Palanisamy, G. A. Ryslik, F. Vandin, B. J. Raphael, L. P. Kunju, D. R. Rhodes, K. J. Pienta, A. M. Chinnaiyan, S. A. Tomlins, The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487**, 239–243 (2012).
22. D. J. Hazelett, S. G. Coetzee, G. A. Coetzee, A rare variant, which destroys a FoxA1 site at 8q24, is associated with prostate cancer risk. *Cell Cycle* **12**, 379–380 (2013).
23. M. P. Creyghton, A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, J. Hanna, M. A. Lodato, G. M. Frampton, P. A. Sharp, L. A. Boyer, R. A. Young, R. Jaenisch, Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21931–21936 (2010).
24. N. D. Heintzman, R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. Van Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, B. Ren, Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**, 311–318 (2007).
25. A. Barski, S. Cuddapah, K. Cui, T.-Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, K. Zhao, High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
26. Z. Li, J. Schug, G. Tuteja, P. White, K. H. Kaestner, The nucleosome map of the mammalian liver. *Nat. Struct. Mol. Biol.* **18**, 742–746 (2011).
27. Z. Li, P. Gadue, K. Chen, Y. Jiao, G. Tuteja, J. Schug, W. Li, K. H. Kaestner, Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. *Cell* **151**, 1608–1616 (2012).
28. R. Barrangou, C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, P. Horvath, CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
29. M. Jinek, A. East, A. Cheng, S. Lin, E. Ma, J. Doudna, RNA-programmed genome editing in human cells. *eLife* **2**, e00471 (2013).
30. P. Mali, L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, G. M. Church, RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
31. T. Wang, J. J. Wei, D. M. Sabatini, E. S. Lander, Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014).
32. A. J. Saldanha, Java Treeview—Extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004).
33. J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
34. H. Thorvaldsdottir, J. T. Robinson, J. P. Mesirov, Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
35. I. Kozarewa, Z. Ning, M. A. Quail, M. J. Sanders, M. Berriman, D. J. Turner, Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nat. Methods* **6**, 291–295 (2009).
36. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
37. P. Cingolani, A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wanga, S. J. Land, X. Lua, D. M. Rudena, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w^{1118}; iso-2; iso-3. *Fly* **6**, 80–92 (2012).
38. C. Wang, J. M. Evans, A. V. Bhagwate, N. Prodduturi, V. Sarangi, M. Middha, H. Sicotte, P. T. Vedell, S. N. Hart, G. R. Oliver, J.-P. A. Kocher, M. J. Maurer, A. J. Novak, S. L. Slager, J. R. Cerhan, Y. W. Asmann, PatternCNV: A versatile tool for detecting copy number changes from exome sequencing data. *Bioinformatics* **30**, 2687–2689 (2014).

Acknowledgments: We thank K. H. Kaestner (University of Pennsylvania), L. A. Chodosh (University of Pennsylvania), and M. Brown (Dana-Farber Cancer Institute and Harvard University) for their mentorship of Z.L. during the K99 phase of the NIH Pathways to Independence Award. We thank J. M. Kachergus (A. Thompson's laboratory) for her instructions on Sanger sequencing. We acknowledge the support of the Epigenomics Program of the Center for Individualized Medicine at Mayo Clinic (led by T. Ordog) and thank J. H. Lee for his technical support. We also thank the staff of the Sequencing Core and Bioinformatics Core facilities at Mayo Clinic for their technical support. We thank K. Viola (editor of the life sciences from Mayo Clinic) for her assistance with manuscript proofreading. **Funding:** This project was supported by the NIH Pathways to Independence Award (K99/R00) to Z.L. (grant R00CA168983). **Author contributions:** Z.L. designed and wrote the manuscript. G.Z., Y.Z., Y.L., L.-P.K., B.S., and X.W. performed the experiments and wrote part of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Gene Expression Omnibus, GSE63109. All data used to obtain the conclusions in this paper are available in GSE63109 or presented in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 19 October 2015
Accepted 27 January 2016
Published 18 March 2016
10.1126/sciadv.1501473

Citation: G. Zhang, Y. Zhao, Y. Liu, L.-P. Kao, X. Wang, B. Skerry, Z. Li, FOXA1 defines cancer cell specificity. *Sci. Adv.* **2**, e1501473 (2016).

Downloaded from http://advances.sciencemag.org/ on January 23, 2019

Science Advances

FOXA1 defines cancer cell specificity

Gaihua Zhang, Yongbing Zhao, Yi Liu, Li-Pin Kao, Xiao Wang, Benjamin Skerry and Zhaoyu Li

Sci Adv **2** (3), e1501473.
DOI: 10.1126/sciadv.1501473

ARTICLE TOOLS http://advances.sciencemag.org/content/2/3/e1501473

SUPPLEMENTARY MATERIALS http://advances.sciencemag.org/content/suppl/2016/03/15/2.3.e1501473.DC1

REFERENCES This article cites 38 articles, 10 of which you can access for free
http://advances.sciencemag.org/content/2/3/e1501473#BIBL

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science Advances* is a registered trademark of AAAS.

Downloaded from http://advances.sciencemag.org/ on January 23, 2019