Cross-Species Regulatory Network Analysis Identifies a Synergistic Interaction between FOXM1 and CENPF that Drives Prostate Cancer Malignancy

Alvaro Aytes,1,8,15 Antonina Mitrofanova,2,16 Celine Lefebvre,5,15,16 Mariano J. Alvarez,2 Mireia Castillo-Martin,9 Tian Zheng,7,16 James A. Eastham,11 Anuradha Gopalan,12 Kenneth J. Pienta,13,14 Michael M. Shen,1,2,3,4,7 Andrea Califano,2,5,7,* and Cory Abate-Shen1,2,6,7,*

1Department of Urology, Columbia University Medical Center, New York, NY 10032, USA
2Department of Systems Biology, Columbia University Medical Center, New York, NY 10032, USA
3Department of Medicine, Columbia University Medical Center, New York, NY 10032, USA
4Department of Genetics & Development, Columbia University Medical Center, New York, NY 10032, USA
5Department of Biochemistry and Molecular Biophysics, Columbia University Medical Center, New York, NY 10032, USA
6Department of Pathology and Cell Biology, Columbia University Medical Center, New York, NY 10032, USA
7Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, New York, NY 10032, USA
8Translational Research Laboratory, Catalan Institute of Oncology, Bellvitge Institute for Biomedical Research, L’Hospitalet de Llobregat, Barcelona 08907, Spain
9Department of Pathology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
10Department of Statistics, Columbia University, New York, NY 10027, USA
11Department of Urology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
12Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
13The University of Michigan, Ann Arbor, MI 48109, USA
14Brady Urological Institute at the Johns Hopkins School of Medicine, Baltimore, MD 21231, USA
15Co-first author
16Present address: Inserm U981, Gustave Roussy Cancer Institute, 94805 Villejuif Cedex, France
*Correspondence: califano@c2b2.columbia.edu (A.C.), cabateshen@columbia.edu (C.A.-S.)
http://dx.doi.org/10.1016/j.ccr.2014.03.017

SUMMARY

To identify regulatory drivers of prostate cancer malignancy, we have assembled genome-wide regulatory networks (interactomes) for human and mouse prostate cancer from expression profiles of human tumors and of genetically engineered mouse models, respectively. Cross-species computational analysis of these interactomes has identified FOXM1 and CENPF as synergistic master regulators of prostate cancer malignancy. Experimental validation shows that FOXM1 and CENPF function synergistically to promote tumor growth by coordinated regulation of target gene expression and activation of key signaling pathways associated with prostate cancer malignancy. Furthermore, co-expression of FOXM1 and CENPF is a robust prognostic indicator of poor survival and metastasis. Thus, genome-wide cross-species interrogation of regulatory networks represents a valuable strategy to identify causal mechanisms of human cancer.

INTRODUCTION

It is widely appreciated that cancer is not a single entity but rather a highly individualized spectrum of diseases characterized by a large number of molecular alterations (Hanahan and Weinberg, 2011). Distinguishing those that constitute true drivers of cancer phenotypes from the multitude that are simply deregulated has proven to be a daunting task, which is further exacerbated by

Significance

Genetically engineered mouse models have been widely used for in vivo analyses of cancer phenotypes as well as preclinical investigations. However, inherent species differences often hinder the appropriate extrapolation of studies performed in mice to human cancer. Here we introduce a strategy using cross-species computational analysis of context-specific regulatory networks for the effective integration of experimental findings from mouse models and human cancer. This approach enables the identification of conserved master regulators of malignant prostate cancer, as well as elucidation of their synergistic interaction. This computational paradigm should be broadly applicable for elucidating causal mechanisms of cancer, as well as integrating preclinical analyses from mouse to man.
the complexity of elucidating how such drivers interact synergistically to elicit cancer phenotypes. In this regard, prostate cancer is particularly challenging because its notorious heterogeneity, combined with a relative paucity of recurrent gene mutations, has made it especially difficult to identify molecularly distinct subtypes with known clinical outcomes (Baca et al., 2013; Schoenbom et al., 2013; Shen and Abate-Shen, 2010). Additionally, whereas most early-stage prostate tumors are readily treatable (Cooperberg et al., 2007), advanced prostate cancer frequently progresses to castration resistance, which is often metastatic and nearly always fatal (Ryan and Tindall, 2011; Scher and Sawyers, 2005). Thus, there is a pressing need to identify bona fide determinants of aggressive prostate cancer as well as prognostic biomarkers of disease outcome.

Analysis of genetically engineered mouse models (GEMMs) can circumvent inherent challenges associated with the intrinsic complexity of more heterogeneous human cancer phenotypes. Indeed, investigations of mouse models of prostate cancer have contributed to characterization of disease-specific pathways, led to the identification of biomarkers of disease progression, and provided useful preclinical models for prevention and therapy (Irshad and Abate-Shen, 2013; Ittmann et al., 2013). Following the description of an initial transgenic model nearly 20 years ago, there are now numerous GEMMs that collectively model key molecular pathways deregulated in human prostate cancer and recapitulate the various stages of disease progression, including preinvasive lesions (prostatic intraepithelial neoplasia; PIN), adenocarcinoma, castration resistance, and metastasis (Irshad and Abate-Shen, 2013; Ittmann et al., 2013).

However, inherent species differences often hinder direct comparative analyses of mouse models and human cancer. Indeed, such analyses would greatly benefit from computational approaches that enable accurate cross-species integration of regulatory information from mouse to man. Recent advances in systems biology have led to the reverse engineering of regulatory networks (interactomes) that integrate large-scale data sets encompassing expression profiles, protein-protein interactions, genomic alterations, and epigenetic changes associated with cancer and other diseases (Lefebvre et al., 2012). However, whereas individual analyses of human and murine interactomes have led to relevant biological discoveries, their cross-species interrogation has not been systematically implemented.

Here, we introduce an approach for accurate cross-species analysis of conserved cancer pathways based on reverse engineering of genome-wide regulatory networks (i.e., interactomes) representing both human and mouse prostate cancer. To accomplish this, we have produced a regulatory network based on in vivo perturbation of a repertoire of mouse cancer models and implemented comparative analysis with a complementary regulatory network generated from human prostate cancer data sets. Cross-species computational interrogation of these paired interactomes, followed by experimental and clinical validation, has elucidated the synergistic interaction of FOXM1 and CENPF as a driver of prostate cancer malignancy. We propose that analyses of genome-wide, cross-species regulatory networks will provide an effective paradigm for elucidating causal mechanisms of human cancer and other complex diseases.

RESULTS

We developed a strategy for genome-wide interrogation of cancer phenotypes based on accurate integration of experimental data from model organisms and human cancer (Figure 1). First, we generated regulatory networks (interactomes) for human and mouse prostate cancer using the Algorithm for the Reconstruction of Accurate Cellular Networks (ARACNe; Basso et al., 2005; Margolin et al., 2006b). We next evaluated the suitability of these mouse and human interactomes for cross-species interrogation using a computational approach to assess the global conservation of their transcriptional programs. We then used the Master Regulator Inference algorithm (MARINa; Carro et al., 2010; Lefebvre et al., 2010) to infer candidate master regulators that act individually or synergistically to drive malignant prostate cancer. Finally, we performed experimental studies to validate synergistic interactions of master regulators, to elucidate underlying mechanisms, and to evaluate their clinical relevance.

Assembly of Interactomes for Human and Mouse Prostate Cancer

ARACNe is an unbiased algorithm that infers direct transcriptional interactions based on the mutual information between each transcriptional regulator and its potential targets. For optimal analyses, ARACNe requires large data sets of gene expression profiles (≥100) having significant endogenous (i.e., genetic and/or exogenous, i.e., perturbation-induced) heterogeneity. To assemble a human prostate cancer interactome, we analyzed the expression profile data set reported elsewhere (Taylor et al., 2010), which is ideally suited for ARACNe because: (1) it is relatively large (n = 185) and diverse, including primary tumors, adjacent normal tissue, metastases, and cell lines; (2) its primary tumors encompass the full range of pathological Gleason scores and have well-annotated clinical outcome data; and (3) it displays extensive genetic diversity and tumor heterogeneity, as shown by t-Distributed Stochastic Neighbor Embedding (t-SNE) analysis (Figure 2; Table S1 available online). Notably, interactomes assembled from three alternative human prostate cancer data sets (Table S1) were neither as complete nor as extensive (data not shown).

To assemble a corresponding mouse prostate cancer interactome, it was first necessary to generate an expression profile data set of appropriate size and representing sufficient expression variability. We selected 13 distinct GEMMs, which together represent the full spectrum of prostate cancer phenotypes, including normal epithelium (wild-type), low-grade PIN (Nkx3.1 and APT), high-grade PIN, and adenocarcinoma (APT-P, APC, Myc, NP, Erg-P, and NPS3), castration resistance (NP-AI), and metastatic prostate cancer (NPB, NPK, and TRAMP; Figure S1A; Table S2). To further increase the variability of the expression profiles, we introduced a controlled set of exogenous perturbations by in vivo administration of 13 small-molecule perturbagens to each GEMM. Perturbagens were selected for their clinical relevance and/or ability to modulate key prostate cancer pathways, including hormone signaling (testosterone, calcitriol, and enzalutamide); P13 kinase activity (MK2206, LY294002, and rapamycin); MAP kinase activity (PD035901); tyrosine kinase activity (imatinib, dasatinib, and sorafenib); NFκB signaling...
I. Assembly of interactomes

Human Interactome
- Human prostate cancer
  - Normal
  - Primary tumors
  - Metastasis
- Dataset: 185 GEPs

Mouse Interactome
- Prostate cancer GEMMs
  - Normal
  - LG-PIN
  - HG-PIN/Cancer
  - CRPC
  - Metastasis
- Perturbagens
  - Hormone signaling
  - Signaling pathways
  - Kinase inhibitors
  - Other pathways
- Dataset: 384 GEPs

II. Genome-wide conservation analysis

Regulon Activity — Human

Regulon Activity — Mouse

Correlation of human and mouse

III. Master regulator analysis

Human Gene Signature
- Repressed Targets
- Activated Targets

Mouse Gene Signature
- Repressed Targets
- Activated Targets

Conserved Master Regulators

Synergy Analysis

IV. Validation of Synergistic Master Regulators

Functional Validation

Molecular Validation

Clinical Validation

Figure 1. Strategy for Genome-wide Cross-Species Analyses of Prostate Cancer
Following pilot studies to define appropriate doses and schedule (Figures S1B–S1D), we adopted a universal treatment schedule of one treatment per day for 5 days with dosage determined independently for each perturbagen (Supplemental Experimental Procedures).

The resulting data set comprises 384 gene expression profiles, corresponding to the 13 GEMMs each treated with the 13 perturbagens or vehicles. t-SNE analysis revealed that the resulting mouse data set represented an extensive range of expression variability, as required for ARACNe (Figure 2). Specifically, whereas expression profiles from the same GEMMs and perturbagens clustered together, the diverse GEMMs and perturbagens provided independent and highly effective axes to modulate gene expression variability.

ARACNe was run independently on the human and mouse data sets using a conservative mutual information threshold (p ≤ 1.0 × 10^-3, i.e., p ≤ 0.05 Bonferroni corrected for all candidate interactions). This resulted in highly robust regulatory networks—in particular, the “human interactome” represented 249,896 interactions between 2,681 transcriptional regulators and their inferred target genes (Figure 3A; Table S3), whereas the “mouse interactome” represented 222,787 interactions for 2,072 transcriptional regulators (Figure 3A; Table S4).

Analysis of Genome-wide Conservation of Transcriptional Regulatory Pathways
Because it has been previously established that target-by-target analysis may not be optimal to evaluate cross-species interactome conservation (e.g., Zhang et al., 2012), we developed a quantitative metric to compare conservation of the human and mouse interactomes. In particular, we developed a modification of the MARINa algorithm that allows for single-sample analysis to infer the differential activity of all 2,028 transcriptional regulators represented in both interactomes. Analysis was performed on 1,009 expression profiles across the four human data sets (Table S1) and the mouse data set (described herein) to determine whether the inferred activities of each regulator were significantly correlated (p ≤ 0.05), indicating that the murine and human regulatory programs were conserved (Supplemental Experimental Procedures). The accuracy of this metric was evident by comparing two equivalent same-species interactomes from the human and mouse data sets (i.e., positive controls), in which virtually all transcriptional regulators were conserved (>90%), in contrast to randomized interactomes (i.e., negative controls), which had virtually no conservation (Figure 3B).

Using this metric, we found that 70% of the transcriptional regulators in the human and mouse prostate cancer interactomes regulate statistically conserved programs (p ≤ 0.05; Figure 3C; Table S5). Notably, among the conserved transcriptional regulators are many genes important in prostate cancer, such as AR, ETS1, ETV4, ETV5, STAT3, MYC, BRCA1, and NKX3.1 (Shen and Abate-Shen, 2010; Figure 3A; Table S5). In particular, AR displayed extensive correlation of its transcriptional activity between the human and mouse interactomes (Figure 3D), consistent with its known role as a key regulator of prostate development and tumorigenesis (Ryan and Tindall, 2011; Shen and Abate-Shen, 2010).

Cross-Species Computational Analysis Identifies Synergistic Master Regulators of Malignant Prostate Cancer
To identify master regulators (MRs) of malignant prostate cancer (Figure 4), we used the MARINa algorithm, which identifies candidate MRs based on the concerted differential expression of the interactomes. This approach allowed us to uncover a set of MRs that were independently and effectively up-regulated in the human and mouse interactomes, thus providing a framework for understanding the cross-species conservation of key regulatory pathways in malignant prostate cancer.
of their ARACNe-inferred targets (i.e., their inferred differential activity, DA). Specifically, “activated” MRs have positively regulated and repressed targets significantly enriched among over- and underexpressed genes, respectively, while “repressed” MRs have the converse. We interrogated the human prostate cancer interactome using a gene signature representing prostate cancer malignancy derived from the Taylor data set as described elsewhere (Aytes et al., 2013), which compares aggressive prostate tumors (Gleason score $\geq 8$ with rapid biochemical recurrence; $n = 10$) to indolent ones (Gleason score 6 tumors with no biochemical recurrence; $n = 39$). This analysis identified 175 candidate MRs, including 49 activated and 126 repressed ($p < 0.05$; Figure 4A; Table S6).

To investigate the robustness of these MRs, we performed MARIna analysis using an independent malignancy signature derived from the Balk data set as described elsewhere (Aytes et al., 2013), which compares castration-resistant ($n = 29$) with hormone-naive disease ($n = 22$). The MRs identified from the Balk malignancy signature (Table S6) significantly overlapped with those identified from the Taylor malignancy signature (36 in common; Fisher exact test $p < 0.0001$; Table S6). Furthermore, MARIna analyses of 15 independent interactomes revealed that the MRs inferred from two independent prostate cancer interactomes significantly overlapped with those inferred from the Taylor prostate cancer interactome ($p < 7 \times 10^{-3}$ and $p < 8 \times 10^{-20}$, Fisher exact test), whereas the overlap of MRs inferred from 13 nonprostate cancer-specific interactomes were orders of magnitude less significant (Supplemental Experimental Procedures). Thus, MRs of human prostate cancer malignancy are consistent across independent prostate cancer malignancy signatures, but dependent on a context-specific prostate cancer interactome.

To identify corresponding mouse MRs of malignancy, we performed MARIna analysis on four independent GEMM signatures, which embody the diverse range of prostate cancer phenotypes represented among the GEMMs (Figure S2A; Table S2; Supplemental Experimental Procedures). Meta-analysis of the resulting MRs from these independent GEMM signatures led to identification of 229 candidate mouse MRs, including 110 activated and 119 repressed ($p < 0.001$; Figure 4A; Table S7).

The independent list of human and mouse MRs were then integrated to produce a ranked list of 20 conserved MRs, including seven activated and 13 repressed (joint $p$ value: $p \leq 0.0074$ by 2006), which compares castration-resistant ($n = 29$) with hormone-naive disease ($n = 22$). The MRs identified from the Balk malignancy signature (Table S6) significantly overlapped with those identified from the Taylor malignancy signature (36 in common; Fisher exact test $p < 0.0001$; Table S6). Furthermore, MARIna analyses of 15 independent interactomes revealed that the MRs inferred from two independent prostate cancer interactomes significantly overlapped with those inferred from the Taylor prostate cancer interactome ($p < 7 \times 10^{-3}$ and $p < 8 \times 10^{-20}$, Fisher exact test), whereas the overlap of MRs inferred from 13 nonprostate cancer-specific interactomes were orders of magnitude less significant (Supplemental Experimental Procedures). Thus, MRs of human prostate cancer malignancy are consistent across independent prostate cancer malignancy signatures, but dependent on a context-specific prostate cancer interactome.

To identify corresponding mouse MRs of malignancy, we performed MARIna analysis on four independent GEMM signatures, which embody the diverse range of prostate cancer phenotypes represented among the GEMMs (Figure S2A; Table S2; Supplemental Experimental Procedures). Meta-analysis of the resulting MRs from these independent GEMM signatures led to identification of 229 candidate mouse MRs, including 110 activated and 119 repressed ($p < 0.001$; Figure 4A; Table S7).

The independent list of human and mouse MRs were then integrated to produce a ranked list of 20 conserved MRs, including seven activated and 13 repressed (joint $p$ value: $p \leq 0.0074$ by
Notably, these conserved MRs were more likely to be associated with disease outcome than the nonconserved ones, as assessed by a univariate COX proportional hazard regression model (p < 0.05), and were also more likely to be differentially expressed in aggressive prostate tumors (Figures 4A and 4C; Supplemental Experimental Procedures).

We focused our subsequent analysis on the activated conserved MRs, each of which has been associated with cancer-related biological processes: CHAF1A (chromatin activity); FOXM1; CENPF; PSRC1; TSFM; and ASF1B.

See also Figure S2 and Tables S6 and S7.

**Figure 4. Conserved Master Regulators of Malignant Prostate Cancer**

(A) (left) Master regulators (MRs) were identified using human or mouse malignancy signatures; differential activity (DA) is based on enrichment of activated (red) and repressed (blue) targets. DE, differential expression. (Right) Venn diagram showing integration of independent lists of activated MRs from human (49) and mouse (110) with an overlap of seven conserved MRs. Clinical features of all human MRs versus the conserved MRs showing the percentage associated with disease outcome (using a COX proportional hazard model) and the percentage that are differentially expressed in advanced prostate cancer (from Oncomine).

(B) Conserved activated MRs are shown for the human (left) and mouse (right) malignancy signatures, depicting their positive (activated; red bars) and negative (repressed; blue bars) targets. The ranks of differential activity (DA) and differential expression (DE) are shown by the shaded boxes; the numbers indicate the rank of the DE in the malignancy signature.

(C) Summary of conserved MRs showing joint p value from human and mouse MARINa analysis, calculated using Stouffer’s method; p value for COX proportional hazard regression model applied to mRNA expression levels and predicted MR activity; and average p values for differential expression of MRs in metastatic versus nonmetastatic primary tumors.

(D) Computational synergy analysis depicting FOXM1 and CENPF regulons from the human (left) and mouse (right) interactomes showing shared and nonshared targets. Red corresponds to overexpressed targets and blue to underexpressed targets; the p value for the enrichment of shared targets is shown.

Stouffer’s method; Figures 4A and 4B; Figure S2B).
To evaluate their individual and potential synergistic functions in prostate cancer, we silenced FOXM1 and/or CENPF individually or together in four human prostate cell lines, DU145, PC3, LNCaP, and 22Rv1, which have differing tumorigenic properties or together in four human prostate cell lines, DU145, PC3, LNCaP, and 22Rv1, which have differing tumorigenic properties and responses to androgen signaling (Figure 5A; Figure S3A). Notably, each of these cell lines express high levels of FOXM1 and CENPF mRNA; however, LNCaP does not have detectable CENPF protein (Figure 5B; Figures S3B–S3E), and therefore provides an excellent negative control for synergy analysis. To silence FOXM1 and/or CENPF, we engineered doxycycline-inducible lentiviral vectors expressing shRNAs for FOXM1 or CENPF or a control shRNA, as well as an RFP or GFP reporter (Figure 5A; Supplemental Experimental Procedures); analyses were done using two independent shRNA to minimize concerns about off-target effects (Figure S3). We distinguish “synergistic” versus “additive” effects of FOXM1 and CENPF by first extrapolating their “predicted additivity” based on their individual silencing using a log-linear model, and then comparing this predicted value to the “actual” (observed) effect following their co-silencing using a one-sample t test; if the “actual” is statistically greater than the “predicted additive,” we conclude that FOXM1 and CENPF are synergistic rather than additive (Supplemental Experimental Procedures).

Whereas individual silencing of FOXM1, and, to a lesser extent, CENPF, resulted in reduced cellular proliferation, the actual reduction following their co-silencing was statistically greater (p < 0.01; one-sample t test) than the predicted additive, and is therefore synergistic for each cell line that expresses both FOXM1 and CENPF proteins (Figure S3F). Similarly, with respect to colony formation, whereas individual silencing of FOXM1 or CENPF reduced the number of colonies, their co-silencing resulted in nearly complete abrogation of colony formation in each cell line expressing both FOXM1 and CENPF proteins (p < 0.001; one-sample t test; Figures 5C and 5D; Figures S3G and S3H). Importantly, co-silencing of FOXM1 and CENPF was not associated with reduced viability, apoptosis, or further cell cycle arrest relative to their individual silencing (Figures S3I–S3K), suggesting that their observed synergy was not simply due to induction of cell death or was secondary to cell cycle arrest.

To investigate their consequences for tumor growth in vivo, we engrafted DU145 cells expressing silencing vectors for FOXM1 and/or CENPF (or controls) into immunodeficient mice and monitored tumor growth (Figures 5E–5H). Consistent with the cell culture studies, individual silencing of FOXM1 or CENPF resulted in a modest but statistically significant reduction in tumor growth (2-fold, p ≤ 0.002 and 1.5-fold, p ≤ 0.002, respectively), as well as tumor weight (2.3-fold, p ≤ 0.007, and 1.6 fold, p ≤ 0.01, respectively) (Figures 5F and 5G). However, co-silencing of FOXM1 and CENPF resulted in a complete abrogation of tumor growth (10.2-fold reduced, p ≤ 0.000013) and a profound reduction in tumor weight (12.9-fold, p ≤ 0.000011; Figures 5F and 5G). Notably, the actual inhibition of tumor growth following co-silencing of FOXM1 and CENPF was significantly greater than the predicted additive inhibition (3.3-fold difference, p ≤ 0.00026; one sample t test; Figure 5F), supporting the conclusion that FOXM1 and CENPF synergistically regulate tumor growth in vivo.

To further evaluate the synergistic activity of FOXM1 and CENPF for tumor growth, we developed an in vivo competition assay (Figures 5I–5K). Specifically, we infected DU145 cells with silencing vectors expressing an FOXM1 shRNA and an RFP reporter (red) or a CENPF shRNA and an GFP reporter (green), or both lentiviruses (yellow; Figure 5I). As negative controls, we infected DU145 cells with control vectors lacking the FOXM1 or CENPF shRNA but expressing the fluorescent reporters. We then implanted equal numbers of viable red, green, or yellow cells from the experimental or control groups into immunodeficient mice. Following 1 month of growth in vivo, the resulting tumors were isolated and the percentage of red, green, and yellow cells were quantified by fluorescence-activated cell sorting.

Tumors derived from control cells (n = 4) were comprised of equivalent numbers of red (34% ± 0.6%), green (34% ± 2.7%), and yellow (33% ± 1.2%) cells, indicating that the respective lentiviral vectors offer no selective growth advantage (p ≤ 0.614; Hotelling’s one-sample T-squared test; Figures 5J and 5K). In striking contrast, tumors derived from the experimental cells (n = 7) were comprised primarily of green CENPF-silenced cells (57% ± 3.5%) and red FOXM1-silenced cells (41% ± 2.6%), whereas there were virtually no yellow co-silenced cells (2.0% ± 0.3%; Figures 5J and 5K). This profound selection against cells co-silenced for FOXM1 and CENPF was highly significant (p ≤ 0.0001; Hotelling’s one-sample T-squared test; Figure 5K), which further supports the conclusion that FOXM1 and CENPF synergistically regulate tumor growth in vivo.

**FOXM1 and CENPF Coregulate Gene Expression and Control Tumorigenic Signaling Pathways in Prostate Cancer**

To investigate the mechanism(s) underlying the observed activities of FOXM1 and CENPF, we assessed the consequences of their individual versus co-silencing for expression of their ARACNe-inferred common (shared) target genes (Table S3).
Figure 5. Functional Validation of FOXM1 and CENPF
(A) Human prostate cancer cells were infected with lentiviral silencing vectors expressing shRNA for FOXM1 and/or CENPF (or control) and either an RFP (red) or GFP (green) reporter. Unless otherwise indicated, analyses were done using two independent shRNAs for each gene and in four independent prostate cancer cell lines (DU145, PC3, LNCaP, 22Rv1); in most cases data using shRNA1 are shown.
(B) Western blot analysis showing expression of FOXM1 or CENPF proteins in DU145 cells with the indicated shRNAs.
(C and D) Colony formation assay. (C) Representative analyses of DU145 cells with an shRNA for FOXM1 or CENPF (or the control) with colonies visualized using crystal violet. (D) Quantification of colonies using ImageJ.
(E–H) Analysis of tumor growth in vivo. (E) DU145 cells expressing an shRNA for FOXM1 and/or CENPF, or the control, were implanted subcutaneously into mouse hosts. Beginning on day 7, mice were administered doxycycline to induce shRNA expression and tumor growth was monitored for 1 month. (F) Tumor growth curves for the indicated shRNA. The dashed line shows the predicted additive effect of cosilencing FOXM1 and/or CENPF. (G) Tumor weights at the time of sacrifice. (H) Representative tumors. In (D), (F), and (G) the predicted additive was estimated based on the consequences of individual silencing of FOXM1 and CENPF using a log-linear model; the p value, calculated using a one-sample t test, indicates the significance between the predicted additive versus the actual (observed) consequences of cosilencing FOXM1 and CENPF.
(I–K) In vivo competition assay. (I) Equal numbers of DU145 cells expressing the control shRNA (control cells), or the experimental shRNA for FOXM1 and/or CENPF (experimental cells) as well as RFP or GFP were implanted into mouse hosts. Beginning on day 7, mice were administered doxycycline to induce shRNA expression and tumor growth was monitored for 1 month, then tumors were collected and fluorescence-activated cell sorting was performed to quantify the total number of red, green, or yellow cells in individual tumors for control and experimental groups. (J) Representative fluorescence-activated cell sorting plots showing the percentage of red, green, or yellow cells relative to the total number of fluorescent cells. (K) Graphs show the average percent of red, green, and yellow cells in the control tumors (n = 4) or experimental tumors (n = 7); p values correspond to a Hotelling’s one-sample t test. (Bottom) Representative tumors. Error bars represent ± SD.
See also Figure S3.
Although target gene expression was somewhat reduced by their individual silencing, cosilencing of FOXM1 and CENPF produced a significantly greater reduction for the majority of targets, consistent with coregulation of target gene expression by FOXM1 and CENPF (Figure 6A; Figure S4). Furthermore, using chromatin immunoprecipitation (ChIP) followed by quantitative
individually silenced or co-silenced (Figure 6D; Table S8). The expression profiles from prostate cancer cells in which they were targets (p = 0.0028 for enrichment of FOXM1 targets; p ≤ 0.001 for CENPF targets), further confirming the accuracy of the ARACNe analysis (Table S8). Inspection of these differentially expressed genes, as well as gene set enrichment analysis of enriched biological pathways confirmed the known individual functions of FOXM1 and CENPF as regulators of cellular proliferation and/or mitosis (Tables S8 and S9).

However, cosilencing of FOXM1 and CENPF revealed an additional repertoire of significantly differentially expressed genes and enriched biological pathways (Figures 6D and 6E; Tables S8 and S9), including several pathways associated with tumorigenesis: “cell cycle” (normalized enrichment score [NES] 1.32; p ≤ 0.001), “stress pathway” (NES 1.58; p ≤ 0.01), “regulation of insulin-like growth factor” (NES 1.89; p ≤ 0.001), “signaling by NGF” (NES 1.25; p ≤ 0.001), “metabolism of amino acids” (NES 1.25; p ≤ 0.01), “PI3-Akt signaling” (NES 1.89; p ≤ 0.001), “MAP kinase pathway” (NES 1.34; p ≤ 0.008), “telomere maintenance” (NES 1.35; p ≤ 0.01), and “cell adhesion molecules” (NES 1.32; p ≤ 0.001).

Notable was the enrichment of PI3-kinase and MAP kinase signaling pathways following cosilencing of FOXM1 and CENPF (Figure 6E; Table S9) because these constitute established hallmarks of aggressive prostate cancer (Aytes et al., 2013; Taylor et al., 2010). As evident by western blot analysis, both pathways are completely abrogated following cosilencing of FOXM1 and CENPF (88% and 90%, respectively; n = 53 informative cases; Figure 6C). Notably, subcellular colocalization of FOXM1 and CENPF in prostate cancer cells may be effective for inactivation of these signaling pathways.

Co-Expression of FOXM1 and CENPF Is a Prognostic Indicator for Human Prostate Cancer
We next asked whether expression of FOXM1 and/or CENPF is associated with cancer progression and/or outcome by analysis of tissue microarrays (TMAs; Figure 7A; Table S1). In particular, we analyzed a high-density TMA containing primary tumors from a large cohort of patients (n = 916) that had undergone prostatectomy at Memorial Sloan-Kettering Cancer Center (MSKCC) from 1985 to 2003 (Donovan et al., 2008). These cases have extensive clinical follow-up data for up to 20 years, including time to biochemical recurrence, prostate cancer-specific survival, and time to metastasis (Table S1). We also evaluated a second TMA from the rapid autopsy program at the University of Michigan, which contains prostate cancer metastases (n = 60), including 6 lung, 11 liver, 22 lymph node, and 14 other sites (Shah et al., 2004).

Analysis of the MSKCC prostatectomy TMA revealed that FOXM1 and CENPF were overexpressed in 33% and 37% of all cases, respectively, (n = 821 informative cases) with a trend toward increased expression in tumors with higher Gleason scores (Figure 7A; Figure S5A). Furthermore, analysis of the Michigan metastasis TMA revealed that FOXM1 and CENPF were expressed in most of the prostate cancer metastases (88% and 90%, respectively; n = 53 informative cases; Figure 7A). Moreover, FOXM1 and CENPF were frequently coexpressed and colocalized in the nucleus in both the MSKCC prostatectomy TMA (Spearman’s Rho = 0.57, p ≤ 2 × 10^{-16}) and the Michigan metastasis TMA (Spearman’s Rho = 0.43, p ≤ 1 × 10^{-3}; Figure 7A). Additionally, at the mRNA level, overexpression of FOXM1 and CENPF was well correlated in advanced prostate cancer and metastases from independent cohorts of human prostate cancer (Figure S5B).

To determine whether expression of FOXM1 and/or CENPF is associated with disease outcome, we first defined four groups of patients from the MSKCC TMA based on their expression levels: (1) low/normal expression of both FOXM1 and CENPF (n = 418); (2) high expression of FOXM1 and low/normal expression of CENPF (n = 47); (3) high expression of CENPF and low/normal expression of FOXM1 (n = 133); and (4) high expression of both FOXM1 and CENPF (n = 173). Kaplan-Meier survival analysis revealed that patients having elevated expression of both FOXM1 and CENPF had the worst outcome for three independent clinical endpoints, namely, time to biochemical-free recurrence (p ≤ 4.4 × 10^{-5}), death due to prostate cancer (p ≤ 5.9 × 10^{-5}), and time to metastasis (p ≤ 1.0 × 10^{-7}; Figure 7B). Notably, subcellular colocalization of FOXM1 and CENPF in prostate tumors was also associated with the worst outcome for all three independent clinical endpoints (Figure S5A). In contrast, elevated expression of only FOXM1 or CENPF was either not significant or marginally significant for biochemical recurrence and prostate-specific survival (p ≤ 0.053 and p ≤ 0.011 for FOXM1, respectively; p ≤ 0.078 and p ≤ 0.402 for CENPF, respectively), and was 10 to 13 orders of magnitude less significant for time to metastasis (p ≤ 0.001 for FOXM1 and p ≤ 3.1 × 10^{-6} for CENPF; Figure 7B).

We independently corroborated the association of FOXM1 and CENPF with disease outcome in two independent human prostate cancer data sets that had not been used for training purposes elsewhere in this study; namely, the Glinsky data set, in which biochemical recurrence is the clinical endpoint (Glinsky et al., 2004), and the Sboner data set, in which the clinical endpoint is prostate cancer-specific overall survival (Sboner et al., 2010; Table S1). Using these independent cohorts, we evaluated the mRNA expression levels of FOXM1 and CENPF as well as their MARINa-inferred transcriptional activity...
We then performed Kaplan-Meier survival analysis comparing four patient groups: (1) those with low inferred activity or expression for FOXM1 and CENPF; (2) those with high inferred activity or expression only for FOXM1; (3) those with high inferred activity or expression only for CENPF; and (4) those with high inferred activity or expression for both FOXM1 and CENPF. Similar to our analysis of the TMA, patients with high inferred activity or mRNA expression for both CENPF and FOXM1 were associated with the worst outcome in both cohorts, as measured by biochemical recurrence (p ≤ 0.000065) and prostate cancer-specific survival (p ≤ 0.000040; Figures 7C and S5C). Notably, these findings reveal that their MARINa-inferred activities are well correlated with the actual protein expression of FOXM1 and CENPF, and further demonstrate the striking association of their co-expression/co-activity with poor disease outcome.

Finally, association of FOXM1 and CENPF protein expression with disease outcome using C-statistics revealed its robust prognostic value for disease-specific survival (C = 0.71; CI 0.59–0.84, p ≤ 0.00024), as well as time to metastasis (C = 0.77;
CI 0.71–0.83, p ≤ 3.0 × 10^{-19}, Figure 7D). Notably, co-expression of FOXM1 and CENPF proteins dramatically improves prognosis over Gleason score alone for both disease-specific survival (C = 0.86; CI 0.80–0.93, p ≤ 1.0 × 10^{-10}; p value for improvement p ≤ 0.00020) and time to metastasis (C = 0.86; CI 0.81–0.89, p ≤ 6.5 × 10^{-58}; p value for improvement, p ≤ 5.3 × 10^{-13}, Figure 7D). Taken together, these analyses of independent clinical cohorts and using distinct statistical models demonstrate that co-expression of FOXM1 and CENPF is a robust prognostic indicator of poor disease outcome and metastasis.

**DISCUSSION**

Recent advances in whole-genome analyses are providing an increasingly high-resolution view of the multitude of genetic, genomic, and epigenetic alterations associated with cancer phenotypes. Given the staggering number of potential interactions, identification of the true causal drivers and essential synergistic interactions represents a considerable challenge. In the current study, we have demonstrated that cross-species interrogation of genome-wide context-specific regulatory networks can address this challenge by dramatically winnowing the candidate gene interactions that implement the regulatory programs underlying cancer phenotypes. In particular, we have introduced a comprehensive systems approach to interrogate complementary regulatory networks for human and mouse prostate cancer to identify conserved causal regulators and to elucidate synergistic interactions among them. These studies have led to the identification of FOXM1 and CENPF as synergistic master regulators of prostate cancer malignancy and robust prognostic biomarkers of aggressive prostate cancer. We propose that this overall approach for genome-wide cross-species analysis will be generally applicable for identifying synergistic interactions that drive physiologic and pathologic phenotypes in cancer and other diseases.

The genome-wide assembly and cross-species interrogation of human and mouse prostate cancer regulatory networks represents a major conceptual advance. A critical requirement was the generation of a mouse prostate cancer interactome from a data set of appropriate size and expression heterogeneity. We incorporated the diversity afforded by genetically and phenotypically distinct mouse models of prostate cancer obtained through a community effort (see Supplemental Experimental Procedures), in combination with exogenous perturbations administered to each mouse model. This strategy has led to the successful construction of a genome-wide, context-specific mouse interactome for the study of prostate cancer, and is generalizable for the generation of interactomes for a wide range of physiologic and pathologic phenotypes.

A second critical requirement was the development of an informative algorithm to determine whether the human and mouse prostate cancer interactomes represented conserved regulatory programs, thus enabling accurate and robust cross-species integrative analysis. Toward this end, we introduced a metric for quantitative assessment of conservation of regulatory networks, which revealed that the large majority of regulatory programs represented by these networks are highly conserved (>70%). Although the current study is focused on prostate cancer interactomes assembled using ARACNe, this general approach for evaluating conservation can be used for cross-species analyses of regulatory networks for other cancers or other diseases and can be readily adapted for analyses of networks inferred using alternative algorithms, such as those based on the Context Likelihood of Relatedness and Bayesian-networks algorithms (Akavia et al., 2010; Faith et al., 2007). Indeed, we envision that the ability to quantitatively evaluate conservation of cross-species regulatory programs will be broadly applicable for other physiological and pathological comparisons, and particularly beneficial for accurate integration of preclinical findings from genetically engineered mice to human clinical trials.

A third critical requirement for the success of our approach was our ability to effectively mine these cross-species regulatory networks to identify conserved master regulators of cancer malignancy and to identify their synergistic interactions. The MARINa algorithm used for these analyses infers “master regulators” as genes that most significantly regulate the transcriptional program associated with a particular phenotype (in this case, prostate cancer malignancy), and hence are rate-limiting drivers of the phenotype (Carro et al., 2010; Lefebvre et al., 2010). Notably, MARINa also provides an effective computational tool for analyses of synergistic interactions among master regulators (Carro et al., 2010); indeed, our unbiased interrogation of ~2,000 transcriptional regulators represented in the interactomes led to identification of a single synergistic pair, namely FOXM1 and CENPF. The power of this approach suggests that it may be of general value in dissecting polygenic dependencies in cancer and other diseases.

Although both FOXM1 and CENPF have been implicated in various cancers, our study has uncovered their unexpected synergistic interaction. FOXM1 encodes a forkhead domain transcription factor that is frequently overexpressed in many different types of cancer, including prostate (Alvarez-Fernandez and Medema, 2013; Halasi and Gartel, 2013a; Kalin et al., 2011; Koo et al., 2012). Many previous studies have established a role for FOXM1 in regulation of cellular proliferation, DNA damage, genomic stability, drug resistance, and metastasis, and have shown that it interacts with other key regulators such as β-catenin and MYB (Lefebvre et al., 2010; Zhang et al., 2011). In prostate cancer, gain- or loss-of-function of FOXM1 in vivo have been shown to elicit modest effects on tumor growth (Cai et al., 2013; Kalin et al., 2006).

CENPF (also known as mitosin or LEK1 in mouse), a known target of FOXM1, has also been implicated in various cancers, although not previously in prostate, and in some cases has been shown to undergo gene amplification and to be associated with disease outcome (Ma et al., 2006; Varis et al., 2006). However, the actual functional role of CENPF has been more elusive and difficult to reconcile. In particular, whereas CENPF is named for its association with the centromere-kinetochore protein complex, such association is transient and, in fact, CENPF has other functions, including regulation of mitosis and cellular proliferation (Bormont et al., 2005; Feng et al., 2006; Holt et al., 2005), which are mediated in part by protein interactions (Ma et al., 2006; Varis et al., 2006).

Thus, although the individual functions of FOXM1 and CENPF in cancer had been well studied, their synergistic interaction
could not have been anticipated from previous analyses. Cumulatively, our findings suggest that co-expression of FOXM1 and CENPF in prostate cancer leads to coregulation of transcriptional programs, which ultimately result in activation of the key signaling pathways associated with prostate cancer malignancy, including the PI3K and MAPK signaling pathways. Because FOXM1 and CENPF can each be targeted pharmacologically (Halasi and Gartel, 2013b; Pan and Yeung, 2005; Radhakrishnan et al., 2006), their inhibition may provide an effective means of treating advanced prostate cancer; indeed, therapeutic targeting of FOXM1 and CENPF may help overcome the complex feedback mechanisms that have hindered therapeutic targeting of PI3K and MAPK signaling pathways.

Furthermore, we envision that by using alternative gene signatures that represent other prostate cancer phenotypes, genome-wide cross-species analysis of master regulators and their potential synergistic interactions may help to define molecular subtypes of prostate cancer, which have thus far been elusive. More broadly, our general approach to elucidate conserved and functionally relevant gene interactions can be applied to many tumor contexts as well as other human diseases for which appropriate model systems are available.

EXPERIMENTAL PROCEDURES

Assembly of Interactomes and Master Regulator Analyses
Expression profile data sets for human prostate cancer are described in Table S1. GEMMs are described in Table S2 and their representative histopathology shown in Figure S1A. A description of perturbagen treatments is provided in the Supplemental Experimental Procedures. Human and mouse interactomes were assembled using the ARACNe algorithm (Margolin et al., 2006a). Details of the resulting human and mouse networks are provided in Tables S3 and S4, respectively. Analysis of cross-species network conservation was done using a modification of the MARINa algorithm described in the Supplemental Experimental Procedures. Master regulator analysis and computational synergy analysis were performed using MARINa (Carro et al., 2010; Lefebvre et al., 2010). Master regulators for the human and mouse interactomes are provided in Tables S6 and S7, respectively.

Functional Validation
Gene silencing of FOXM1 and CENPF as well as forced expression of FOXM1 were done using lentiviral shRNAs or expression vectors (Open Biosystems and CCSB Human ORFeome Library, respectively). Human cancer cell lines used for functional studies were obtained from ATCC. All experiments using animals were performed according to protocols approved by the Institutional Animal Care and Use Committee at Columbia University Medical Center.

Tissue Microarray Analyses
All studies involving human subjects were approved by the Institutional Review Board of MSKCC or University of Michigan. TMAs were constructed with approval from the Human Biospecimen Utilization Committee and Institutional Review Board; consent was obtained from all patients, as required. Analysis of protein expression of FOXM1 and CENPF was performed using a high-density TMA (Donovan et al., 2013) and a metastasis TMA (Shah et al., 2004). Available clinicopathological features of these TMAs are summarized in Table S1.

Statistical Methods
Statistical analysis was performed with survcomp package using R v2.14.0. Cox proportional hazard model was estimated with the surv and coxph functions. Kaplan-Meier survival analysis was performed using surv, survfit, and survdiff functions. Concordance indexes (c-index) were estimated and compared using coxph and concordance.index (counting ties) and cindex.comp functions. Details of all statistical analyses and all computational procedures are provided in the Supplemental Experimental Procedures. An executable SWEAVE document and R data objects are deposited in Figshare at http://dx.doi.org/10.6084/m9.figshare.928353.

ACCESSION NUMBERS

The Gene Expression Omnibus accession number for the mouse expression profiles reported in this paper is GSE53202.

SUPPLEMENTAL INFORMATION

Supplemental information includes Supplemental Experimental Procedures, five figures, and nine tables and can be found with this article online at http://dx.doi.org/10.1016/j.ccr.2014.03.017.

AUTHOR CONTRIBUTIONS

A.A., A.M., M.J.A., M.M.S., A.C., and C.A.-S. designed experiments, analyzed results, and wrote the manuscript; C.L. contributed to initial interactome assembly and analyses; M.C.-M. provided pathological expertise; T.Z. provided statistical expertise; and J.A.E., A.G., and K.L.P. provided critical reagents.

ACKNOWLEDGMENTS

We are indebted to our colleagues who generously provided genetically engineered mouse models for assembly of the mouse prostate cancer interactome, namely Drs. Bart Williams (Van Andel), Terry Van Dyke (NCI-Frederick), Barbara Foster (Roswell Park Cancer Center), Yu Chen (MSKCC), and Charles Sawyers (MSKCC). We are also indebted to Dr. Victor Reuter (MSKCC) for generously sharing the prostatectomy tissue microarray. We thank Jose Silva and members of his laboratory for invaluable help with establishment of the shRNA validation studies and Drs. Adolfo Ferrando and Ricardo Dalla Favera for critical reading of the manuscript.

This work was supported by grants CA084294 (to C.A.S., M.M.S., and A.C.), U54 CA121852 (to A.C., C.A.S., and M.M.S.), CA154293 (to M.M.S. and C.A.S.), DK076602 (to M.M.S.), Cancer Target Discovery and Development Centers NCI U01 CA168426 (to A.C.), the Michigan Center for Translational Pathology, SPORE grant P50 CA69568 (to J.K.P.), and an award from the V-Foundation for Cancer Research (to A.C.S.). A.A. was a recipient of a Marie Curie International Outgoing Fellowship (PIOF-GA-2009-253290), cosponsored by the Catalan Institute of Oncology-Bellvitge Institute for Biomedical Research, Barcelona, Spain. A.M. is a recipient of a Prostate Cancer Foundation Young Investigator Award. C.A.S. is an American Cancer Society Research Professor supported in part by a generous gift from the F.M. Kirby Foundation.

Received: August 31, 2013
Revised: January 9, 2014
Accepted: March 14, 2014
Published: May 12, 2014

REFERENCES

Cancer Cell

Genome-wide Regulatory Networks of Prostate Cancer


Cancer Cell 25, 638–651, May 12, 2014 ©2014 Elsevier Inc. 651