

# Mechanism-centric regulatory network identifies NME2 and MYC programs as markers of Enzalutamide resistance in CRPC

---

Received: 13 August 2022

---

Accepted: 22 December 2023

---

Published online: 08 January 2024

---

 Check for updates

---

Sukanya Panja<sup>1,8</sup>, Mihai Ioan Truica<sup>2,8</sup>, Christina Y. Yu<sup>1</sup>, Vamshi Saggurthi<sup>1</sup>, Michael W. Craige<sup>1</sup>, Katie Whitehead<sup>1</sup>, Mayra V. Tuiche<sup>1,3</sup>, Aymen Al-Saadi<sup>4</sup>, Riddhi Vyas<sup>1</sup>, Shridar Ganesan<sup>5</sup>, Suril Gohel<sup>1</sup>, Frederick Coffman<sup>1</sup>, James S. Parrott<sup>1</sup>, Songhua Quan<sup>2</sup>, Shantenu Jha<sup>4</sup>, Isaac Kim<sup>5,6</sup>, Edward Schaeffer<sup>2</sup>, Vishal Kothari<sup>2,9</sup>✉, Sarki A. Abdulkadir<sup>2,7,9</sup>✉ & Antonina Mitrofanova<sup>1,5,9</sup>✉

---

Heterogeneous response to Enzalutamide, a second-generation androgen receptor signaling inhibitor, is a central problem in castration-resistant prostate cancer (CRPC) management. Genome-wide systems investigation of mechanisms that govern Enzalutamide resistance promise to elucidate markers of heterogeneous treatment response and salvage therapies for CRPC patients. Focusing on the de novo role of MYC as a marker of Enzalutamide resistance, here we reconstruct a CRPC-specific mechanism-centric regulatory network, connecting molecular pathways with their upstream transcriptional regulatory programs. Mining this network with signatures of Enzalutamide response identifies NME2 as an upstream regulatory partner of MYC in CRPC and demonstrates that NME2-MYC increased activities can predict patients at risk of resistance to Enzalutamide, independent of co-variables. Furthermore, our experimental investigations demonstrate that targeting MYC and its partner NME2 is beneficial in Enzalutamide-resistant conditions and could provide an effective strategy for patients at risk of Enzalutamide resistance and/or for patients who failed Enzalutamide treatment.

The seminal discovery of the role of the androgen receptor (AR) in prostate tumor growth and progression nominated androgen-deprivation therapy (ADT) as the current mainstay for the treatment of advanced prostate cancer (PCa)<sup>1–3</sup>. Despite an initial response to ADT in ~80% of patients<sup>4</sup>, relapse and progression to castration-resistant prostate cancer (CRPC) are often inevitable. Interestingly, CRPC tumors retain a high level of AR signaling<sup>5</sup>, thus allowing CRPC

patients to benefit from more potent second-generation AR targeting by androgen receptor-signaling inhibitors (ARSIs)<sup>6–11</sup>, with Enzalutamide (Enza) as one of the most commonly used ARSIs<sup>9</sup>.

Enzalutamide is an AR signaling inhibitor that blocks binding of androgen to AR in addition to inhibiting AR nuclear translocation and its binding to DNA<sup>12</sup>. Administration of Enzalutamide has been shown to improve patient survival overall, yet response to Enzalutamide is

---

<sup>1</sup>Department of Health Informatics, Rutgers School of Health Professions, Newark, NJ 07107, USA. <sup>2</sup>Department of Urology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA. <sup>3</sup>Rutgers Biomedical and Health Sciences, Rutgers School of Graduate Studies, Newark, NJ 07039, USA.

<sup>4</sup>Department of Electrical and Computer Engineering, Rutgers School of Engineering, New Brunswick, NJ 08854, USA. <sup>5</sup>Rutgers Cancer Institute of New Jersey, New Brunswick, NJ 08901, USA. <sup>6</sup>Department of Urology, Yale School of Medicine, New Haven, CT 06510, USA. <sup>7</sup>Robert H. Lurie Comprehensive Cancer Center, Chicago, IL 60611, USA. <sup>8</sup>These authors contributed equally: Sukanya Panja, Mihai Ioan Truica. <sup>9</sup>These authors jointly supervised this work: Vishal Kothari, Sarki A. Abdulkadir, Antonina Mitrofanova. ✉e-mail: [vkothari@piantrx.com](mailto:vkothari@piantrx.com); [sarki.abdulkadir@northwestern.edu](mailto:sarki.abdulkadir@northwestern.edu); [amitrofa@shp.rutgers.edu](mailto:amitrofa@shp.rutgers.edu)

heterogeneous with nearly half of CRPC patients either not responding and/or developing resistance to Enzalutamide within 8 to 18 months<sup>12–17</sup>. Thus, prioritization of patients based on their risk of developing resistance to Enzalutamide could provide an effective avenue for a personalized line of treatment and potential improvement in overall PCa management.

Recently, several groups have tackled mechanisms involved in response to Enzalutamide, including investigation of the chromatin remodeling protein *CHD1*<sup>18</sup>, TGF- $\beta$  signaling<sup>19</sup>, the Persist gene signature<sup>20</sup> (which contains a cohort of genes involved in cell proliferation, cell cycle regulation, stemness, chromatin remodeling and reorganization, and DNA repair), Wnt/ $\beta$ -catenin pathway<sup>21</sup>, stemness programs<sup>12</sup> etc. These discoveries provide substantial advances in uncovering mechanisms involved in Enzalutamide response, yet their clinical utility and potential therapeutic intervention remain to be further investigated.

Since MYC (i.e., c-MYC) has been known to play a central role in PCa development and metastasis<sup>3,22–28</sup>, is overexpressed in 60% of CRPC patients<sup>27,29</sup>, has been shown to be associated with general ARSI response<sup>27</sup>, and could potentially be therapeutically targeted<sup>30,31</sup>, we sought to investigate MYC mechanisms for their role in response to Enzalutamide.

In this work, we developed TR-2-PATH algorithm to reconstruct a CRPC-specific mechanism-centric regulatory network using the Stand Up to Cancer (SU2C) East Coast cohort<sup>32,33</sup>, which identified a network of transcriptional regulatory programs (comprised of transcriptional regulators and their target genes) upstream of the MYC pathway. Querying this network with signatures of favorable and poor Enzalutamide response nominated the NME2 transcriptional regulatory program as the most significant upstream regulatory partner of MYC in CRPC Enzalutamide-associated conditions. We demonstrated consistent association of MYC pathway and NME2 transcriptional regulatory activities across multiple CRPC patient cohorts and showed that increased activity levels of MYC pathway and NME2 transcriptional regulatory program could be utilized as markers to identify primary resistance to Enzalutamide, independent of clinical and molecular variables. Further, we evaluated MYC and NME2 partnership in response to Abiraterone (a second-generation AR signaling inhibitor of androgen biosynthesis)<sup>8,33,34</sup> and did not observe association to the risk of resistance to Abiraterone. Functional studies further demonstrated that therapeutic targeting of MYC using the small-molecule MYC inhibitor MYCi975 and concurrent NME2 knock-down could reverse Enzalutamide resistance. Thus, we propose that MYC-associated mechanisms could be utilized to identify CRPC patients

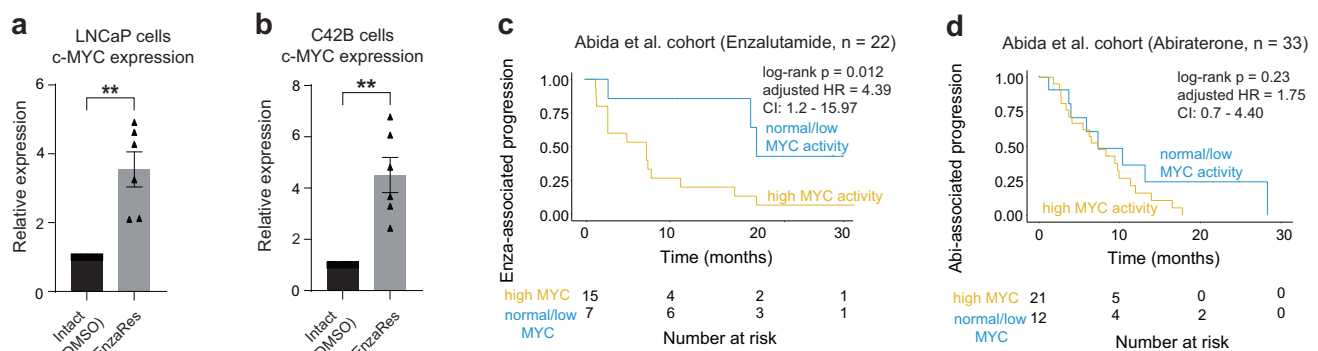
that are at risk of developing primary resistance to Enzalutamide and that therapeutic targeting of these mechanisms could provide an effective strategy for patients at risk of Enzalutamide resistance and/or for patients who failed Enzalutamide treatment.

## Results

### Increased MYC activity is associated with the risk of Enzalutamide resistance in CRPC patients

We observed that the levels of MYC are increased in Enzalutamide-resistant conditions (treatment with 20  $\mu$ M Enzalutamide for up to 3 months, EnzaRes) compared to control (DMSO) conditions in LNCaP (a cell line derived from prostate cancer metastasis to lymph-node<sup>35</sup>, commonly used to study Enzalutamide resistance) and C42B (LNCaP metastatic CRPC derivative) cell lines (Fig. 1a, b, one-tailed Welch t-test  $p$  value = 0.002 and  $p$  value = 0.0018 for LNCaP and C42B cells, respectively, see Methods), which were accompanied by increased levels of AR (Supplementary Fig. 1a, b, one-tailed Welch t-test  $p$  value = 0.011 and  $p$  value = 0.001 for LNCaP and C42B cells, respectively, see Methods).

To evaluate if this observation could be translated to human patients, we tested if high MYC pathway activity was characteristic of CRPC patients at risk of developing resistance to Enzalutamide. For this, we utilized RNA-seq profiles of CRPC patients from ref. 33, specifically selecting samples from CRPC patients that did not receive any ARSI treatment prior to sample collection (i.e., ARSI-naïve). We further selected patients that after biopsy (i.e., sample collection), were treated with Enzalutamide and monitored for Enzalutamide-associated disease progression (see Methods, Supplementary Data 1,  $n$  = 22). We subjected this patient cohort (referred to as Enzalutamide-specific Abida et al. cohort hereafter) to single-sample Gene Set Enrichment Analysis (GSEA)<sup>36</sup> to estimate activity levels of MYC pathway (i.e., HALLMARK\_MYC\_TARGETS\_V2 from Hallmark collection in MSigDB 3.0,  $n$  = 58) in each patient, that were then subjected to unsupervised clustering (see Methods), separating patients with high MYC pathway activity (Fig. 1c, yellow,  $n$  = 15) and normal/low MYC pathway activity (Fig. 1c, blue,  $n$  = 7). We then compared Enzalutamide-associated disease progression (which was defined in ref. 33 as the time on Enzalutamide without being subjected to another agent, such as taxane) between these groups using Kaplan-Meier survival analysis<sup>37</sup> and Cox proportional hazards modeling<sup>38</sup>, which demonstrated a significant difference between patients from high MYC and normal/low MYC pathway activity groups (Fig. 1c, log-rank  $p$  value = 0.012, adjusted HR (hazard ratio) = 4.39, CI (confidence interval): 1.2–15.97, see Methods), indicating that increased MYC pathway activity is characteristic for

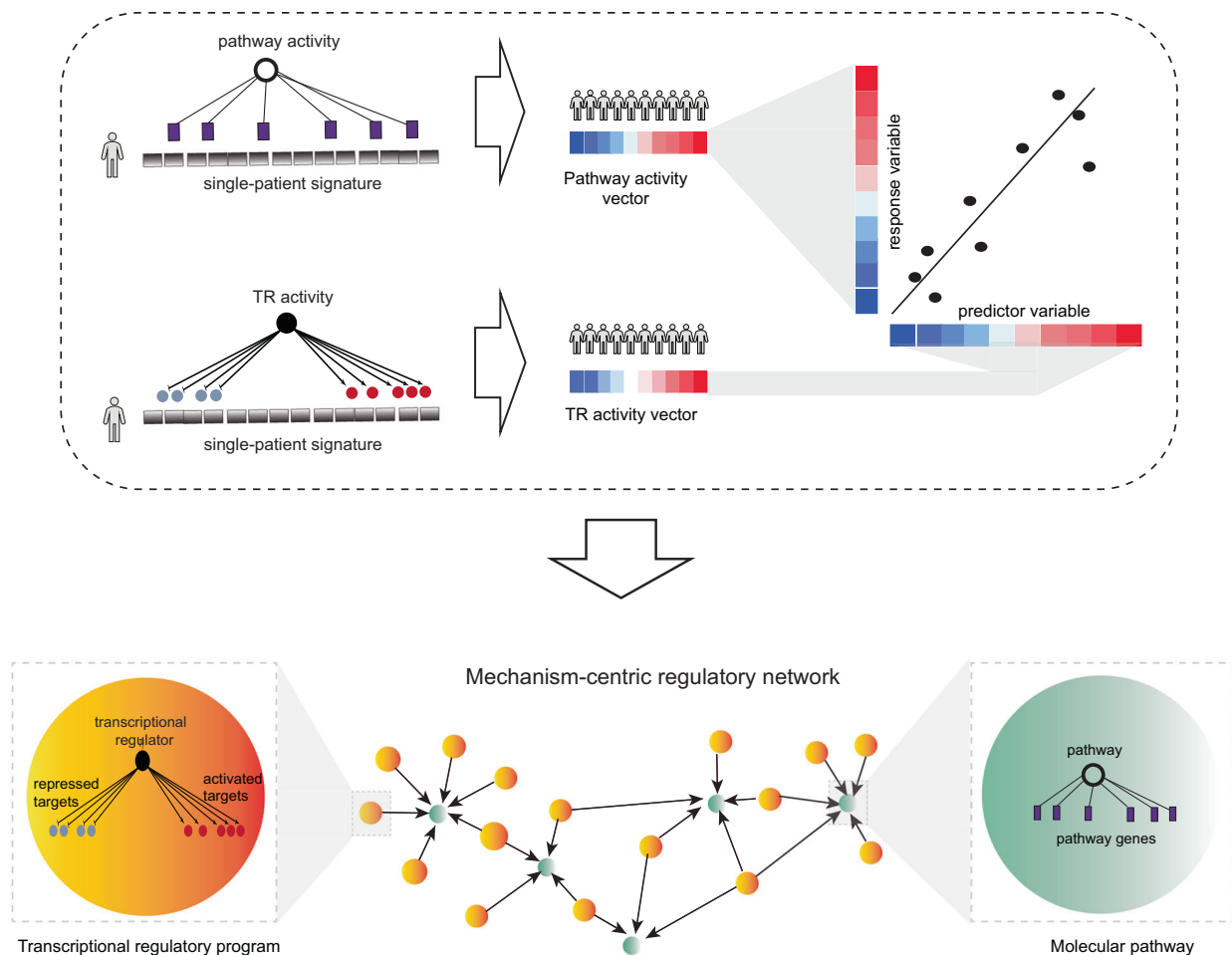


**Fig. 1 | MYC pathway activity is specific for predicting response to Enzalutamide in CRPC patients.** c-MYC expression in Intact (DMSO treated) and Enzalutamide-resistant (EnzaRes) (a) LNCaP and (b) C42B cells as shown using qRT-PCR.  $P$  values were estimated using a one-tailed Welch t-test. Data are presented as mean values  $\pm$  SEM from  $n$  = 6 independent biological replicates. \*\* $p$  value  $\leq$  0.01. Source data are provided as a Source Data file. Kaplan-Meier survival analysis

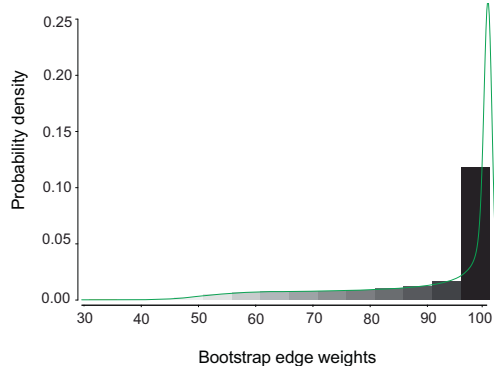
comparing CRPC patients that received (c) Enzalutamide or (d) Abiraterone after sample collection from the Abida et al. cohort with high (yellow) and normal/low (blue) MYC pathway activity levels. Log-rank  $p$  value, adjusted HR (hazard ratio), and CI (confidence interval) are indicated. Source data are provided as a Source Data file.

a

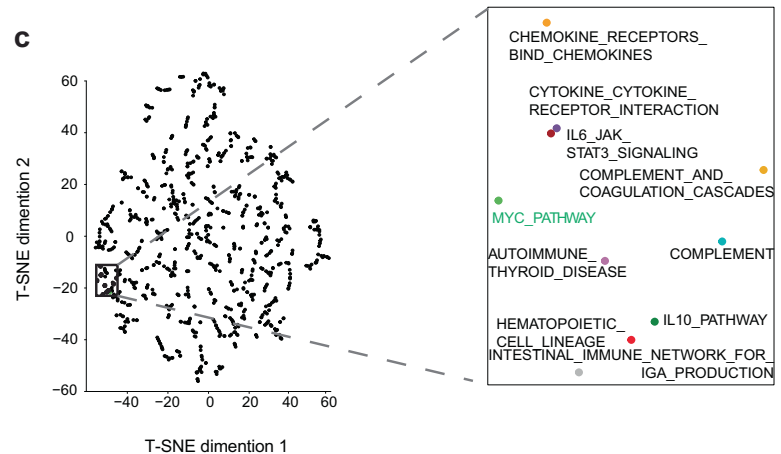
## TR-2-PATH mechanism-centric network reconstruction



b



c



patients with a higher risk of resistance to Enzalutamide. The patient group with high MYC pathway activity also demonstrated increased levels of AR expression and activity (estimated as in ref. 39, see Methods) (Supplementary Fig. 1c, d, one-tailed Welch t-test  $p$  value = 0.002 and  $p$  value = 0.01, for AR expression and AR activity, respectively) and significant correlation was observed between MYC pathway activity and AR expression/activity levels in the Enzalutamide-specific

Abida et al. cohort<sup>33</sup> ( $n = 22$ ), as described above (Spearman correlation  $\rho = 0.482$ ,  $p$  value = 0.024; and Spearman correlation  $\rho = 0.484$ ,  $p$  value = 0.023, for AR expression and AR activity, respectively). To ensure that the correlation between MYC pathway and AR transcriptional regulatory program is not due to their compositional similarity, we evaluated overlap between MYC pathway genes ( $n = 58$ , Supplementary Data 2) and AR transcriptional regulon ( $n = 231$ ,

**Fig. 2 | Reconstruction of a mechanism-centric regulatory network for CRPC patients.** **a** Schematic representation of the TR-2-PATH workflow. (First row) Single-patient pathway enrichment analysis and single-patient transcriptional regulatory activity identifies pathway activity vector and transcriptional regulatory activity vector respectively, pairs of which are then subjected to linear regression analysis to reconstruct a mechanism-centric regulatory network. (Second row) In the network, transcriptional regulatory programs are represented as orange nodes and molecular pathways as green nodes. An edge (black arrow) illustrates that a

significant relationship was defined between a transcriptional regulatory program and molecular pathway. **b** Distribution of edge weights across the network, as defined by the bootstrap analysis. The x-axis corresponds to the edge weight and the y-axis to its frequency (probability). **c** (Left) t-SNE clustering of molecular pathways (dots), based on the weights of their incoming edges. (Right) Pathways around MYC are shown as a zoom-in and MYC pathway is shown in green. Source data are provided as a Source Data file.

Supplementary Data 2), which demonstrated negligible similarity between these two mechanisms (common genes = 2, Jaccard similarity<sup>40</sup> index = 0.007 out of 1).

Interestingly, in Abiraterone-specific Abida et al. cohort<sup>33</sup>, we did not observe a correlation of MYC pathway activity with Abiraterone response. In particular, we utilized RNA-seq profiles of CRPC patients from Abida et al. cohort<sup>33</sup>, selecting patients that did not receive any ARSI treatment prior to sample collection and then sub-selecting patients that after biopsy (i.e., sample collection) were treated with Abiraterone and monitored for Abiraterone-associated disease progression (which was defined in ref. 33 as the time on Abiraterone without being subjected to another agent, see Methods, Supplementary Data 1,  $n = 33$ , referred to as Abiraterone-specific Abida et al. cohort). We identified no significant difference in Abiraterone-associated disease progression between patients from high MYC and normal/low MYC groups (Fig. 1d, log-rank  $p$  value = 0.23, adjusted HR = 1.75, CI: 0.7–4.40, see Methods) in the Abiraterone-specific Abida et al. cohort.

Taken together, these analyses demonstrate that increased activity of MYC pathway could potentially serve as a marker to stratify patients for their risk of developing resistance to Enzalutamide and would benefit from validation in larger patient cohorts for future clinical use.

### Defining MYC upstream regulatory programs through mechanism-centric network analysis

To elucidate MYC regulation implicated in Enzalutamide resistance and define potential additional axes for salvage therapeutics, we investigated transcriptional regulatory mechanisms upstream of MYC pathway that might affect MYC while also governing Enzalutamide resistance. For this, we developed “TR-2-PATH” algorithm to reconstruct a CRPC-specific mechanism-centric regulatory network, which connects molecular pathways (Fig. 2a, green nodes) with potential upstream transcriptional regulatory (TR) programs (Fig. 2a, orange nodes). Nodes in this mechanism-centric network do not correspond to individual genes or alterations, but rather represent mechanisms: such as transcriptional regulatory programs or molecular pathways, promising to uncover complexity underlying therapeutic resistance and identify more effective biomarkers and optimized therapeutic interventions (Fig. 2a). Our objective was to reconstruct a mechanism-centric regulatory network that would capture a wide-array of CRPC-specific phenotypes, constituting a valuable resource for the community and allowing effective utilization across different CRPC-related questions (e.g., primary and secondary therapeutic resistance, metastases to different sites etc.) For this, we utilized RNA-seq profiles from CRPC patients in the Stand Up to Cancer (SU2C) East Coast cohort<sup>32,33</sup>, excluding repeated samples and samples from ref. 33 (see Methods, Supplementary Data 1,  $n = 153$ ). The selected SU2C East Coast cohort was well-suited for CRPC mechanism-centric network reconstruction as it (i) constitutes one of the largest-to-date cohorts of CRPC patients, essential for statistical learning/inference; (ii) is characterized by wide-ranged age ( $59.2 \pm 8.38$  years) and prostate specific antigen (PSA) levels ( $234.5 \pm 1574.4$  ng/ml); (iii) includes different metastatic sites, including bone ( $n = 39$ ), liver ( $n = 26$ ), lymph-node ( $n = 57$ ), prostate ( $n = 4$ ), lung ( $n = 2$ ), adrenal ( $n = 1$ ), other soft tissue ( $n = 19$ ), etc.; and (iv) represents different stages of therapeutic intervention,

including samples from patients previously exposed to ARSIs (including Enzalutamide and Abiraterone,  $n = 67$ ), ARSI-naïve at the time of sample collection ( $n = 75$ ), currently on treatment ( $n = 4$ ), etc.; all together capturing a wide range of clinical variables necessary for accurate statistical inference.

For network reconstruction, to evaluate relationships between molecular pathways and their upstream transcriptional regulatory programs, we first needed to estimate pathway activity levels and transcriptional regulator activity levels in each sample in the SU2C East Coast cohort (Fig. 2a). To estimate pathway activity levels, we performed single-sample GSEA<sup>36</sup> on the scaled SU2C East Coast cohort, so that each sample ( $n = 153$ ) was used as a reference signature and each molecular pathway ( $n = 883$ , from KEGG<sup>41</sup>, BioCarta<sup>42</sup>, Reactome<sup>43</sup>, and Hallmark<sup>44</sup> collections) was used as a query (see Methods). To estimate activity of TRs in the SU2C East Coast cohort, we performed VIPER analysis<sup>45</sup> on the scaled SU2C East Coast cohort (as above) utilizing each sample ( $n = 153$ ) as a reference and transcriptional regulatory programs ( $n = 2678$ ) from a prostate cancer specific interactome<sup>39</sup> as a query (see Methods). For each molecular pathway, its activity level (defined as Normalized Enrichment Scores, NES) across all patients in the cohort then defined a “pathway activity vector” (Fig. 2a, top, see Methods). Similarly, for each transcriptional regulatory program, its activity level across all patients in the cohort defined a “TR activity vector” (Fig. 2a, bottom, see Methods). All pairs of TR-pathway activity vectors were then subjected to linear regression analysis<sup>46</sup> (see Methods), where “TR activity vector” was utilized as a predictor (independent) variable and “pathway activity vector” was used as a response (dependent) variable, with an objective to identify TR programs whose changes could potentially explain changes in the activity of molecular pathways in CRPC-specific manner. Significant relationships, corrected for multiple hypotheses testing (see Methods), between TRs and pathways (both positive and negative) were then considered for network reconstruction (Fig. 2a).

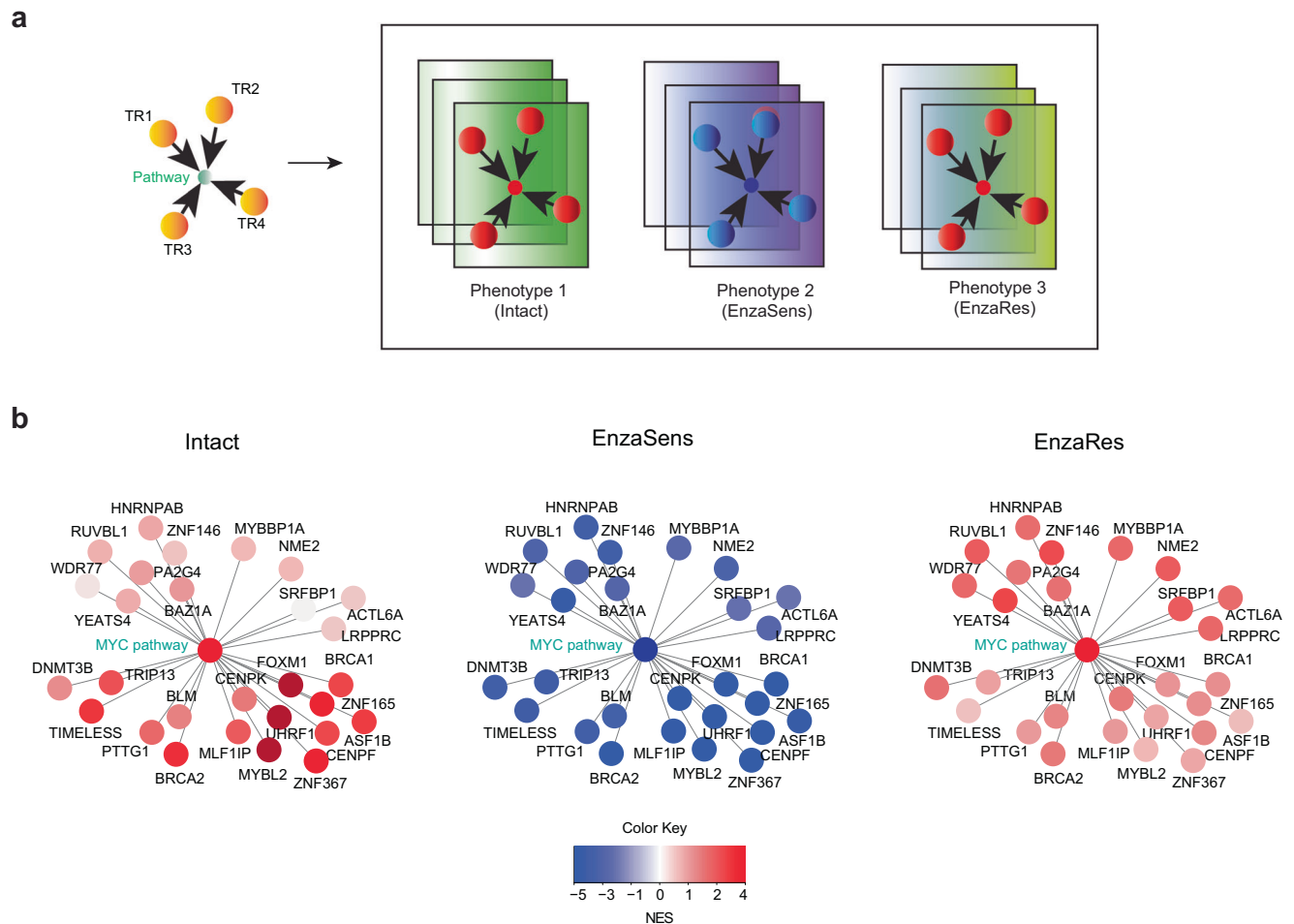
To ensure that the network is robust to experimental and sampling noise, we enhanced our network reconstruction with bootstrap analysis. For this, patients from the SU2C East Coast cohort ( $n = 153$ ) were sampled with replacement (see Methods,  $k = 100$ ) and each bootstrap was subjected to TR-2-PATH network reconstruction. A comparison of edge distributions across the 100 bootstrapped networks showed their similarity (Supplementary Fig. 2), indicating the method’s overall reproducibility. A total of 100 bootstrapped networks were then used to assign “weight” to each edge in the network reconstructed from the whole dataset ( $n = 153$ ), reflecting the number of times an edge appears (and thus could be recovered) across the bootstrapped networks (see Methods, Supplementary Data 3, Fig. 2b). Unsupervised t-distributed stochastic neighbor embedding clustering<sup>47</sup> (t-SNE) was utilized to cluster molecular pathways based on weights of their incoming edges, demonstrating co-clustering of MYC pathway with Chemokine<sup>48</sup>, Cytokine<sup>49</sup>, IL-6<sup>50</sup>, JAK STAT 3 signaling<sup>50</sup>, and IgA<sup>51</sup> pathways (see Methods, Fig. 2c), and revealing their potential cross-talk in the CRPC setting<sup>52</sup>.

### Network mining I: identifying differentially altered sub-networks

The next step was to utilize this network to identify TR programs upstream of MYC pathway that are involved in Enzalutamide response



## Network mining I: identifying differentially altered sub-networks



**Fig. 3 | Network mining I: identifying upstream transcriptional regulatory programs that affect MYC pathway and are associated with response to Enzalutamide.** **a** Schematic representation of the changes in activity levels of molecular pathways and their upstream transcriptional regulatory programs (i.e., sub-networks) as they transition from Intact (treated with DMSO) to Enzalutamide-sensitive (EnzaSens) to Enzalutamide-resistant (EnzaRes) phenotypes. Red depicts

up-regulation and blue depicts down-regulation of TR and pathway activities. **b** Identified upstream transcriptional regulatory programs (MYC-centered sub-network) associated with Enzalutamide treatment affecting MYC pathway, depicted across Intact, EnzaSens, and EnzaRes phenotypes. Red depicts up-regulation and blue depicts down-regulation of TR and pathway activities. Source data are provided as a Source Data file.

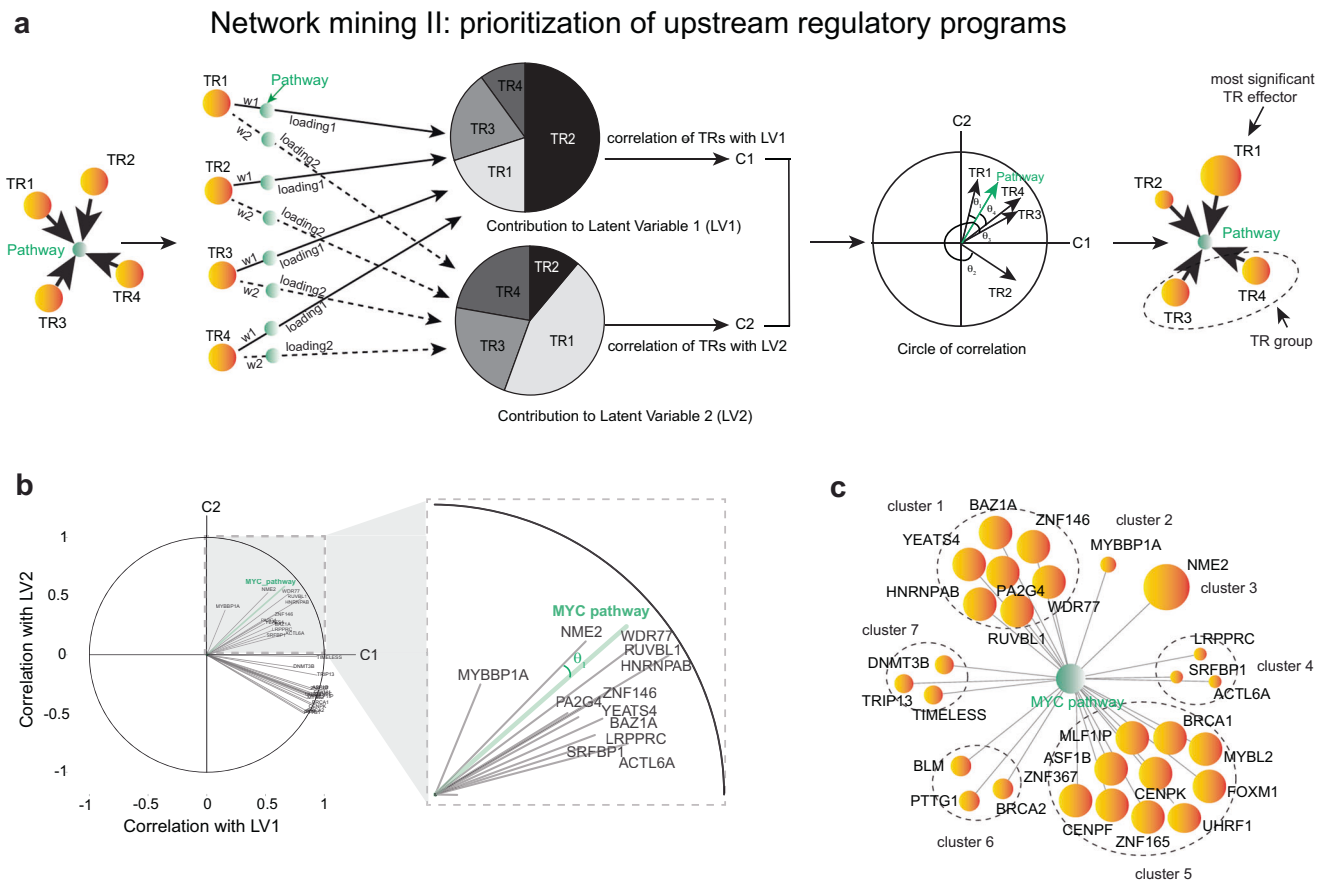
and resistance. To achieve this, we aimed to identify parts (sub-networks) of the mechanism-centric network that significantly change (alter) between phenotypes of interest, in our case—phenotypes that describe response to Enzalutamide (Fig. 3a). To accurately capture response and resistance to Enzalutamide in a controlled setting, we utilized gene expression profiles from ref. 53 (see Methods, Supplementary Data 1,  $n = 12$ ), which is based on the LNCaP experimental system that has been widely used to study Enzalutamide-resistance previously<sup>9,54–56</sup>. These profiles included (i) LNCaP parental intact samples subjected to DMSO (phenotype 1 - Intact, Fig. 3a); (ii) LNCaP samples treated for 48 h with Enzalutamide, where their survival and proliferation were sensitive to Enzalutamide (phenotype 2 - EnzaSens, Fig. 3a); and (iii) LNCaP samples treated with Enzalutamide for 6 months, where their survival and proliferation did not depend on Enzalutamide (phenotype 3 - EnzaRes, Fig. 3a).

We hypothesized that regulatory programs that are active in the intact state, then are repressed by Enzalutamide treatment (EnzaSens phenotype) and further re-activated as Enzalutamide resistance develops (EnzaRes phenotype) would be effective candidates to uncover mechanisms that govern Enzalutamide-resistance (Fig. 3a). Such network mining (using pairwise phenotype comparison, see Methods, Supplementary Fig. 3a, b, Supplementary Data 4A–F)

identified TR mechanisms ( $n = 28$ ) upstream of the MYC pathway, which constitutes of MYC-centric TR sub-network with a significant “active→repressed→reactivated” activity changes across the Enzalutamide-related phenotypes (Fig. 3b).

### Network mining II: Prioritization of upstream regulatory programs

The next essential step was to prioritize the identified transcriptional regulatory programs upstream of a pathway of interest (e.g., MYC pathway) for experimental validation and potential salvage therapeutic targeting. We developed such a prioritization step to overcome several important drawbacks, commonly present in widely utilized statistical analyses. First of all, our method considers potential multi-collinearity among input variables (TRs), which is often naturally present in biological systems, yet can substantially obstruct statistical learning. In fact, variance inflation factor (VIF) analysis<sup>57</sup> of the 28 TR programs identified upstream of MYC demonstrated significant multi-collinearity (all TRs had  $VIF > 10$ , a multi-collinearity threshold suggested in refs. 58,59) (see Methods, Supplementary Fig. 4) and thus requires special methods to avoid information loss or model misinterpretation. Commonly utilized methods for handling multi-collinearity (e.g., regularization techniques), often keep one of the



**Fig. 4 | Network mining II: NME2 is prioritized as TR with the most significant effect on MYC pathway.** **a** Schematic representation of the PLS-inspired approach to prioritize TR programs, based on their effect on a molecular pathway of interest. (Left) TR activity vectors are utilized as inputs, which are then regressed on a pathway to identify non-collinear latent variables (pie charts), which include a linear combination of TR programs, based on their effect on the pathway (slices in each pie). (Middle) These latent variables are utilized to build a circle of correlation, which depicts the relationship between each latent variable and each TR and

pathway. (Right) Effect scores are defined to group and prioritize TRs, based on their effect on a pathway. **b** A circle of correlation is utilized to determine the degree of closeness between TR programs and the MYC pathway, based on their effect on each latent variable. **c** Grouping and prioritization of the MYC upstream TR programs. Circle sizes correspond to the TR effect scores. NME2 is determined to have the most significant effect on MYC pathway. Source data are provided as a Source Data file.

collinear variables in the model and eliminate others at random, thus limiting biological interpretability and translatability of the model. Our technique keeps all input variables intact, instead identifying their potential groups (based on their effect on the pathway of interest) and preventing information loss. Furthermore, our prioritization method not only tests for direct regulatory relationships but also considers that a meaningful regulatory relationship can exist between entities that do not necessarily have direct (but rather indirect) interactions, which are widely present in biological systems, potentially including relationships between TRs and biological pathways.

To overcome these limitations, we developed a prioritization step, inspired by the Partial Least Squares (PLS) approach<sup>60</sup>, which has been mostly utilized in social sciences<sup>61–64</sup> (sometimes referred to as a “supervised PCA”) but has not been used for network-based mining in oncology to date. Briefly, to identify TR groups and prioritize the effect of the TR programs on the MYC pathway, our approach considers TR activity vectors as predictor (input) variables and utilizes MYC pathway activity vector as a response (output) variable (see Methods, Fig. 4a, left). It then regresses TR activity vectors on the MYC pathway vector so that a linear combination of TRs defines a latent variable (see Methods, Fig. 4a, left). This latent variable is then “subtracted” from the TR activity vectors, leaving the residuals to be utilized for defining the next latent variable (see Methods). Identified latent variables do not express collinearity or multi-collinearity and are utilized as axes to

build a “circle of correlation” (see Methods, Fig. 4a, middle). Such a circle of correlation depicts the association of TR programs and the MYC pathway (defined as arrows on the circle of correlation, see Methods) to each latent variable. We then developed a method that utilized unsupervised hierarchical clustering on the degrees of closeness (angles) between TR and pathway arrows (see Methods) so that TRs in high proximity to one another (thus having similar effects on latent variables) are grouped as they express simultaneous effect on the MYC pathway (see Methods, Fig. 4a, right). Such TR groups/clusters (which also include groups with one TR) are then “prioritized” based on their effect on the MYC pathway activity (see Methods, Fig. 4a, right) using effect scores, which are defined as a combination of (i) degree of closeness between a TR group/cluster and the MYC pathway on the circle of correlation (i.e., angle between their arrows), which reflects effect of each TR group activity changes on MYC pathway; (ii) association (i.e., Pearson correlation) between a TR group/cluster and each evaluated latent variable, which reflects contribution of each TR group to each latent variable; and (iii) edge weight between TR group/cluster and the MYC pathway from the TR-2-PATH mechanism-centric network reconstruction step, which reflects robustness of their regulatory relationship (see Methods, Fig. 4b).

This approach identified 7 TR groups/clusters, based on their effect on the MYC pathway activity (two of the clusters had single TRs, Fig. 4c): group/cluster 1 (HNRNPAB, YEATS4, BAZ1A, ZNF146, WDR77,

RUVBL1, and PA2G4), group/cluster 2 (MYBBP1A), group/cluster 3 (NME2), group/cluster 4 (ACTL6A, LRPPRC and SRFBP1), group/cluster 5 (FOXMI, MYBL2, BRCA1, MLF1IP, ASF1B, ZNF367, CENPF, ZNF165, CENPK, and UHRF1), group/cluster 6 (BRCA2, PTTG1, and BLM), and group/cluster 7 (TIMELESS, TRIP13, and DNMT3B) (Fig. 4c). Each group/cluster was then assigned an effect score, with group/cluster 3 (NME2)<sup>65</sup> having the highest effect (score) on the MYC pathway (Fig. 4c, Supplementary Fig. 5a, b, Supplementary Data 5). While our analysis nominated NME2 transcriptional regulatory program to have the highest effect on MYC pathway in Enzalutamide-associated CRPC context, it has also been previously shown to bind to the MYC promoter region and upregulate MYC transcription<sup>44,66</sup>, suggesting that further investigation of this relationship may uncover aspects of the transcriptional regulatory mechanisms governing the MYC pathway which could potentially provide an additional axis for therapeutic targeting for CRPC patients.

### Validation in clinical cohorts

We next sought to confirm and evaluate if activation of NME2 TR program and MYC pathway are present in patients at risk of resistance to Enzalutamide and if they can be used as markers to risk-stratify patients prior to Enzalutamide administration. First, to evaluate if activity of the MYC pathway and NME2 TR program are high in treatment-naïve patients, who are at risk of developing resistance to Enzalutamide, we have evaluated single-cell profiles from two sequential samples from a CRPC patient (O1115655) that eventually failed Enzalutamide: neoadjuvant sample (prior to Enzalutamide treatment, Fig. 5a) and adjuvant sample (after developing resistance to Enzalutamide, i.e., EnzaRes, Fig. 5b) from ref. 19 (see Methods, Supplementary Data 1). After subjecting single-cell transcriptomic data to unsupervised uniform manifold approximation and projection (UMAP) clustering<sup>67</sup> (see Methods), to identify adenocarcinoma cells among the cell populations we assessed activity levels of AR, alongside expression of CK8 and CD45 (Supplementary Fig. 6a–d). Following adenocarcinoma identification, we evaluated activity levels of the MYC pathway and NME2 TR program in both neoadjuvant and adjuvant samples. Our analysis indicated significantly higher levels of both NME2 TR activity and MYC pathway activity in adenocarcinoma cells, compared to other cells in Enzalutamide-naïve (neoadjuvant, Fig. 5a, one-tailed Welch t-test  $p$  value < 2.26E-16 for NME2 and  $p$  value = 1.74E-7 for MYC, see Methods) and Enzalutamide-resistant, i.e., EnzaRes (adjuvant, Fig. 5b, one-tailed Welch t-test  $p$  value < 2.26E-16 for NME2 and  $p$  value < 2.26E-16 for MYC, see Methods) samples, indicating that (i) both high-MYC pathway and high-NME2 TR activity levels were present prior to treatment in a patient who was at risk of developing subsequent resistance to Enzalutamide, nominating them as markers to identify patients at potential risk of Enzalutamide resistance; and (ii) both high-MYC pathway and high-NME2 TR activity levels were also observed after resistance to Enzalutamide developed (similar to our observation in LNCaP and C42B cell lines, Fig. 1a, b, Supplementary Fig. 5a, b), cautiously nominating a MYC-centered salvage therapeutic line for patients that fail Enzalutamide.

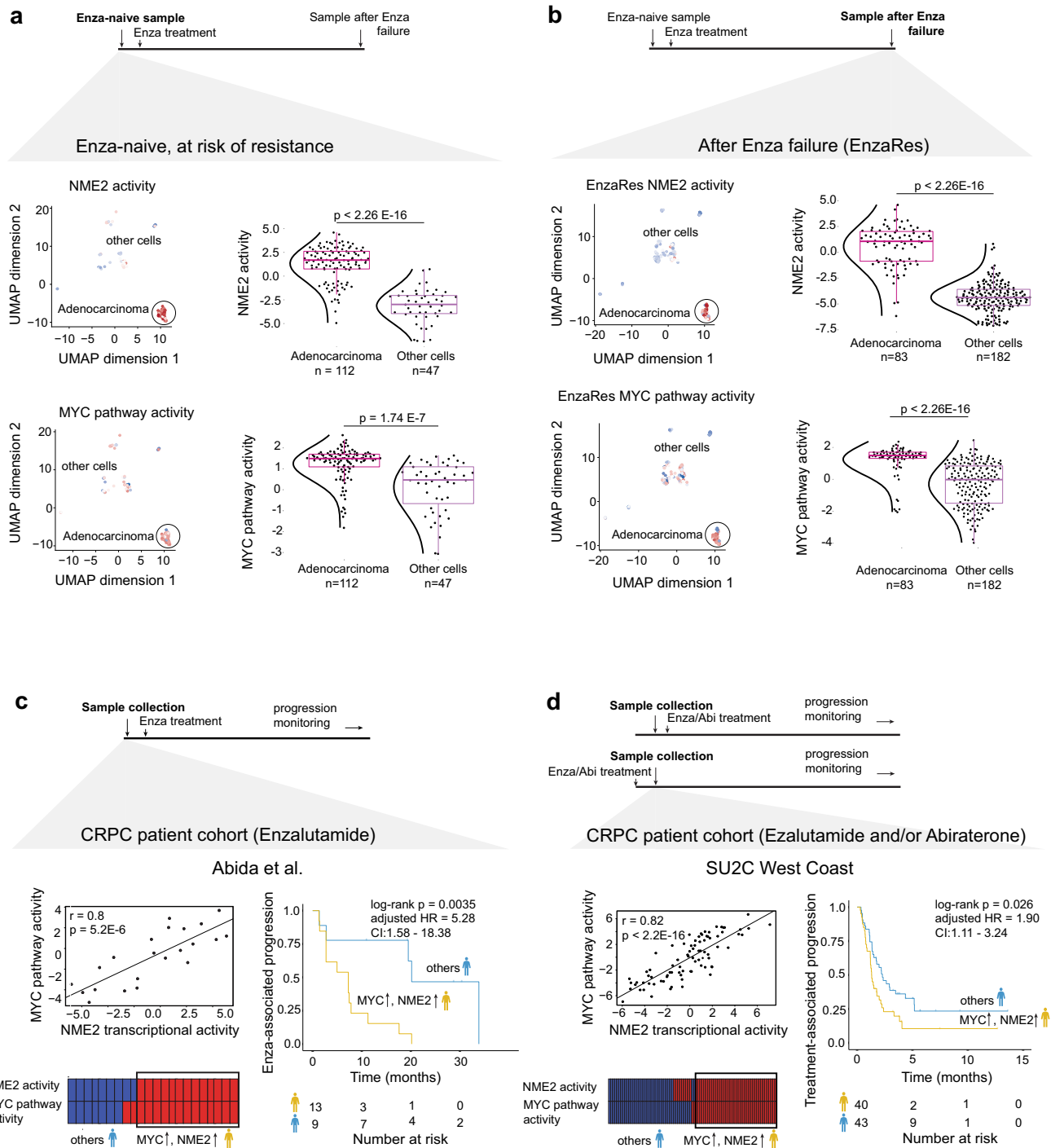
Given increased activity levels of both NME2 TR and MYC pathway in a single-cell sample from a treatment-naïve patient that later developed resistance to Enzalutamide, we sought to confirm their collective ability to predict CRPC patients at the treatment-naïve stage for risk of developing primary resistance to Enzalutamide using the Abida et al.<sup>33</sup> cohort, which was also utilized in Fig. 1c (see Methods, Supplementary Data 1,  $n = 22$ ). As previously described, we used Enzalutamide-specific Abida et al. cohort<sup>33</sup> (i.e., ARSI-naïve CRPC patients that were later subjected to Enzalutamide and monitored for Enzalutamide-associated disease progression) to estimate NME2 TR and MYC pathway activity levels in each patient (see Methods). The NME2 TR and MYC pathway demonstrated

concordance of their activity levels while maintaining largely distinct regulatory programs (overlap between NME2 transcriptional regulon  $n = 412$  and MYC pathway  $n = 58$  is equal to 7 genes, Jaccard similarity index = 0.01 out of 1, Supplementary Data 6) (Fig. 5c, left top, Pearson  $r = 0.8$ ,  $p$  value = 5.2E-6, see Methods). The association between NME2 transcriptional activity and MYC pathway activity was also confirmed in refs. 12 ( $n = 24$ ),<sup>68</sup> (CRPC samples  $n = 34$ ) and<sup>69</sup> ( $n = 157$ ) CRPC patient cohorts (Alumkal et al. Pearson  $r = 0.74$ ,  $p$  value = 3.53E-5; Beltran et al. Pearson  $r = 0.86$ ,  $p$  value = 5.13 E-11; Kumar et al. Pearson  $r = 0.65$ ,  $p$  value < 2.2E-16, Supplementary Fig. 7, Supplementary Data 1, see Methods).

Next, we subjected the activity levels of the NME2 TR and MYC pathway in Enzalutamide-specific Abida et al.<sup>33</sup> cohort to unsupervised clustering (see Methods) that identified (i) patients with both high levels of NME2 transcriptional activity and MYC pathway activity (Fig. 5c, left bottom,  $n = 13$ , yellow group) and (ii) patients with at least one low/normal NME2 transcriptional activity and/or MYC pathway activity, categorized as “others” (Fig. 5c, left bottom,  $n = 9$ , blue group). A comparison of these groups using Kaplan-Meier survival analysis<sup>37</sup> and Cox proportional hazards model analysis<sup>38</sup>, demonstrated a significant difference in their Enzalutamide-associated disease progression (Fig. 5c, right, log-rank  $p$  value = 0.0035, adjusted HR = 5.28, CI = 1.58–18.38, see Methods).

To extend our validation studies to an additional patient cohort, we utilized SU2C West Coast cohort<sup>70,71</sup> (see Methods, Supplementary Data 1,  $n = 83$ ) which comprises of CRPC patients that were subjected to Enzalutamide and/or Abiraterone (~67% of patients in this cohort were pre- or post-treated with Enzalutamide, either alone or in combination with Abiraterone) either before or after biopsy (i.e., before or after sample collection) and were subsequently monitored for treatment-associated disease progression (which was defined in refs. 70,71 as increase in PSA level (minimum 2 ng/mL) that has risen at least twice, in an interval of at least one week or soft tissue progression (nodal and visceral) based on RECIST v1.1). For each of these CRPC patients, we first estimated their NME2 TR and MYC pathway activity levels (see Methods) followed by evaluating association between NME2 TR and MYC pathway activity. Our analysis demonstrated concordance between NME2 TR and MYC pathway activity (Fig. 5d, left top, Pearson  $r = 0.82$ ,  $p$  value < 2.2E-16, see Methods). Next, we subjected the activity levels of the NME2 TR and MYC pathway to unsupervised clustering (see Methods) that identified (i) patients with both high NME2 TR and MYC pathway activity levels (Fig. 5d, left bottom,  $n = 40$ , yellow group) and (ii) patients with at least one low/normal NME2 TR and/or MYC activity levels, categorized as “others” (Fig. 5d, left bottom,  $n = 43$ , blue group). A comparison of these groups using Kaplan-Meier survival analysis<sup>37</sup> and Cox proportional hazards model analysis<sup>38</sup> demonstrated a significant difference in treatment-associated disease progression (Fig. 5d, right, log-rank  $p$  value = 0.026, adjusted HR = 1.90, CI = 1.11–3.24, see Methods), which supports our previous observations.

To evaluate association of NME2 TR and MYC pathway with response to Abiraterone, we analyzed RNA-seq profiles from two Abiraterone-specific cohorts: (i) ARSI-naïve CRPC patients from ref. 33 that were subjected to Abiraterone after sample collection and monitored for Abiraterone-associated disease progression, as in Fig. 1d (see Methods, Supplementary Data 1,  $n = 33$ ) and (ii) PROMOTE<sup>34</sup> cohort, which is comprised of ARSI-naïve CRPC patients, subjected to Abiraterone for 12 weeks after sample collection and then assessed for disease progression (binary outcomes, where disease progression was defined based on the combined score that included serum PSA level, bone and CT imaging and symptom assessments at week 12, see Methods, Supplementary Data 1,  $n = 77$ ). In both cohorts, we estimated the NME2 TR and MYC pathway activity in each sample and subjected them to similar analyses as above. Kaplan-Meier survival analysis<sup>37</sup> and Cox proportional hazards model analysis<sup>38</sup> on the Abida et al.<sup>33</sup> cohort

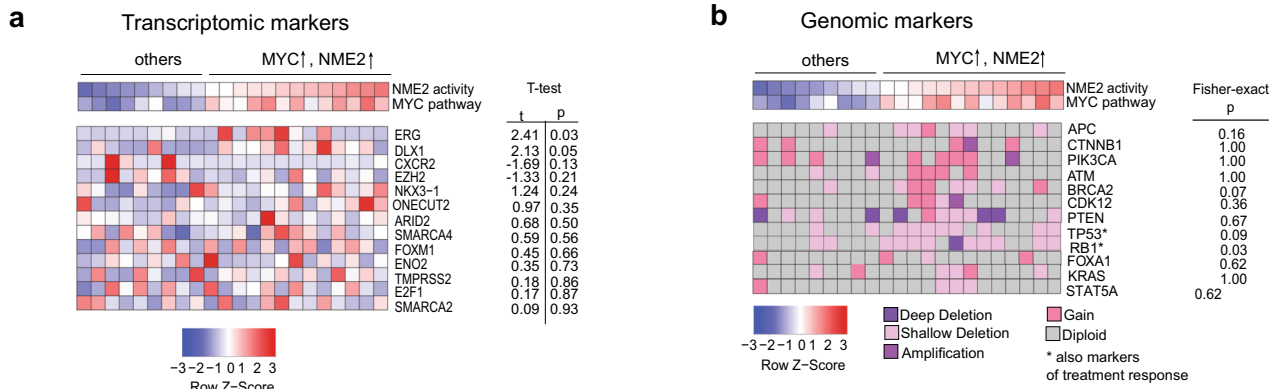


**Fig. 5 | Activities of MYC and NME2 are associated with poor response to Enzalutamide.** **a, b** Comparing NME2 TR and MYC pathway activity between adenocarcinoma cells (dark magenta) and other cells (dark orchid) in neoadjuvant and adjuvant samples of a CRPC patient obtained from He et al.  $P$  value was estimated using a one-tailed Welch t-test. In boxplots, the center line corresponds to the median and the box limits correspond to the first and third quartiles (the 25th and 75th percentiles). The upper and lower whiskers extend to the maximum or minimum value within 1.5 times the interquartile range, respectively. **c** (Left, top) Pearson correlation analysis between NME2 TR and MYC pathway activity in the Abida et al. cohort subjected to adjuvant Enzalutamide. Pearson  $r$  and  $p$  value are indicated. (Left, bottom) Patients with high-MYC and high-NME2 (yellow) and the rest of the patients (blue) were identified. (Right) Kaplan-Meier survival analysis, comparing high-MYC and high-NME2 group (yellow) to the rest of the patients (blue) in SU2C West Coast cohort. Log-rank  $p$  value, adjusted HR (hazard ratio), and CI (confidence interval) are indicated.

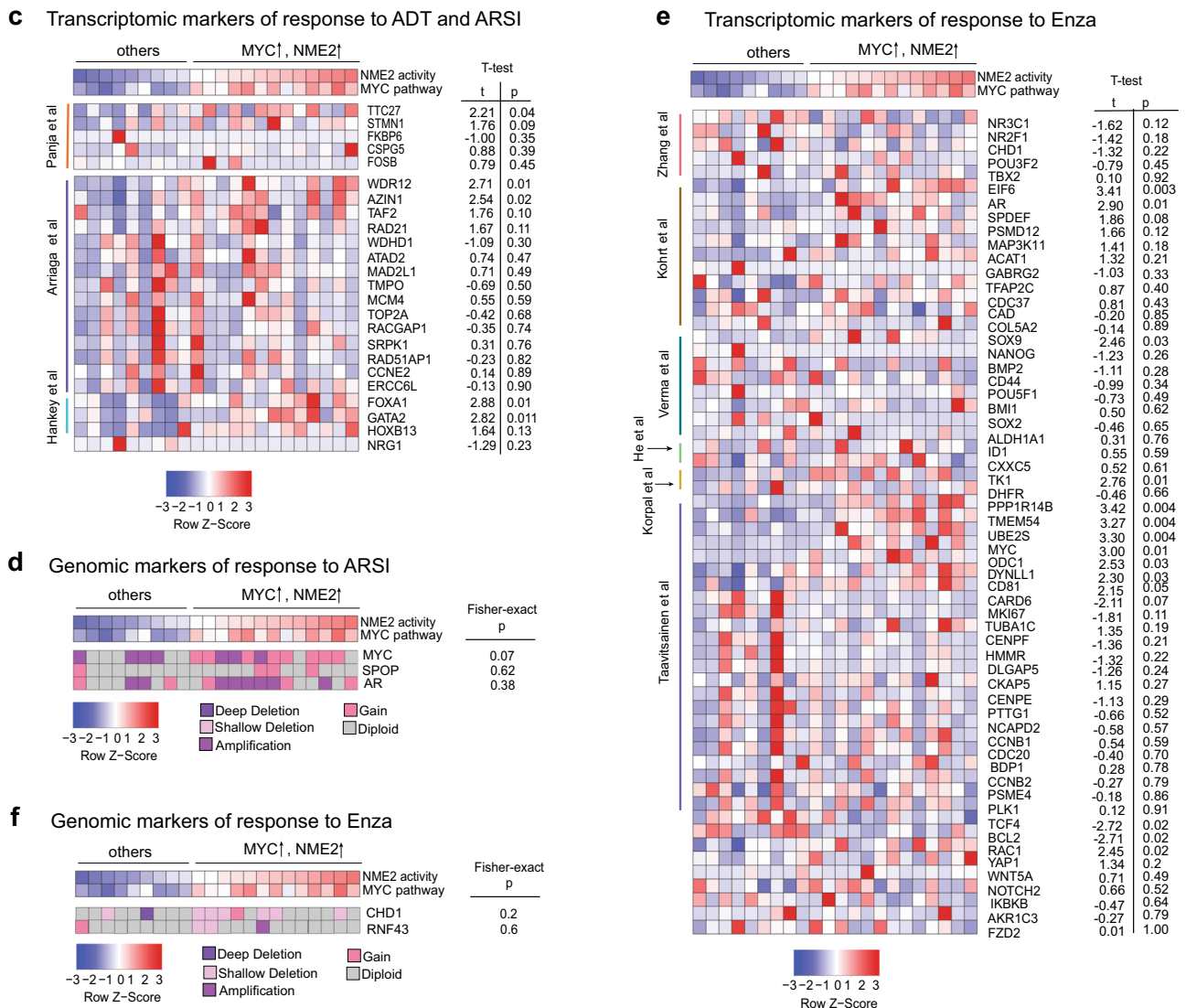
comparing high-MYC and high-NME2 group (yellow) to the rest of the patients (blue) in Abida et al. cohort subjected to adjuvant Enzalutamide. Log-rank  $p$  value, adjusted HR (hazard ratio), and CI (confidence interval) are indicated. **d** (Left, top) Pearson correlation analysis between NME2 TR and MYC pathway activity in SU2C West Coast cohort subjected to Enzalutamide and/or Abiraterone either before or after sample collection. Pearson  $r$  and  $p$  value are indicated. (Left, bottom) Patients with high-MYC and high-NME2 (yellow) and the rest of the patients (blue) were identified. (Right) Kaplan-Meier survival analysis, comparing high-MYC and high-NME2 group (yellow) to the rest of the patients (blue) in SU2C West Coast cohort. Log-rank  $p$  value, adjusted HR (hazard ratio), and CI (confidence interval) are indicated.



### Comparison to markers of overall PCa aggressiveness



### Comparison to markers of PCa treatment response



demonstrated no significant difference in Abiraterone-associated disease progression between the two identified patient groups (Supplementary Fig. 8, left, log-rank  $p$  value = 0.09, adjusted HR = 2.37, CI = 0.92–6.09, see Methods). ROC analysis<sup>72</sup> in the PROMOTE<sup>34</sup> cohort (see Methods) demonstrated that activation of NME2 and MYC did not

classify patients based on their binary response to Abiraterone (Supplementary Fig. 8, right, AUROC = 0.58, where AUROC = 0.5 indicates a random classifier). Taken together, these analyses suggest that partnership between NME2 TR and MYC pathway is associated with resistance to Enzalutamide.

**Fig. 6 | Ability of MYC and NME2 to predict Enzalutamide-response outperforms known markers of PCa progression and treatment response.** Comparison of MYC and NME2 ability to predict response to Enzalutamide in Abida et al. cohort to known markers of PCa aggressiveness, including (a) transcriptomic and (b) genomic markers. Comparison of MYC and NME2 ability to predict response to Enzalutamide in Abida et al. cohort to known markers of response to ADT and ARSIs including (c) transcriptomic and (d) genomic markers. Comparison of MYC and NME2 ability to predict response to Enzalutamide in Abida et al. cohort to known

markers of Enzalutamide-response, including (e) transcriptomic and (f) genomic markers. Two-tailed Welch t-test was utilized to calculate *p* values to estimate the difference in expression levels (red corresponds to over-expression and blue to under-expression) between high-MYC and high-NME2 group and the rest of the patients for transcriptomic markers in (a, c, e). Two-tailed Fisher-exact test was utilized to calculate *p* values to estimate the difference in the frequency/occurrence of any genomic alterations between high-MYC and high-NME2 group and the rest of the patients for genomic markers in (b, d, f).

### Comparison to common markers of PCa aggressiveness and treatment response

We next compared the ability of NME2 TR and MYC pathway to predict Enzalutamide resistance in Abida et al. cohort<sup>33</sup>, with the predictive ability of known markers of prostate cancer (i) aggressiveness, (ii) response to first-generation ADT and ARSIs (not specific to any particular drug), and (iii) Enzalutamide-specific response (Fig. 6). Transcriptomic and genomic markers were considered separately.

First, we evaluated the enrichment of transcriptomic markers of PCa aggressiveness<sup>73–86</sup> ( $n = 13$ , Fig. 6a) in high NME2 and MYC group, out of which *ERG*, and *DLX1* showed significant differential expression in the high-MYC and high-NME2 group, compared to the rest of the patients, “others” (two-tailed Welch t-test *p* value < 0.05, Supplementary Data 7A, see Methods). Interestingly, these genes have a direct relationship to MYC: *ERG* fusion (which eventually leads to over-expression of *ERG*) was shown to be correlated to *MYC* expression<sup>87</sup> and *DLX1* is a known transcriptional target of *ERG*<sup>76,88,89</sup> (Fig. 6a). To assess an independent association of the transcriptomic markers to Enzalutamide resistance (independent of MYC and NME2), we performed a univariable Cox proportional hazards model analysis<sup>38</sup>, using Enzalutamide-associated disease progression as the end-point in Enzalutamide-specific Abida et al.<sup>33</sup> cohort, which showed no significant association of these markers to Enzalutamide resistance (Supplementary Data 7A, see Methods). Among genomic markers<sup>21,32,90–92</sup> of PCa aggressiveness ( $n = 12$ , Fig. 6b, where *TP53* and *RBI* have also been shown to be markers of response to ARSIs<sup>33</sup>), *RBI* with shallow and deep deletion was significantly enriched in patients with high-NME2 and high-MYC pathway activity (Fisher’s exact test *p* value = 0.03) (Fig. 6b, Supplementary Data 7B, see Methods). Interestingly, *RBI* loss has been shown to correlate with *MYC* expression in small-cell lung carcinoma<sup>93</sup>, indicating potential cross-talk with the MYC pathway. Independent Cox proportional hazards model analysis<sup>38</sup> identified a shallow and deep deletion of *KRAS* to be significantly associated with response to Enzalutamide (Wald *p* value = 0.01, Supplementary Data 5B, see Methods), which has also been previously shown to be associated with MYC<sup>27</sup>.

Comparison to transcriptomic markers of first-generation ADT and ARSIs ( $n = 24$ ), taken from refs. 27,94–96), demonstrated that five markers (*TTC27*, *WDR12*, *AZINI*, *FOXAI*, and *GATA2*) were significantly differentially expressed in the high-MYC and high-NME2 patients (two-tailed Welch *t* test *p* value < 0.05, Supplementary Data 5C, Fig. 6c, see Methods). Interestingly, *WDR12* and *AZINI* are members of the MYC pathway and *TTC27*, *FOXAI*, and *GATA2* are MYC transcriptional targets<sup>86</sup>. Furthermore, Cox proportional hazards model analysis<sup>38</sup> indicated that five of the transcriptomic markers (*STMNI*, *WDR12*, *AZINI*, *MAD2L1*, and *MCM4*) had a significant association with response to Enzalutamide (Wald *p* value < 0.05, Supplementary Data 5C, see Methods), yet many of them were borderline significant and did not outperform MYC and NME2 (Supplementary Data 7C). Additionally, genomic markers of first-generation ADT and ARSIs described in refs. 27, 33 ( $n = 3$ ), had no significant enrichment in the high-MYC and high-NME2 group (Fig. 6d, Supplementary Data 7D, see Methods) or independent response to Enzalutamide.

Comparison to transcriptomic markers of Enzalutamide-specific response ( $n = 60$ ), described by refs. 18,97–100,20, in addition to

refs. 101–108) demonstrated that a group of 14 markers (*EIF6*, *AR*, *SOX9*, *TKI*, *PPP1R14B*, *TMEM54*, *UBE2S*, *MYC*, *ODCI*, *DYNLL1*, *CD81*, *BCL2*, *TCF4* and *RAC1*) had significant differential expression in the high-MYC and high-NME2 group (two-tailed Welch t-test *p* value < 0.05, Supplementary Data 7E, Fig. 6e, see Methods). Interestingly, 12 of these (*EIF6*, *SOX9*, *TKI*, *PPP1R14B*, *TMEM54*, *UBE2S*, *MYC*, *ODCI*, *DYNLL1*, *CD81*, *BCL2*, and *TCF4*) were transcriptional MYC targets, determined from ChEA transcription factor targets dataset<sup>86</sup>. Another member of this group, *AR*, as we have shown previously (Supplementary Fig. 1c), was also differentially expressed between the groups. Finally, *RAC1* (Fig. 6e, Supplementary Data 7E) is a member of the RAS pathway which has been shown to be associated with MYC pathway in CRPC samples<sup>27</sup>. Cox proportional hazards model analysis<sup>38</sup> demonstrated that 10 of the transcriptomic markers (i.e., *EIF6*, *ACAT1*, *TKI*, *PPP1R14B*, *TMEM54*, *UBE2S*, *DYNLL1*, *TUBA1C*, *RAC1*, and *WNT5A*) had significant association with response to Enzalutamide (Wald *p* value < 0.05, Supplementary Data 7E, see Methods), yet many of them were borderline significant and none of them outperformed NME2 and MYC (Supplementary Data 7E). None of the genomic markers of Enzalutamide-specific response ( $n = 2$ ), (described by refs. 18,109) were significantly enriched in the high-MYC and high-NME2 group (Fig. 6f, Supplementary Data 7F) or independently associated with response to Enzalutamide. Taken together, these findings indicate that the majority of the markers of PCa aggressiveness and therapeutic response that are enriched in the high-MYC and high-NME2 group are associated with MYC-related mechanisms and none of them outperform the ability of MYC and NME2 to predict risk of Enzalutamide resistance.

### Comparison to gene-centric computational methods

To evaluate if TR-2-PATH mechanism-centric predictions (i.e., activity levels of NME2 TR and MYC pathway) outperform predictive ability of commonly used gene-centric methods, we compared TR-2-PATH to differential expression analyses, Random (survival) Forest (RF)<sup>100</sup>, and Support Vector Machine (SVM)<sup>111</sup> methods all utilized on the Enzalutamide-specific Abida et al. cohort. For differential gene expression analysis, we considered genes that were differentially expressed between the three phenotypes (Intact, EnzaSens, and EnzaRes) in the mining step I and considered genes at (i) Welch t-test *p* value < 0.05; (ii) top 470 differentially expressed genes (comparable to the total number of target genes and pathway genes used for activity estimation) and *not excluding* target/pathway genes from NME2 TR and MYC pathway; (iii) top 470 differentially expressed genes, *excluding* target/pathway genes from NME2 TR and MYC pathway. For RF and SVM analysis, we utilized 470 genes from (iii) to avoid overfitting and then selected top 10 most significant genes/features from the outputs. Final gene list from each of these analyses was subjected to Kaplan-Meier survival analysis<sup>37</sup> and Cox proportional hazards model analysis<sup>38</sup> (crude and adjusted for age and Gleason), which did not show a significant association with Enzalutamide-associated disease progression using log-rank test, Wald test, or crude/adjusted hazards models (Supplementary Fig. 9a–d, see Methods). Such analysis demonstrates that mechanisms identified by TR-2-PATH (NME2 TR and MYC pathway) have significant advantage in predicting the risk of Enzalutamide resistance, compared to commonly used gene-centric methods.

## Targeting MYC and NME2 is beneficial in Enzalutamide-resistant conditions

Given that NME2 and MYC are upregulated in both patients that are at risk of Enzalutamide resistance and patients that fail Enzalutamide, we experimentally evaluated the benefits of therapeutic targeting of MYC and NME2 in similar experimental conditions. For this, we utilized LNCaP and C42B cell lines, as our experimental systems in Enzalutamide-naïve and Enzalutamide-resistant (EnzaRes) conditions (see Methods). To target MYC, we used MYCi975, a small molecule that directly inhibits MYC activity<sup>30</sup>. To understand the dose-dependent effect of MYCi975 in Enzalutamide-naïve and Enzalutamide-resistant conditions, we performed dose-response assays in both LNCaP and C42B cell lines (Fig. 7a, Supplementary Fig. 10a, see Methods) with varying doses of both drugs. This analysis demonstrated a striking reduction of the cell viability when treated with MYCi975 both in Enzalutamide-naïve and Enzalutamide-resistant conditions in a dose-dependent manner (Fig. 7a, Supplementary Fig. 10a).

Next, we utilized identified sub-IC<sub>50</sub> concentrations of Enzalutamide (10 μM) alone, MYCi975 (2 μM) alone, or Enzalutamide and MYCi975 in combination, to perform colony formation assay using Enzalutamide-resistant LNCaP and C42B cells (see Methods). Inhibition of MYC using MYCi975 reduced the colony formation of LNCaP-EnzaRes cells (one-tailed Welch t-test  $p$  value = 0.013, Supplementary Fig. 10b) and C42B-EnzaRes cells (one-tailed Welch t-test  $p$  value = 3.98E-6, Fig. 7b), compared to Intact (DMSO). Interestingly, the colony formation ability was significantly reduced when MYCi975 and Enzalutamide were administered in combination on LNCaP-EnzaRes cells (one-tailed Welch t-test  $p$  value = 6.39E-5, Supplementary Fig. 10b) and C42B-EnzaRes (one-tailed Welch t-test  $p$  value = 4.7E-8 Fig. 7b) compared to Intact (DMSO).

To evaluate the impact of Enzalutamide alone, MYCi975 alone, or in combination on the migratory capacity of Enzalutamide-resistant cells we performed Boyden chamber-based in vitro migration assay (see Methods) using C42B-EnzaRes cells (LNCaP cells do not migrate) and observed a significant reduction in cell migration when treated with MYCi975 (one-tailed Welch t-test  $p$  value = 0.02, Fig. 7c, see Methods) and even greater reduction when treated with a combination of MYCi975 and Enzalutamide (one-tailed Welch t-test  $p$  value = 0.003, Fig. 7c, see Methods). Positive control for MYC protein level after treatment with MYCi975 in C42B-EnzaRes cells was performed after 24 h of treatment (Supplementary Fig. 11).

Subsequently, to evaluate the effect of NME2 silencing on MYC expression, we first evaluated the expression of NME2 in LNCaP/C42B-EnzaRes cells, compared to wild-type LNCaP/C42B cells, which showed elevated levels of NME2 in both LNCaP-EnzaRes cells (Supplementary Fig. 10c, one-tailed Welch t-test  $p$  value = 0.00188, see Methods) and C42B-EnzaRes cells (Fig. 7d, one-tailed Welch t-test  $p$  value = 0.0005, see Methods). NME2 knockdown in C42B-EnzaRes cells using two different siRNAs (Fig. 7e, see Methods), demonstrated a significant reduction in expression of NME2 (Fig. 7e, left) and MYC (Fig. 7e, right), supported by the previously identified NME2 upstream regulation of MYC<sup>112-114</sup>.

Finally, to evaluate if silencing of NME2 could enhance the negative effect of MYC inhibition on tumor cell metastatic potential, we performed Boyden chamber-based in vitro migration assay using C42B-EnzaRes cells (see Methods), which demonstrated that cell migration was further reduced when NME2 knockdown was added to MYCi975 administration (Fig. 7f).

To evaluate whether NME2 tumor expression has a direct role in promoting resistance to Enzalutamide in vivo, we sought to test the hypothesis that loss of NME2 in Enzalutamide-resistant C42B cells is sufficient to restore Enzalutamide sensitivity. First, we tested this in vitro, by generating CRISPR KO of NME2 in the C42B-EnzaRes cells (Supplementary Fig. 12) and subjecting them to Enzalutamide treatment for 72 h. The cells with impaired NME2 expression displayed

increased sensitivity to Enzalutamide, compared to cells transfected with non-targeting sgRNA control (Fig. 7g).

To test this hypothesis in vivo, we engineered C42B-EnzaRes cells with a doxycycline-inducible shNME2 knockdown construct (C42B-EnzaRes Dox/shNME2). We validated NME2 knockdown and the role of NME2 in regulating MYC by measuring MYC protein levels after NME2 knockdown by Western blot (Fig. 7h). Mice were implanted with the C42B-EnzaRes Dox/shNME2 cells. When tumors were established, mice were subjected to Enzalutamide treatment in the presence or absence of Doxycycline in the drinking water or the appropriate control treatments. Enzalutamide alone had a minimal effect under the tested dosing conditions (Fig. 7i, Supplementary Fig. 13), while NME2 knockdown (with Doxycycline in drinking water) re-sensitized tumors to Enzalutamide, with an average tumor growth inhibition of 62%. Taken together, these results using relevant pre-clinical models of Enzalutamide resistance indicate that therapeutic targeting of the NME2/MYC axis is beneficial in Enzalutamide-resistant conditions and MYC inhibition could be combined with concurrent Enzalutamide administration for improved efficacy.

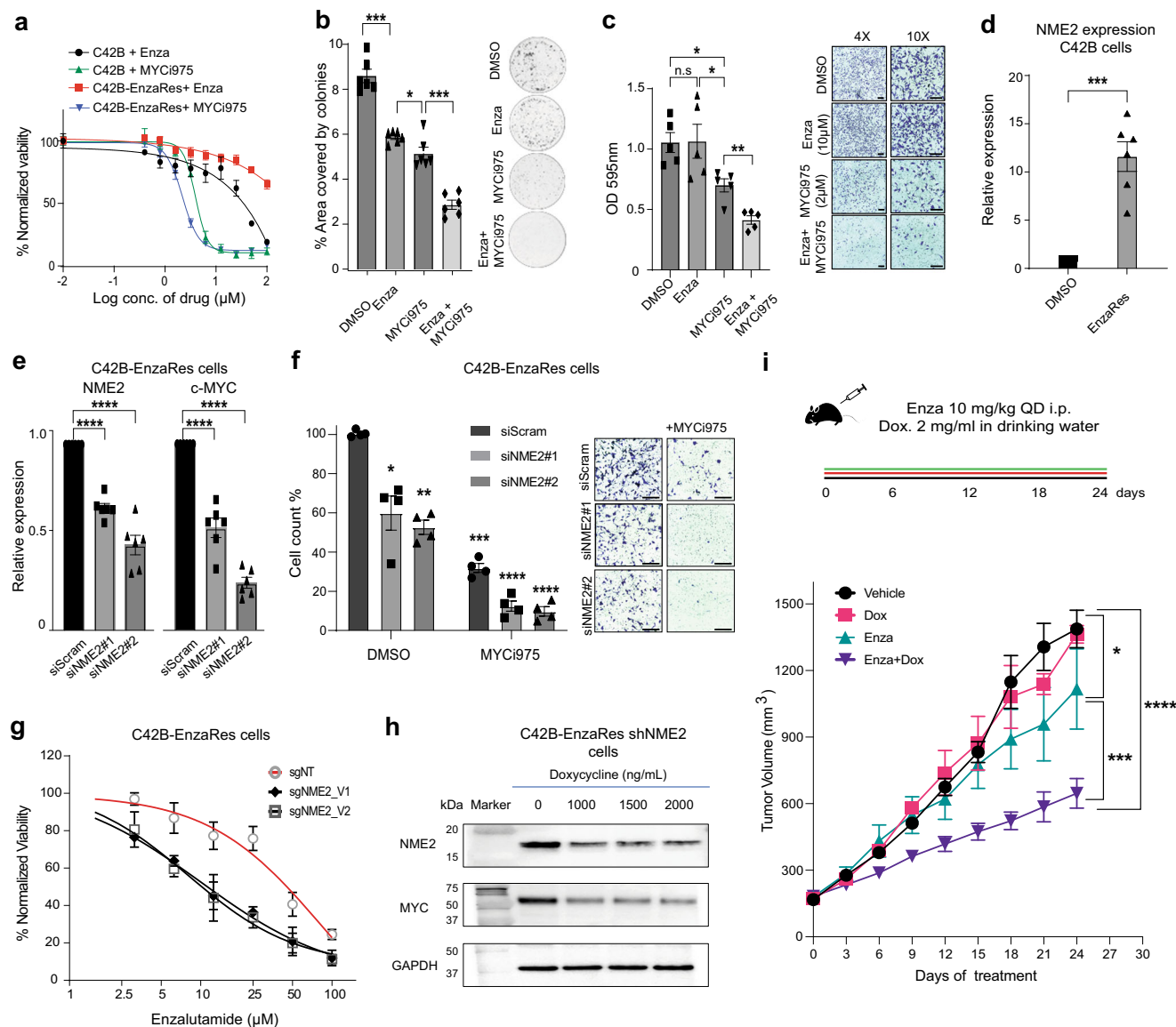
## Discussion

To investigate the mechanisms of Enzalutamide resistance in CRPC patients, we have reconstructed a CRPC-specific mechanism-centric regulatory network that encodes relationships between molecular pathways and their upstream transcriptional regulatory programs (i.e., TR-2-PATH). Such network has been reconstructed in a genome-wide manner for CRPC patients and was mined using signatures of Enzalutamide sensitivity and resistance. Our TR-2-PATH approach has several advantages that distinguish it from previously utilized methods. First, its focus on mechanism-centrality, which connects molecular pathways to their upstream regulatory programs, opens a door for the discovery of more effective biomarkers and optimized therapeutic interventions. Second, through a network mining/interrogation step our method overcomes several important drawbacks present in commonly utilized statistical analyses, including the multi-collinearity of the TRs and their effective prioritization for enhanced therapeutic targeting. Finally, our approach constitutes a valuable community resource to investigate CRPC phenotypes and could potentially be applicable to other cancer types.

We have interrogated the mechanism-centric CRPC-specific regulatory network using signatures of Enzalutamide sensitivity and resistance and identified MYC pathway and its upstream transcriptional regulator NME2 as markers of increased risk of resistance to Enzalutamide - a discovery that promises to prioritize patients for Enzalutamide intervention. Further, we have demonstrated that therapeutic targeting of MYC and NME2 could potentially provide an effective strategy for high-MYC and high-NME2 patients at risk of developing resistance to Enzalutamide and/or as a salvage therapy for patients that fail Enzalutamide.

The MYC oncogene<sup>30,115</sup> is a prominent early response gene and is known as a “super-transcription factor”, potentially regulating transcription of at least 15% of the entire genome<sup>116</sup> and includes a basic-region/ helix-loop-helix/ leucine-zipper (BR/HLH/LZ) motif at the C terminus and three conserved elements known as MYC boxes (MB) 1–3 at the N terminus<sup>117</sup>. To work as a transcription regulator, MYC interacts with another helix-loop-helix/ leucine-zipper protein, MAX, through the C-terminal BR/HLH/LZ motif to form a complex that binds to the conserved sequences (CACGTG) in the transcriptional regulatory regions of the target genes<sup>118-120</sup>, known to be involved in proliferation<sup>121</sup>, differentiation<sup>122</sup>, cell cycle<sup>121,122</sup>, metabolism<sup>123</sup>, apoptosis<sup>121</sup>, angiogenesis<sup>124</sup>, and therapeutic resistance<sup>23,27</sup> across different cancers.

Interestingly, a recent study by ref. 20 observed that MYC targets are open for binding in Enzalutamide-resistant conditions<sup>30</sup>, which prompted us to evaluate MYC TR activity in the Enzalutamide-



**Fig. 7 | MYC targeting is beneficial in Enzalutamide-resistant conditions.** **a** Drug sensitivity curves of Enzalutamide-naïve, or Enzalutamide-resistant (EnzaRes) C42B cells treated with MYC1975. Data are presented as mean values  $\pm$  SEM from  $n = 3$  biologically independent experiments. **b** Colony formation assay using Enzalutamide-resistant (EnzaRes) C42B cells in Intact (i.e., treated with DMSO), treated with Enzalutamide (10  $\mu$ M), MYC1975 (2  $\mu$ M), or a combination of Enzalutamide+MYC1975 (10  $\mu$ M + 2  $\mu$ M). Cells were grown in the presence of respective drugs. Representative images are shown, and data are presented as mean values  $\pm$  SEM from  $n = 6$  biologically independent experiments.  $P$  value was estimated using a one-tailed Welch t-test. \* $p$  value  $< 0.05$ , \*\*\* $p$  value  $\leq 0.001$ . **c** Boyden chamber-based in vitro migration assay using Enzalutamide-resistant (EnzaRes) C42B cells in Intact (i.e., treated with DMSO), treated with Enzalutamide (10  $\mu$ M), MYC1975 (2  $\mu$ M), or a combination of Enzalutamide+MYC1975 (10  $\mu$ M + 2  $\mu$ M). Representative images are shown, data are presented as mean values  $\pm$  SEM from  $n = 5$  biologically independent experiments, indicating the quantification of Crystal Violet trapped by migrated cells. Scale bars, 100  $\mu$ m.  $P$  value was estimated using a one-tailed Welch t-test. \* $p$  value  $< 0.05$ , \*\* $p$  value  $\leq 0.01$ . **d** Expression of NME2 in Intact and Enzalutamide-resistant (EnzaRes) C42B cells, using qRT-PCR, data are presented as mean values  $\pm$  SEM from  $n = 6$  biologically independent experiments.  $P$  value was estimated using the one-tailed Welch t-test. \*\*\* $p$  value  $\leq 0.001$ . **e** Two different siRNAs targeting NME2 were used to downregulate NME2 (left panel) and its effect on MYC expression using qRT-PCR is shown (right panel), data are presented as

mean values  $\pm$  SEM from  $n = 6$  biologically independent experiments.  $P$  value was estimated using one-tailed Welch t-test. \* $p$  value  $< 0.05$ . **f** Boyden chamber-based in vitro migration assay using Enzalutamide-resistant (EnzaRes) C42B cells treated with DMSO or MYC1975 (2  $\mu$ M) with or without knockdown of NME2. Representative images are shown, data are presented as mean values  $\pm$  SEM from  $n = 4$  biologically independent experiments, indicating the cell count quantification of Crystal Violet trapped by migrated cells, normalized to control (DMSO with siScramble). Scale bars, 100  $\mu$ m.  $P$  value was estimated using 2-way ANOVA, \* $p$  value  $< 0.05$ , \*\* $p$  value  $\leq 0.01$ , \*\*\* $p$  value  $\leq 0.001$ , \*\*\*\* $p$  value  $\leq 0.0001$ . **g** Comparison of anti-proliferative effects of Enzalutamide on C42B EnzaRes cells that have NME2 knocked down via CRISPR (sgNME2\_V1 and sgNME2\_V2) versus C42B EnzaRes which have received non-targeting sgRNA control (sgNT). Data are presented as mean values  $\pm$  SEM from  $n = 3$  independent biological replicates for each cell line. **h** Western blot of NME2, MYC and GAPDH protein levels in C42B shNME2 EnzaRes cells treated with Doxycycline for 96 h at the indicated concentrations. ( $n = 3$  biologically independent experiments, representative blot shown). **i** Schematic representation and tumor volumes for mice bearing established C42B EnzaRes shNME2 xenografts treated with vehicle ( $n = 7$  mice), Enzalutamide (10 mg/kg QD i.p) and/or Doxycycline (2 mg/mL in drinking water) ( $n = 6$  mice for Dox, Enza, Enza + Dox arms) for 24 days. Statistical significance was performed via two-way ANOVA. \* $p$  value  $< 0.05$ , \*\*\* $p$  value  $\leq 0.001$ , \*\*\*\* $p$  value  $\leq 0.0001$ . Source data for all the panels in this figure are provided as a Source Data file.



associated Abida et al.<sup>33</sup> cohort. When MYC TR activity was utilized as a predictive factor, we observed its modest ability to identify patients at risk of resistance to Enzalutamide (log-rank  $p$  value = 0.01), yet at a smaller scale compared to the MYC pathway and its NME2 TR partnership (log-rank  $p$  value = 0.0035), indicating a potential cross-talk between MYC transcriptional regulatory machinery and the MYC pathway.

Several groups have been working on direct and indirect strategies to target MYC-related mechanisms<sup>30,125,126</sup>, with ref. 30 developing a small-molecule MYCi975 that directly inhibits MYC through disrupting MYC/MAX interaction (i.e., heterodimerization of MYC with MAX) to suppress the expression of MYC target genes and reduce MYC protein stability by enhancing phosphorylation of threonine 58. Administration of MYCi975 significantly reduces MYC-dependent cancer-cell proliferation, expression of MYC target genes, and tumor growth. Here, we demonstrate that administration of MYCi975 is beneficial in high-MYC Enzalutamide-resistant conditions and constitutes a potential therapeutic option for high-MYC CRPC patients who are at risk of Enzalutamide-resistance and/or who failed Enzalutamide. Our results agree with recent observations by Holmes et al.<sup>127</sup> that demonstrated MYCi975 regulates AR, AR-V7 and FOXA1 and sensitizes AR-V7+, Enzalutamide-resistant 22Rv1 tumors to Enzalutamide. Furthermore, a positive correlation of MYC pathway activity levels with AR expression and activity, previously observed by refs. 26,128, has also been confirmed by our investigations. Since MYC is known to bind to the regulatory regions of AR<sup>26</sup>, MYC targeting (alone or in combination with AR inhibitors) in AR-associated conditions is a promising therapeutic avenue for a high-MYC subset of CRPC patients.

Our approach identified NME2 as a transcriptional regulatory program upstream of MYC, involved in Enzalutamide resistance in CRPC. Previously, NME2 has been shown to bind to the nuclease-hypersensitive element (NHEIII) in the promoter region of MYC interacting with G-quadruplex structure<sup>65</sup> and increase MYC expression, nominating NME2 as a transcriptional regulator of MYC<sup>65,129</sup>. NME2 is a member of the NME family<sup>65</sup> that has been shown to be involved in tumorigenesis and treatment response across different cancer types. In particular, ref. 130 demonstrated that NME2 plays a vital role in maintaining stemness of gastric cancer stem cells by enhancing the expression of anti-apoptosis genes, ref. 131 demonstrated suppression of apoptosis via NME2-mediated miR-100 upregulation in the development of gastric cancer, and ref. 132 demonstrated that over-expression of NME2 is associated with acquired resistance to 5-fluorouracil in colorectal cancer, yet its role (alongside partnership with MYC) in Enzalutamide resistance in CRPC patients has not been explored yet.

Here, we propose that knockdown of NME2, alongside targeting of MYC, is beneficial in Enzalutamide-resistant conditions for patients with high-MYC and high-NME2 profiles. Recently, ref. 129 have shown that a small molecule stauprimide binds to NME2 and inhibits its nuclear localization, thus affecting MYC transcription in renal cancer cell lines RXF 393 and CAKI-1. While shown to be effective in renal-specific cells, a potent NME2 inhibitor in PCa/CRPC settings is yet to be discovered and would constitute an effective combination strategy for high-MYC and high-NME2 CRPC patients.

In conclusion, we propose that MYC-centric mechanisms could be effectively utilized as markers to identify CRPC patients at risk of resistance to Enzalutamide. Pre-screening patients prior to Enzalutamide administration could potentially enhance personalized therapeutic planning, preclude harmful side effects, and extend patient survival. Further, we nominate MYC and NME2 as potential therapeutic targets for CRPC patients with high-MYC and high-NME2 profiles who are at risk of resistance to Enzalutamide and/or as a salvage therapy for patients that have failed Enzalutamide treatment.

## Methods

### Computational methods

**Datasets utilized in this work.** Datasets utilized for network construction, mining, validation, and negative control analysis are summarized in Supplementary Data 1 (supplied separately).

- (i) **Dataset to associate activity levels of MYC pathway with response to Enzalutamide:** To determine if increased activity levels of MYC pathway (i.e., HALLMARK\_MYC\_TARGETS\_V2 pathway from Hallmark collection in MSigDB 3.0) were associated with Enzalutamide resistance, we utilized Enzalutamide-associated CRPC metastatic samples from ref. 33 cohort (fresh-frozen needle biopsies), profiled on Illumina HiSeq 2500 and downloaded from [https://github.com/cBioPortal/datahub/tree/master/public/prad\\_su2c\\_2019](https://github.com/cBioPortal/datahub/tree/master/public/prad_su2c_2019). We specifically selected samples that at the time of biopsy (sample collection) were ARSI-naïve (i.e., not subjected to any ARSI treatment), treated with Enzalutamide after sample collection, and then followed up for Enzalutamide-associated disease progression ( $n = 22$ , one sample per patient, as described in ref. 33). In this sub-group, the mean age at diagnosis was 59 years with a standard deviation of 6.85, the mean age at biopsy was 67.6 years with a standard deviation of 8.3, and the mean prostate-specific antigen (i.e., PSA) was 189.4 ng/ml with a standard deviation of 526.18. Metastatic composition of this sub-group included lymph node ( $n = 13$ ), bone ( $n = 6$ ), lung ( $n = 1$ ), other soft tissue ( $n = 1$ ), and liver ( $n = 1$ ) samples. We utilized Enzalutamide-associated disease progression, defined as the time on Enzalutamide treatment without being subjected to another agent such as taxane, as the clinical end-point (as defined and suggested in ref. 33).
- (ii) **Dataset to associate activity levels of MYC pathway with response to Abiraterone:** To determine if elevated activity levels of MYC pathway (i.e., MYC pathway was utilized as a geneset from HALLMARK\_MYC\_TARGETS\_V2 pathway in Hallmark collection in MSigDB 3,  $n = 58$ , Supplementary Data 2) were specifically associated with Enzalutamide (and not Abiraterone) resistance, we utilized Abiraterone-associated metastatic CRPC sample from ref. 33 cohort. We specifically selected samples that at the time of biopsy (sample collection) were ARSI-naïve (as above), treated with Abiraterone after sample collection, and then followed up for Abiraterone-associated disease progression ( $n = 33$ , one sample per patient) for negative control analysis. The mean age at diagnosis for this patient sub-group was 61.38 years with a standard deviation of 5.94, the mean age at biopsy was 66.73 years with a standard deviation of 7.02, and the mean PSA was 51.4 ng/ml with a standard deviation of 91.05. Metastatic composition of this sub-group included lymph node ( $n = 18$ ), bone ( $n = 11$ ), liver ( $n = 2$ ), and other soft tissue ( $n = 2$ ) samples. We utilized Abiraterone-associated disease progression, defined as the time on Abiraterone treatment, without being subjected to other agents such as taxane, as the clinical end-point (as defined and suggested in ref. 33).
- (iii) **Dataset for network reconstruction:** To construct a mechanism-centric network, we utilized the Stand Up to Cancer (SU2C) East Coast cohort<sup>32,33</sup>, profiled on Illumina HiSeq 2500 and downloaded from dbGaP phs000915.v2.p2. This cohort included metastatic CRPC samples, obtained as fresh-frozen needle biopsies. We examined 280 samples available at dbGaP, and to avoid any overlap with treatment-associated analysis in ref. 33 (which we have utilized in part for validation and in part for a negative control), we removed all SU2C East Coast cohort samples that were present in ref. 33 ( $n = 29$ ). Subsequently, we also removed samples that were duplicated (i.e., when the same sample was sequenced by different facilities) and selected one sample per patient to avoid signal duplication for our final network-building. Our final cohort comprised 153 patients with a

- mean age at diagnosis of 59.2 years with a standard deviation of 8.38, a mean age at biopsy of 66.1 years with a standard deviation of 8.07, and a mean PSA of 234.5 ng/ml with a standard deviation of 1574.4. Metastatic composition of this cohort included adrenal ( $n=1$ ), bone ( $n=39$ ), liver ( $n=26$ ), lymph node ( $n=57$ ), other soft tissue ( $n=19$ ), prostate ( $n=4$ ), lung ( $n=2$ ) and unknown origin ( $n=5$ ) samples. At the time of biopsy (sample collection), patients either were exposed to ARSI ( $n=67$ ), were ARSI-naïve ( $n=75$ ), were on treatment ( $n=4$ ), or their treatment was unknown ( $n=7$ ).
- (iv) **Datasets for network mining:** For network mining (i.e., query/interrogation), we utilized LNCaP cell line samples from ref. 53 ( $n=12$ ), that were profiled with the HumanHT-12 v4 Expression BeadChip and downloaded from GEO GSE78201. These dataset included three phenotypes: (i) LNCaP cells subjected to DMSO (referred to as Intact to indicate that they were not subjected to Enzalutamide treatment) ( $n=4$ ); (ii) LNCaP cells subjected to Enzalutamide for 48 h and sensitive to it (referred to as Enzalutamide-sensitive, EnzaSens) ( $n=4$ ); and (iii) LNCaP cells subjected to Enzalutamide for 6 months and having developed resistance to it (referred as Enzalutamide-resistant, EnzaRes) ( $n=4$ ).
- (v) **Datasets for clinical validation:** For validation purposes, we utilized (i) ref. 19 (ii) Enzalutamide-associated ref. 33, and (iii) SU2C West Coast<sup>70,71</sup> datasets. First, to confirm that upregulation of the NME2 transcriptional regulatory program and MYC pathway characterize Enzalutamide-naïve samples (i.e., before patients were exposed to Enzalutamide) from patients that were later exposed to Enzalutamide and eventually failed it, we selected two sequential single-cell samples from the same CRPC patient (01115655) from ref. 19 cohort. These samples were profiled on Illumina NextSeq 500 and downloaded from [https://singlecell.broadinstitute.org/single\\_cell/study/SCP1244/transcriptional-mediators-of-treatment-resistance-in-lethal-prostate-cancer](https://singlecell.broadinstitute.org/single_cell/study/SCP1244/transcriptional-mediators-of-treatment-resistance-in-lethal-prostate-cancer). In particular, the first sample was collected before the patient was subjected to Enzalutamide and second sample was collected after the same patient received Enzalutamide and developed resistance to it. Both samples were collected from the lymph-node metastatic site. Finding from ref. 19 were confirmed in Enzalutamide-associated Abida et al.<sup>33</sup> cohort. Briefly, we selected a subset of patients that were ARSI-naïve at biopsy, treated with Enzalutamide after sample collection, and subsequently monitored for Enzalutamide-associated disease progression ( $n=22$ , as described above). Further, we validated the predictive ability of NME2 TR and MYC pathway in SU2C West Coast cohort<sup>70,71</sup>, which comprises samples from CRPC patients (obtained from fresh frozen image-guided core needle biopsies), profiled on Illumina HiSeq 2500 or NextSeq 500, accessed through dbGap phs001648.v2.p1 and downloaded from GDC (<https://portal.gdc.cancer.gov/projects/WCDT-MCRPC>). The samples in this cohort were subjected to Enzalutamide and/or Abiraterone either before biopsy (sample collection) or after biopsy (sample collection). Subsequently, all patients were monitored for disease progression ( $n=83$ , one sample per patient). The mean age for the patients in this cohort was 70.59 years with standard deviation of 8.14. The patients in this cohort were from various races, including, white ( $n=70$ ), Asian ( $n=2$ ), African American ( $n=5$ ) and unknown ( $n=6$ ). Metastatic composition of this cohort included bone ( $n=36$ ), liver ( $n=7$ ), lymph node ( $n=31$ ), and unknown ( $n=9$ ). Further, samples were obtained from patients who were either in M1b stage (i.e., when prostate cancer has spread to bone,  $n=36$ ) or M1c stage (i.e., when prostate cancer has spread to other parts of the body,  $n=47$ ). We utilized treatment-associated disease progression (defined as an increase in PSA level (minimum 2 ng/mL) that has risen at least twice in an interval of least one week or soft tissue progression (nodal and visceral) based on RECIST v1.1) as the clinical end-point (as defined and suggested in refs. 70,71).
- (vi) **Additional datasets for validation of association between NME2 TR and MYC pathway activities:** To further validate the association between NME2 TR and MYC pathway activities, we utilized (i) Alumkal et al.<sup>12</sup>; (ii) Beltran et al.<sup>68</sup> and (iii) Kumar et al.<sup>69</sup> CRPC patient cohorts. Alumkal et al.<sup>12</sup> cohort includes CRPC samples isolated by laser capture microdissection from frozen biopsies ( $n=24$ ), profiled on Illumina HiSeq 2500 or NextSeq 500, and downloaded from Supplementary material of the associated paper<sup>12</sup>. Beltran et al.<sup>68</sup> cohort includes CRPC samples ( $n=34$ , neuroendocrine samples excluded) profiled on HiSeq 2500 and downloaded from dbGaP (phs000909.v1.p1). These samples were obtained at biopsy from different metastatic sites including adrenal ( $n=1$ ), bone ( $n=11$ ), liver ( $n=3$ ), lung ( $n=2$ ), lymph node ( $n=9$ ), prostate ( $n=7$ ), and skull base ( $n=1$ ). ref. 69 cohort includes CRPC samples obtained at rapid autopsy, profiled on Agilent 44 K whole human genome expression oligonucleotide microarray and downloaded from GEO (GSE77930). The samples were obtained from different metastatic sites which included liver ( $n=21$ ), lymph node ( $n=69$ ), lung ( $n=22$ ), bone ( $n=20$ ), retroperitoneal ( $n=4$ ), kidney ( $n=1$ ), appendix ( $n=1$ ), peritoneum ( $n=2$ ), adrenal ( $n=4$ ), bladder ( $n=5$ ), bladder neck ( $n=2$ ), pelvic mass ( $n=1$ ), peritoneal ( $n=1$ ), renal ( $n=1$ ), scrotum ( $n=1$ ), skin ( $n=1$ ), spleen ( $n=1$ ).
- (vii) **Datasets for negative control analysis:** To evaluate if the predictive ability of NME2 TR and MYC pathway are indeed Enzalutamide specific, we utilized (i) Abiraterone-associated ref. 33 cohort (as described above); and (ii) PROMOTE<sup>34</sup> cohort, as negative controls. As described above, Abiraterone-associated ref. 33 cohort included ARSI-naïve CRPC samples obtained at biopsy, treated with Abiraterone after sample collection, and subsequently monitored for Abiraterone-associated disease progression ( $n=33$ , as described above). PROMOTE<sup>34</sup> cohort included samples from patients with CRPC profiled on Illumina HiSeq 2500 and downloaded from dbGaP phs001141.v1.p1. These samples were obtained at biopsy from different metastatic sites ( $n=77$ , one sample per patient), including bone ( $n=56$ ), soft tissue ( $n=2$ ), liver ( $n=2$ ), prostate bed ( $n=2$ ), lymph-node ( $n=14$ ), and lung ( $n=1$ ) and were ARSI-naïve at the time of sample collection. After sample collection, the patients were subjected to Abiraterone for 12 weeks, and were assessed for Abiraterone-associated disease progression right after that, which was defined based on the score that combined serum PSA level, bone and CT imaging, and symptom assessment at week 12. Patients that developed disease progression at week 12 were classified as non-responders ( $n=32$ ) and those that did not develop disease progression at week 12 were classified as responders ( $n=45$ ).
- Data download, processing, and normalization.** Abida et al.<sup>33</sup> RNA-seq samples profiled on Illumina HiSeq 2500, were downloaded from [https://github.com/cBioPortal/datahub/tree/master/public/prad\\_su2c\\_2019](https://github.com/cBioPortal/datahub/tree/master/public/prad_su2c_2019) as Fragments Per Kilobase of transcript per Million mapped reads (FPKM). The clinical and treatment data were downloaded from the supplementary material of the corresponding paper<sup>33</sup> and from cBioPortal (<https://www.cbioportal.org/>).
- SU2C East Coast<sup>32</sup> cohort RNA-seq samples profiled on Illumina HiSeq 2500, were requested and downloaded from dbGaP phs000915.v2.p2 as SRA files using the *prefetch* command and were converted to FASTQ files utilizing the *fastq-dump* command from *sra*

*toolkit* (version 10.8.2)<sup>133</sup>. Following this, the FASTQ files were aligned to a reference genome hg19 using STAR aligner<sup>134</sup> with the *quantMode* option, which generated raw count files. The raw counts were normalized using R *DESeq*<sup>135</sup> package for further statistical analysis. The clinical data were obtained from the supplementary material of the corresponding paper<sup>33</sup> and from cBioPortal (<https://www.cbioportal.org/>).

Kregel et al.<sup>33</sup> LNCaP cell line samples were profiled on HumanHT-12 v4 Expression BeadChip Kit and their quantile-normalized<sup>136</sup> gene expression data were downloaded from GEO GSE78201. The phenotype information was obtained from GEO GSE78201.

He et al.<sup>19</sup> single-cell RNA-seq samples profiled on Illumina NextSeq 500, were downloaded from [https://singlecell.broadinstitute.org/single\\_cell/study/SCP1244/transcriptional-mediators-of-treatment-resistance-in-lethal-prostate-cancer](https://singlecell.broadinstitute.org/single_cell/study/SCP1244/transcriptional-mediators-of-treatment-resistance-in-lethal-prostate-cancer), as single-cell Transcripts Per Million (TPM) data matrix. The clinical data were obtained from the main body and supplementary material of the corresponding paper<sup>19</sup>.

SU2C West Coast<sup>70</sup> cohort RNA-seq samples profiled on either Illumina HiSeq 2500 or NextSeq 500 were downloaded from GDC (<https://portal.gdc.cancer.gov/projects/WCDT-MCRPC>) as BAM files. These BAM files were then converted to FASTQ files utilizing *bam2-fastq* from *bedtools*<sup>137</sup>. Subsequently, the FASTQ files were aligned to a reference genome hg19 using STAR aligner<sup>134</sup> with the *quantMode* option, which generated raw count files. The raw counts were normalized using R *DESeq*<sup>135</sup> for further statistical analysis. The clinical and treatment data were obtained from GDC (<https://portal.gdc.cancer.gov/projects/WCDT-MCRPC>).

Alumkal et al. CRPC samples profiled on either Illumina HiSeq 2500 or NextSeq 500 were downloaded from the Supplementary file of the corresponding paper<sup>12</sup> as Transcripts Per Million (TPM) data matrix. The clinical file was obtained from the main body and Supplementary Material of the corresponding paper<sup>12</sup>.

Beltran et al. cohort RNA-seq samples profiled on Illumina HiSeq 2500 were requested and downloaded from dbGaP phs000909.v1.p1 as SRA files using the *prefetch* command and were converted to FASTQ files utilizing the *fastq-dump* command from *sra toolkit* (version 10.8.2)<sup>133</sup>. Following this, the FASTQ files were aligned to a reference genome hg19 using STAR aligner<sup>134</sup> with the *quantMode* option, which generated raw count files. The raw counts were normalized using R *DESeq*<sup>135</sup> package for further statistical analysis. The clinical data were obtained from dbGaP phs000909.v1.p1.

Kumar et al. cohort CRPC samples profiled on Agilent 44 K whole human genome expression oligonucleotide microarray and their gene expression data were obtained from GEO (GSE77930). The clinical data was obtained from GEO (GSE77930).

PROMOTE<sup>34</sup> RNA-seq samples profiled on Illumina HiSeq 2500, were requested and downloaded from dbGaP phs001141.v1.p1 as SRA files using the *prefetch* command and then converted to FASTQ files using the *fastq-dump* command from *sra toolkit* (version 10.8.2)<sup>133</sup>. Subsequently, the FASTQ files were aligned to the reference genome hg19 using STAR aligner<sup>134</sup> with the *quantMode* option to generate raw count files. The raw count files were normalized using R *DESeq*<sup>135</sup> package. The clinical data were obtained from dbGaP phs001141.v1.p1.

**Estimating activity levels of molecular pathways.** A list of molecular pathways and their corresponding genes were obtained from Molecular Signatures Database (MSigDB)<sup>138</sup>, available from Broad institute, and included C2 pathway collection (KEGG<sup>41</sup>, BioCarta<sup>42</sup>, and Reactome<sup>43</sup>) and Hallmark<sup>44</sup> gene sets. To estimate activity levels of each molecular pathway, we utilized signature-based or single-patient (i.e., single-sample) based Gene Set Enrichment Analysis (GSEA)<sup>36</sup>, similarly to refs. 139,140. For the signature-based GSEA analysis, a signature of interest (e.g., defined as a list of genes ranked by their differential expression using two-tailed Welch t-test between any two phenotypes of interest, such as Enzalutamide-resistant and Enzalutamide-sensitive phenotypes) is used as a

reference signature and genes from a specific pathway are used as a query gene set. For single-patient (i.e., single-sample) GSEA analysis, gene expression profiles were scaled/standardized (i.e., z-scored) on gene-level so that mean of values for each gene was 0 and the standard deviation was 1, allowing for comparison of gene ranks across different samples. A single-sample signature was defined as a list of genes ranked by their z-scores and utilized as a reference signature in single-sample GSEA analysis (pathway genes were utilized for query, in the same manner as above). For signature-based and single-sample GSEA analysis, Normalized Enrichment Score (NES) and *p* values were estimated using 1000 gene permutations. NESs from this analysis were utilized as pathway activity values, where positive NES corresponds to an enrichment of pathway genes in the over-expressed part of the signature and negative NES corresponds to an enrichment of pathway genes in the under-expressed part of the signature.

#### Estimating activity levels of transcriptional regulatory programs.

To estimate the activity levels of transcriptional regulators we utilized MARINA<sup>141</sup> (for a signature-based analysis) and VIPER<sup>45</sup> (for a single-sample-based analysis). Signatures were defined in the same manner as for the pathway enrichment analysis and were utilized as a reference for MARINA/VIPER. Instead of utilizing pathway data, MARINA and VIPER analyses require tissue-specific prostate cancer transcriptional regulatory network (interactome), as reconstructed previously in ref. 39 This interactome comprises of transcriptional regulators (TR, transcription factors and co-factors) and their potential transcriptional targets, connected by the transcriptional regulatory relationships. During MARINA/VIPER analysis, these transcriptional targets (for each transcriptional regulator separately) are utilized as a query gene set. We refer to the TR and the set of its corresponding transcriptional targets as a transcriptional regulatory program. Similar to GSEA, NESs/z-scores from MARINA and VIPER analysis were utilized to define activity levels of TRs. MARINA was implemented using *msvip* function and VIPER was implemented using *viper* function from R *VIPER* package in Bioconductor<sup>45</sup>.

**TR-2-PATH: reconstruction of a mechanism-centric regulatory network.** To identify potential regulatory relationships between molecular pathways and their upstream transcriptional regulatory programs in CRPC patients, we have reconstructed a CRPC-specific mechanism-centric regulatory network using TR-2-PATH method. In this network, each node represents a mechanism: a molecular pathway or transcriptional regulatory program. SU2C East Coast cohort (as described above) was first scaled/standardized on the gene level and then subjected to single-sample pathway enrichment analysis (as described above) and single-sample transcriptional regulatory analysis (as described above). We then defined activity vectors for each molecular pathway (where each pathway vector corresponds to the NESs for this pathway across all patients in the SU2C East Coast cohort) and for each TR program (where each TR vector corresponds to the NESs/z-scores for this TR across all patients in SU2C East Coast cohort). Specifically, let us assume that we have *n* samples. If the activity level of pathway *i* in sample *j* is  $NES_{i,j}$ , then the activity vector for pathway *i*,  $P_{i\ activity}$  is defined as.

$$P_{i\ activity} = \begin{bmatrix} NES_{i,1} \\ NES_{i,2} \\ \vdots \\ NES_{i,n} \end{bmatrix} \quad (1)$$



Similarly, if the activity level of a TR  $t$  in a sample  $j$  is  $a_{t,j}$ , then the activity vector for TR  $t$ ,  $TR_{t\ activity}$  is defined as

$$TR_{t\ activity} = \begin{bmatrix} a_{t,1} \\ a_{t,2} \\ \vdots \\ a_{t,n} \end{bmatrix} \quad (2)$$

To estimate potential regulatory relationships between transcriptional regulatory programs and molecular pathways, we first performed a pairwise comparison of each TR activity vector and each pathway activity vector using linear regression analysis, where a TR activity vector was used as a predictor variable (independent variable) and pathway activity vector was used as a response variable (dependent variable), as below. For each pathway  $i$  and TR  $t$ :

$$P_{i\ activity} = \alpha + \beta * TR_{t\ activity} \quad (3)$$

The positive Beta ( $\beta$ ) coefficient from the linear regression analysis (which corresponds to a positive slope for the fitted line between TR activity vector and pathway activity vector) indicated a positive relationship/association from the TR to the pathway and a negative Beta coefficient (negative slope) indicated a negative relationship/association from the TR to the pathway. Following the regression analysis for all TR-pathway pairs, we subjected it to multiple hypotheses FDR correction, which was performed for each pathway separately. If this relationship showed  $FDR < 0.05$ , it was added as an edge to the final network. Otherwise, it was discarded. Linear regression analysis was performed using the R *lm* function<sup>142</sup> and multiple hypotheses testing per pathway was performed using the R *p.adjust* function<sup>142</sup>.

#### Bootstrap analysis for the mechanism-centric regulatory network.

To evaluate if the edges in the mechanism-centric regulatory network could be “recovered” in the presence of noise (re-sampling), we performed bootstrap analysis. For this, SU2C East Coast cohort gene expression profiles ( $n=153$ ) were sampled with replacements 100 times. Each sampled/bootstrapped gene expression profile was then used to reconstruct a bootstrapped mechanism-centric regulatory network using the TR-2-PATH method (as above). We then utilized results from these 100 networks to assign weights to each edge, which was defined as the number of times this edge appeared (i.e., was recovered) across 100 bootstrapped networks (i.e., edge frequency). In particular, the edge weights were defined as the percent (%) of times an edge identified in the original network was also identified across the bootstrapped networks while maintaining the same direction of the relationship (positive/negative) between a particular TR program and a particular molecular pathway, across all 100 bootstrapped networks. These edge weights were then added to the original network (making it a weighted mechanism-centric network) and further utilized in the network query step.

The R functions *hist* and *density*<sup>142</sup> were utilized to depict weight distributions. To cluster the molecular pathways based on their edge weights, we utilized t-distributed stochastic neighbor embedding clustering<sup>47</sup> (t-SNE), a common dimensionality reduction technique that clustered pathways with similar edge weight patterns as nearby points and pathways with dissimilar edge weight patterns as distal points. t-SNE was implemented using the *Rtsne* function from R *Rtsne* package<sup>143</sup>.

#### Network mining I: identifying differentially altered sub-networks.

To identify parts of the mechanisms-centric network (i.e., sub-networks comprising of the molecular pathways and their upstream

TR programs) that significantly alter their activity across the response to Enzalutamide, we queried (i.e., mined) the mechanism-centric regulatory network using signatures of Enzalutamide-response. In particular, we specifically utilized gene expression profiles from ref. 53 (as described above), which consists of (i) Intact (DMSO subjected) LNCaP cells ( $n=4$ ), (ii) Enzalutamide-sensitive (EnzaSens) LNCaP cells ( $n=4$ ); and (iii) Enzalutamide-resistant (EnzaRes) LNCaP cells ( $n=4$ ). We hypothesized that if a particular sub-network is up-regulated (positive NES) in the intact state, then becomes down-regulated (negative NES) in the sensitive state, yet “recovers” and again become up-regulated (positive NES) in the resistant state (we call this “up-down-up” behavior), then such sub-network is important in Enzalutamide-resistance and could potentially constitute a functional marker and a therapeutic vulnerability. To identify such sub-networks and establish the significance of this change, we defined two gene expression query signatures (i) signature between intact and sensitive phenotype; and (ii) signature between sensitive and resistant phenotype. These signatures were defined utilizing two-tailed Welch t-test and implemented using the R *t.test* function<sup>142</sup>.

To identify sub-networks with such “up-down-up” behavior, we evaluated their enrichment in the “intact to sensitive” signature (i.e., looking for “up-down” behavior, corresponding to the down-regulation as a result of response to Enzalutamide) and enrichment in the “sensitive to resistant” signature (i.e., looking for “down-up” behavior, corresponding to the subsequent up-regulation as a result of resistance to Enzalutamide). To achieve this, we first estimated pathway activity levels and TR activity levels in each signature and overlaid them with our mechanism-centric regulatory network relationships/structure to identify parts of the network that exercise “up-down-up” behavior, as described above. To estimate if such “up-down-up” changes were statistically significant, we performed pathway-on-pathway and TR-on-TR GSEA, where pathways from “intact to sensitive” signature were compared to pathways from “sensitive to resistant” signature (same for the TR programs). Parts of the network with significant negative enrichment in “intact to sensitive” signature and significant positive enrichment in “sensitive to resistant” signature (GSEA  $p$  value  $< 0.001$ ) were utilized for Network mining step II.

#### Network mining II: prioritization of upstream regulatory programs

**Variance Inflation factor analysis.** Sub-networks identified in “Network mining I” include molecular pathways and their potential upstream TR programs. Such TR programs might exercise multi-collinearity in their effect on the pathway and could obstruct further statistical analysis (by making results not interpretable), yet deserve to remain in the analysis (as opposed to simply being eliminated). First, to check for multi-collinearity among TRs, we subjected the activity level of these TRs to Variance Inflation Factor analysis (VIF)<sup>57</sup> in the SU2C East Coast cohort. VIF runs a multivariable regression analysis, iteratively using each TR (activity vector) as a response variable and activity vectors from the rest of the TRs as predictor variables. The percentage of variation that the predictor variables could explain about the response variable is defined by the coefficient of determination,  $R^2$ , where higher  $R^2$  values indicate a higher degree of multi-collinearity and VIF is defined as  $1/(1 - R^2)$ . Typically, the multi-collinearity is observed if  $VIF > 10$ . VIF analysis was implemented utilizing the *vif* function from the R *usdm* package<sup>144</sup>.

**PLS regression analysis.** To address TR multi-collinearity, we developed a Partial Least Squares (PLS) -inspired method. To prioritize the effect of TR programs on a specific pathway  $i$ , our approach considers TR activity vectors  $TR_{t\ activity}$ ,  $t=1..m$  (where  $m$  is the number of TRs upstream of a specific pathway  $i$ ) as predictor variables and utilizes a pathway  $i$  activity vector  $P_{i\ activity}$  as a response variable. TR activity vectors are then regressed (linear regression) on the pathway vector so



that their  $\beta$  coefficients (slopes), indicating the effect of each TR on a pathway  $i$ , are denoted as weights  $w_t$ . Next, utilizing the TR activity vectors and weights associated with each TR, first latent variable  $LV1$  is defined as:

$$LV1 = \sum_{t=1}^m TR_{t \text{ activity}} * w_t \quad (4)$$

Further, the contribution (also referred to as loadings) of each TR on the  $LV1$  is determined through a multivariable regression analysis, where the activity vectors of all the transcriptional regulators  $TR_t \text{ activity}$  are utilized as independent variables and the  $LV1$  is utilized as a dependent variable. The  $\beta$  coefficients associated with each TR in this multivariable analysis, indicating the contribution of each  $TR_t \text{ activity}$  to  $LV1$ , adjusted for the effect of all other TRs, are denoted as *loadings*. Loadings are most often utilized in social science analyses.

This latent variable  $LV1$  is then “subtracted” from the TR activity vectors and the pathway  $i$  activity vector, leaving the residuals to be utilized for defining the next latent variable. In particular, the first latent variable  $LV1$  is utilized as an independent variable to be regressed on the activity vectors of each TR program  $TR_t \text{ activity}$ , as well as activity of the molecular pathway  $P_i \text{ activity}$ , so that the residuals from this analysis explain amount of information that has not been explained by  $LV1$ . The residuals are then utilized to define the second latent variable  $LV2$  in the similar fashion. This process is repeated until latent variables can explain a significant amount of information about a pathway  $i$ . PLS was implemented utilizing the *plsreg1* function from the R *plsdepot* package<sup>145</sup>.

**PLS-inspired circle of correlation analysis.** Identified latent variables do not express collinearity or multi-collinearity and are utilized as axes to build a “circle of correlation”, which depicts the association of TR programs and a specific pathway  $i$  (defined as arrows on the circle of correlation) to each latent variable. In particular, axes of the circle of correlation depict Pearson correlation  $r$  values, defined between latent variables and TR/pathway activity vectors. Each TR and a pathway  $i$  are indicated as arrows on the circle of correlation, with  $x$  and  $y$  coordinates that correspond to the values of Pearson correlation between their vectors and the latent variables.

To identify TRs that effect a specific pathway  $i$  as a group, we developed a method that utilized unsupervised hierarchical clustering on the degree of closeness (angle) between TR and pathway arrows so that TRs in high proximity to one another (thus having similar effects on latent variables) are grouped as they express simultaneous effect on the pathway  $i$ . In particular, for each TR and pathway arrow we first calculated their angle of inclination i.e.,  $(\cos^{-1} \theta)$ . To calculate the  $\cos^{-1} \theta$  we utilized R *acos* function<sup>142</sup>. Following this, angle of inclination in radian was converted to a degree using the *rad2deg* function from the R *rCAT* package<sup>146</sup>. To determine the degree of closeness, we subtracted the angle of inclination of each TR arrow from angle of inclination of a pathway  $i$  arrow. These degrees of closeness for TRs were then subjected to hierarchical clustering, which identified groups of TR programs with similar effects on the pathway  $i$ . For hierarchical clustering we utilized the R *hclust* function<sup>142</sup>.

**Prioritizing TR groups.** The TR groups/clusters (which also include groups with one TR) are then “prioritized” based on their effect on a pathway  $i$  using “effect scores”, which are defined as a combination of (i) degree of closeness between a TR group/cluster and a pathway  $i$  on the circle of correlation; (ii) association (i.e., Pearson correlation  $r$ ) between a TR group/cluster and each evaluated latent variable; and (iii) edge weight between a TR group/cluster and a pathway  $i$  from the TR-2-PATH mechanism-centric network reconstruction step. For clusters that contained more than one TR, average values for all TRs in that cluster were considered. Each of these categories assigned a “rank” for

each cluster and then ranks were combined (using geometric mean) to define the final effect score for each cluster. Geometric mean was implemented utilizing the *geometric.mean* function from the R *psych* package<sup>147</sup>.

**Validation in independent cohorts and Enzalutamide specificity analysis.** For validation and negative control analysis, we utilized refs. 19, 33, SU2C West Coast<sup>70,71</sup>, and PROMOTE<sup>34</sup> cohorts. Clinical characteristics and data normalization for these cohorts are described above and in Supplementary Data 1.

In He et al., (i.e., single-cell profiles), we reproduced data analysis performed in the original manuscript<sup>19</sup>. In particular, we applied UMAP<sup>148</sup> dimensionality reduction technique on single-cell Transcripts Per Million (TPM) data for each sample of the selected patient. We then utilized the AR activity and CK8 and CD45 expression on the UMAP projected data to identify adenocarcinoma cell clusters. Next, we estimated NME2 TR activity and MYC pathway activity on a single-cell level, in a manner similar to the single-sample analysis (as described above) and compared their activities between adenocarcinoma cells and the rest of the cells utilizing one-tailed Welch t-test, using the *t.test* function in R<sup>142</sup>.

In Abida et al.<sup>33</sup>, we subjected the cohort samples to a single-sample pathway and single-sample TR analysis to estimate activity levels of MYC pathway and NME2 TR program across all samples. For Enzalutamide-associated subset, we first performed Cook’s distance<sup>149</sup> analysis to identify outliers that can influence the regression analysis results (no outliers identified) utilizing R *cooks.distance* function<sup>142</sup>. Following this, we performed association analyses between activity vectors of NME2 TR and MYC pathway using the R *cor.test* function<sup>142</sup>. Next to identify patients with high-NME2 TR and high-MYC pathway activities in Enzalutamide-associated subset, we performed hierarchical and kmeans clustering on MYC pathway and NME2 TR activity vectors. For Abiraterone-associated subset, we also performed the Cook’s distance analysis (one outlier identified and removed) to identify outliers, followed by identification of patients with high-NME2 TR and high-MYC pathway activities. To identify patients with high-NME2 TR and high-MYC pathway activities, we applied the same thresholds that were estimated in the Enzalutamide-associated subset. Hierarchical clustering was implemented using the R *hclust* function<sup>142</sup> and kmeans clustering was performed using the R *kmeans* function<sup>142</sup> and identified two clusters of patients (i) patients with high-NME2 activity and high-MYC pathway activity and (ii) the rest of the patients (e.g., patients with low-NME2 and low-MYC pathway activity; patients with low-NME2 and high-MYC pathway activity; and patients with high-NME2 and low-MYC pathway activity). Further, to evaluate the difference in treatment response between the two identified groups, we utilized Kaplan-Meier survival analysis<sup>37</sup> and Cox proportional hazards model analysis<sup>38</sup>, where treatment-associated disease progression (as described above) was utilized as the clinical end-points, as defined in ref. 33. For Kaplan-Meier survival analysis, we utilized the *Surv* and the *ggsurvplot* functions from R *survival*<sup>150</sup> and *survminer*<sup>151</sup> packages, respectively. The Cox proportional hazards model analysis was adjusted for age and Gleason score and utilized the *coxph* function from the R *survival* package<sup>150</sup>.

In SU2C West Coast<sup>70,71</sup> cohort, similar to analysis on Abida et al., we subjected the cohort samples to a single-sample pathway and single-sample TR analysis to estimate activity levels of MYC pathway and NME2 TR program across all samples. As above, we first performed Cook’s distance analysis<sup>149</sup> to identify outliers (three outliers identified and removed) using R *cooks.distance* function followed by performing association analyses between activity vectors of NME2 TR and MYC pathway using the R *cor.test* function. Next, to identify patients with high-NME2 TR and high-MYC pathway activities we utilized hierarchical and kmeans clustering on MYC pathway and NME2 TR activity vectors. Hierarchical clustering was implemented using R *hclust*

function and kmeans clustering was implemented using the R *kmeans* function and identified two clusters of patients (i) patients with high-NME2 activity and high-MYC pathway activity and (ii) the rest of the patients (e.g., patients with low-NME2 and low-MYC pathway activity; patients with low-NME2 and high-MYC pathway activity; and patients with high-NME2 and low-MYC pathway activity). Further, to evaluate the difference in treatment response between the two identified groups, we utilized Kaplan-Meier survival analyses<sup>37</sup> and Cox proportional hazards model analysis<sup>38</sup>, where treatment-associated disease progression (as described earlier) was utilized as the clinical endpoints. For Kaplan-Meier survival analysis, we utilized the *Surv* and the *ggsurvplot* functions from the R *survival*<sup>150</sup> and *survminer*<sup>151</sup> packages respectively. The Cox proportional hazards model analysis was adjusted for race, Mstage, age and metastatic site and utilized the *coxph* function from the R *survival* package<sup>150</sup>.

In PROMOTE<sup>34</sup> cohort, we first performed Cook's distance analysis<sup>149</sup> to identify outliers as above (two outliers identified and removed) using R *cooks.distance* function. Since PROMOTE cohort has binary outcomes (responders vs non-responders), to evaluate the ability of NME2 TR and MYC pathway activities to classify patients based on their binary response to Abiraterone treatment, we performed ROC analysis using a multiplicative logistic regression model<sup>152</sup>, where the product of activity level of the NME2 TR program and activity level of the MYC pathway was utilized as predictor (independent) variable and responder/non-responder classification was utilized as response (dependent) variable. ROC curves were evaluated using area under the curve (AUROC), with AUROC = 0.5 being a random classifier. The logistic regression analysis was implemented using the R *glm* function<sup>142</sup> and ROC analysis was implemented using the *roc* function from the R *pROC* package<sup>153</sup>.

**Comparison to markers of aggressiveness and therapeutic response.** To compare the ability of MYC and NME2 to predict Enzalutamide resistance to the predictive ability of known transcriptomic and genomic markers of aggressiveness and therapeutic response we utilized patients from Enzalutamide-associated Abida et al.<sup>33</sup> cohort (as described above). In particular, comparisons were done in two ways: (i) comparison between high-NME2 and high-MYC pathway patients and the rest of the patients ("others"), as described above using two-tailed Welch t-test (for transcriptomic markers) and Fisher exact test<sup>154</sup> (for genomic markers); and (ii) direct independent association with the Enzalutamide-associated disease progression using Cox proportional hazards model. For transcriptomic markers, we utilized their gene expression/normalized counts. For genomic markers, we utilized genomic alterations (obtained from cBioportal), including deep and shallow deletions, gains, and amplifications, as available in cBioportal.

Two-tailed Welch t-test was implemented using the R *t.test* function<sup>142</sup>, Fisher exact test<sup>154</sup> was implemented using the R *fisher.test* function<sup>142</sup>, and Cox proportional hazards model analysis was implemented using the *coxph* function from the R *survival* package<sup>150</sup>.

**Comparative analysis of gene-centric computational methods.** To evaluate if TR-2-PATH mechanism-centric predictions (i.e., activity levels of NME2 TR and MYC pathway) outperform predictive ability of commonly used gene-centric methods, we compared TR-2-PATH to differential expression analyses, Random (survival) Forests (RF)<sup>110</sup>, and Support Vector Machine (SVM)<sup>111</sup> methods all utilized on the Enzalutamide-associated Abida et al.<sup>33</sup> cohort. For differential gene expression analysis, we considered genes that were differentially expressed between the three phenotypes (Intact, EnzaSens, and EnzaRes) in the mining step I and considered genes at (i) Welch t-test  $p$  value < 0.05; (ii) top 470 differentially expressed genes (comparable to the total number of target genes and pathway genes used for activity estimation) and *not excluding* target/pathway genes from NME2 TR and MYC pathway; (iii) top 470 differentially expressed genes,

*excluding* target/pathway genes from NME2 TR and MYC pathway. For RF and SVM analysis, we utilized 470 genes from (iii) to avoid overfitting and then selected top 10 most significant genes/features from the outputs. Final gene list from each of these analyses were utilized to cluster patients using hierarchical and kmeans clustering (as above), and then subjected these groups to Kaplan-Meier survival analysis<sup>37</sup> and Cox proportional hazards model analysis<sup>38</sup>. For Kaplan-Meier survival analysis, we utilized the *Surv* and the *ggsurvplot* functions from the R *survival*<sup>150</sup> and the *survminer*<sup>151</sup> packages, respectively. Additionally, for Cox proportional hazards model analysis, we utilized the *coxph* function from the R *survival* package<sup>150</sup>. For adjusted Cox proportional hazards model analysis, the model was adjusted for age and Gleason score. Random (survival) Forests were constructed utilizing *rfsrc* function from R *randomForestSRC* package<sup>155</sup>. The tuning parameters for Random (survival) Forests included (i) the maximum number of trees (i.e., "ntrees"), (ii) the number of variables assessed at each split (i.e., "mtry"), and (iii) maximum number of samples in the terminal (leaf) nodes (i.e., "nodesize"). The optimization of mtry and nodesize variables was performed utilizing *tune* function from R *randomForestSRC* package, which determined optimal value for mtry as 100 and nodesize as 5 and iterations of ntrees converged to a stable C-index around 3000, thus 3000 was selected as an optimal value for ntrees. For SVM, we utilized *fit* function from R *rminer* package<sup>156</sup> with default parameters.

**Statistical analysis and data visualization for computational methods.** Statistical analysis was performed using R studio version 4.0.4 for statistical computing. For differential gene expression we utilized one-tailed Welch t-test, with t-values and  $p$  values reported, corrected for multiple hypotheses testing. Linear regression analysis was utilized to evaluate relationships between TR programs and molecular pathways (i.e., network edges); the significance that the slope coefficient is non-zero was estimated using two tailed t-test, with multiple hypothesis testing. To ensure the edges identified were robust to sampling noise, 100 bootstrapped networks were generated, and each edge was assigned a weight, defined as the number of times an edge identified in the original network was also identified across the bootstrapped networks while maintaining the same direction of the relationship (positive/negative) between a particular TR program and a particular molecular pathway, across all 100 bootstrapped networks. To evaluate compositional similarity between MYC pathway genes and target genes of AR/NME2 transcriptional regulatory programs, we utilized Jaccard similarity index, defined as:

$$\text{Jaccard similarity index} = \frac{\text{number of common targets between A and B}}{\text{total number of targets in A and B}} \quad (5)$$

where  $A$  depicts MYC pathway genes and  $B$  depicts target genes of either NME2 TR or AR TR.

Kaplan-Meier survival analysis was utilized to estimate difference in treatment response between two patient groups, with log-rank test utilized for significance. Cox proportional hazards model was utilized to evaluate association with therapeutic response and its significance was reported as hazard ratio, confidence interval, hazard  $p$  value and Wald test. Further, to compare the ability of MYC and NME2 to predict Enzalutamide resistance to the predictive ability of known transcriptomic and genomic markers of aggressiveness and therapeutic response we utilized two-tailed Welch t-test (for continuous variables) and Fisher exact test (for categorical variables). We utilized the *geom\_violin* and the *geom\_boxplot* function from the *ggplot2*<sup>157</sup> in R for data visualization.

## Experimental methods

**Generation of Enzalutamide-resistant cell lines.** LNCaP (clone FDG) and C42B cells were purchased from ATCC and were grown in RPMI 1640 media (GIBCO # 11875093) supplemented with 10% Fetal Bovine Serum (FBS, Corning Cat#35-011-CV) and maintained at 37 °C and 5% CO<sub>2</sub>. Enzalutamide powder was purchased from Sellekchem (cat #S1250) and re-suspended in DMSO. Cells were plated in 6-well plates and treated either with DMSO, or with Enzalutamide (20 μM), refreshed every 4 days for up to 3 months until the resistance emerged. RNA from cells was extracted on indicated days using the methods described below.

**Generation of inducible shNME2 cell line and NME2 CRISPR KO cell line.** We constructed C42B EnzaRes Dox/shNME2 cells by infecting C42B EnzaRes cells with pLKO-Tet-On-shNME2 virus followed by continuous puromycin selection. The lentiviral shRNA plasmid was constructed by inserting target shNME2 oligonucleotide sequence into Tet-pLKO-puro (Addgene, #21915) plasmid. For generating NME2 CRISPR KO cells, sgRNA oligonucleotide sequences were cloned into pLentiCRISPR-v2 plasmid (Genscript).

For both NME2 KD and KO cells, virus was generated as follows: plasmid DNA was extracted using a DNA extraction kit (Vazyme, DC112-01). The lentivirus was packaged by transfecting the plasmids with packaging vectors (psPAX and pMD2.G) and Lipofectamine 2000 (Invitrogen, #11668-019) in Opti-MEM media (Gibco) into HEK293T cells. Afterward, the virus supernatant was collected, filtered with a 0.45 μm strainer, concentrated with PEG6000 (Sigma, #81253), resolved in PBS and then aliquoted for subsequent transfection. Cells were infected with viruses and selected for 72 h with puromycin (2 μg/mL, Sigma-Aldrich, P8833). The following oligonucleotide sequences were used: TCATGGCAGTGATTTCAGTAAA for the shNME2 sequence (flanked by start ACCGGT and end GAATTC sequences for a total of 33 bp) and CACCTTCATCGCCATCAAGC for sgNME2\_V1 and GCACTCACCATGGCCACAAC for sgNME2\_V2. Cells transfected with the second sgRNA guide, sgNME2\_V2, were used for further *in vitro* studies.

**RNA extraction, cDNA preparation, transcript knockdown, and qRT-PCR analysis.** RNA was isolated from cells by the Quick-RNA miniprep kit (Zymogen# R1054) and digested with DNase (provided in the kit). cDNA was synthesized from 1 μg RNA, using an All-in-One 5X RT-master mix (Abm # G592), per the manufacturer's protocol. qRT-PCR was carried out on the StepOne Real-Time PCR system (Applied Biosystems) using gene-specific primers designed with Primer-BLAST and synthesized by IDT Technologies. ON-TARGETplus SMARTpool (cat# L005102-00-0005) was obtained from Dharmacon and was used at 100 nmol/L. Cells were transfected in 6-well plates at a density of 100,000 cells per well using Lipofectamine RNAiMax (Invitrogen #13778075), according to the manufacturer's protocol. RNA was extracted and converted to cDNA as described above. qRT-PCR data were analyzed using the relative quantification method using 18sRNA as an internal reference, and plotted as average fold-change compared with DMSO or the non-targeting siRNA (i.e., Relative Quantity or RQ). Determination of transcript levels was carried out using Fast SYBR Green Master Mix (Invitrogen), using specific primer sets for c-MYC: c-MYC (F) 5'- CCTGGTCTCCATGAGGAGAC-3'; c-MYC (R) 5'- CAGACTCTGACCTTTTGCAGG.

**Evaluating expression of AR.** To evaluate the expression of AR in Enzalutamide-naïve and Enzalutamide-resistant conditions, we utilized cells from LNCaP and C42B cell lines under Enzalutamide-naïve and Enzalutamide-resistant conditions (as described above) and determined the expression level of AR under both conditions using qRT-PCR assay (described above). The specific set of primers used for AR includes: AR (F) 5'- TCTTGTCTCTTCGGAAATGTT-3'; AR (R) 5'- AAGCCTCTCCTTCTCTCTGTA-3'.

**Evaluating expression of NME2 in Enzalutamide-resistant vs Enzalutamide-naïve cells.** To evaluate the expression of NME2 in Enzalutamide-naïve and Enzalutamide-resistant conditions, we utilized cells from LNCaP and C42B cell lines under Enzalutamide-naïve and Enzalutamide-resistant conditions (as described above) and determined the expression level of NME2 under both conditions using qRT-PCR assay (as described above). The specific set of primers used for NME2 is: NME2 (F) 5'- AGGATTCGCCTTGTGGTCTG-3'; NME2 (R) 5'- CGGCAAAGAATGGACGGTCCTT-3'.

**Knockdown of NME2.** Two different siRNA against NME2 (siNME2#1 AAUAAGAGGUGGACACAAC; siNME2#2 CUGAAGAACCUGAAGCA), or non-targeting control (siScram) were obtained from Dharmacon and used at 100 nmol/L.

**Drug response curves.** Treatment-naïve or Enzalutamide-resistant C42B and LNCaP cells were plated in 96-well plates at a density of 5000 cells/well. A day later, cells were treated with indicated doses of drugs (6 technical replicates/dose) and were incubated for 96 h. After 96 h, cell viability was determined using CCK-8 reagent kit (Bimake # 34304) according to manufacturer's protocol. Graphs were plotted using GraphPad PRISM software and IC<sub>50</sub> values were determined. The experiment was repeated three times.

**Colony formation assay.** C42B-EnzaRes and LNCaP-EnzaRes cells were treated with siRNAs, or Enzalutamide or MYCi975, or a combination of both (Enzalutamide + MYCi975) as indicated for 24 h. Cells were then trypsinized (TrypLE Express, GIBCO # 12604013) and live cells were counted using Trypan Blue exclusion method. Five thousand live cells per condition were plated in 60 mm plates (in duplicates) and were grown in complete media with or without drugs as indicated. Cells were allowed to form colonies for two weeks, following which, the resulting colonies were fixed using 4% paraformaldehyde (Sigma Aldrich) for 30 min. Cells were washed with PBS and stained with 0.25% Crystal Violet for 10 min. Plates were washed with PBS to wash off unbound Crystal Violet. The percentage of area occupied by colonies was calculated by ImageJ software and graphs were plotted in GraphPad PRISM software. The experiment was repeated three times.

**Boyden chamber-based migration assay.** Boyden transwell chambers for migration were purchased from Corning Inc. (Cat# 353097). C42B-EnzaRes cells were transfected with siRNAs as described above, or treated with 2 μM MYCi975<sup>30</sup> and/or 10 μM Enzalutamide for 24 h. Cells were then trypsinized and re-suspended in serum-free media at a density of 100,000 cells/chamber, in the upper chamber in duplicates. The lower chamber was supplied media containing 10% FBS (in the case of NME2 knockdown), or supplemented with 10% FBS containing DMSO, or 2 μM MYCi975 (ref) and/or 10 μM Enzalutamide wherever indicated. Cells were allowed to migrate for 48 h, following which they were fixed using 4% paraformaldehyde for 30 min at room temperature. Wells were washed and membranes were stained with 0.25% crystal violet made in 25% methanol for 20 mins. Wells were washed thoroughly, and images were taken on an Echo-Revolve-R4 microscope. For quantification, cells stained with crystal violet were quantified using the image processing software suite ImageJ. Graphs were plotted using GraphPad Prism software. The experiment was repeated three times.

**Western Blot analysis.** Cells were lysed in RIPA (Sigma) lysis buffer containing 1x Halt™ Protease and Phosphatase Inhibitor Cocktail (ThermoFisher). Protein concentration was measured by Bio-Rad Bradford reagent. Protein samples were prepared by addition of 4x Laemmli Sample buffer (Bio-Rad) and 2-mercaptoethanol (Bio-Rad) and resolved on 4–12% SDS-PAGE (Sodium dodecyl sulfate–polyacrylamide) gels, which were subsequently transferred to



PVDF membranes (Bio-Rad) using Trans-Blot Turbo transfer buffer (Bio-Rad) and a Trans-Blot Turbo Transfer System (Bio-Rad). Membranes were blocked for 1 h at room temperature with 5% blotting-grade blocker non-fat dry milk (Bio-Rad), followed by overnight 4 °C incubation with the appropriate primary antibody and 1 h room temperature incubation with an anti-rabbit or anti-mouse IgG (H + L)-HRP conjugate (Bio-Rad) secondary antibody. Blots were imaged using Supersignal West Femto Maximum Sensitivity Substrate detection system (Thermo) and the ChemiDoc Imaging System (Bio-Rad). The following primary antibodies were used: c-Myc (Y69) (Abcam #ab32072, 1:1000 dilution), NME2 (4G7A8) (Abcam #ab204958, 1:1000 dilution), GAPDH (Cell Signaling Technology #3683, 1:5000 dilution), Actin (Cell Signaling Technology #5125 S, 1:5000 dilution).

**In vivo studies.** All animal experiments and procedures were performed in compliance with ethical standards and the approval of Northwestern University Animal Care and Use Committee (IACUC). All mice were obtained from Jackson Laboratory. All the mice used in this study were maintained in a pathogen-free animal barrier facility. The mice used in this study were housed under standardized conditions in a dedicated animal facility accredited by AAALAC, International, in compliance with institutional and ethical guidelines for the humane treatment of animals. The housing facility maintained a 12-h light/dark cycle, with lights on at 6:00 A.M. and off at 6:00 P.M. The ambient temperature was consistently maintained at  $22 \pm 2$  °C, and humidity was kept within the range of 45% to 55%. All animals were housed in standard polypropylene cages with access to standard rodent chow and water ad libitum. The mice were acclimatized to these conditions for a minimum of 7 days before the commencement of any experimental procedures. All the experiments were initiated with mice of age 6 to 8 weeks. C42B EnzaRes Dox/shNME2 inducible cells were suspended at a concentration of 10 million cells/mL in 50% BD Matrigel–50% PBS and 100  $\mu$ L of this solution for a total of 1 million cells subcutaneously injected into the flanks of FVB mice. Mice were grouped and treatment was started when the tumor size reached around 150 to 180 mm<sup>3</sup>. The maximal tumor size burden permitted by Northwestern IACUC for mice is 2000 mm<sup>3</sup> and this was not exceeded for any of the animals involved in the study. Sex was not considered in the study design since these are prostate cancer models that only grow in male mice.

Enzalutamide was dissolved in 5% DMSO (Sigma-Aldrich #276855), 10% Solutol/Kolliphor HS15 (Sigma-Aldrich #42966) and 85% DPBS (Gibco #14190-144) at a concentration of 1.25 mg/mL for I.P. administration. Doxycycline hyclate (Sigma-Aldrich #D9891) was dissolved in drinking water at a concentration of 2 mg/mL and the solution supplemented with 5% sucrose w/v.

**Statistical analysis for in vitro and in vivo data.** All statistical tests in in vitro analyses were performed using one-tailed Welch t-test and implemented using the R *t.test* function<sup>42</sup>. For in vivo data, statistical comparison between groups was performed using two-way ANOVA, implemented in GraphPad Prism 9 software. Data are presented as mean  $\pm$  standard error of the mean (S.E.M.) and statistical significance was defined by  $P < 0.05$  (two-tailed). In figures, “\*” indicates  $P < 0.05$ , “\*\*” indicates  $P < 0.01$ , “\*\*\*” indicates  $P < 0.001$ , “\*\*\*\*” indicates  $P < 0.0001$ , and not significant “ns” indicates  $P \geq 0.05$  for the indicated pairwise comparison.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The mechanism-centric CRPC-specific network is available in Supplementary Data 2. Data supporting the findings of this study were

obtained from (a detailed description of the datasets is available in Supplementary Data 1): (i) dbGaP: 1. Stand Up To Cancer East Coast Prostate Cancer cohort<sup>32,33</sup>, RNA-sequencing data, [phs000915.v2.p2](https://doi.org/10.1038/s41467-024-44686-5). Access could be obtained through dbGap portal. 2. PROMOTE Prostate Cancer cohort<sup>34</sup>, RNA-sequencing data, [phs001141.v1.p1](https://doi.org/10.1038/s41467-024-44686-5). Access could be obtained through dbGap portal. 3. Beltran et al.<sup>68</sup>, RNA-sequencing data, [phs000909.v1.p1](https://doi.org/10.1038/s41467-024-44686-5). Access could be obtained through dbGap portal. 4. SU2C West Coast Prostate Cancer cohort<sup>70,71</sup>, RNA-sequencing data, [phs001648.v2.p1](https://doi.org/10.1038/s41467-024-44686-5), downloaded from [<https://portal.gdc.cancer.gov/projects/WCDT-MCRPC>]. Access could be obtained through dbGap portal. (ii) cBioPortal: 1. Abida et al.<sup>33</sup>, RNA-sequencing data [[https://github.com/cBioPortal/datahub/tree/master/public/prad\\_su2c\\_2019](https://github.com/cBioPortal/datahub/tree/master/public/prad_su2c_2019)]. Data could be downloaded directly from cBioPortal. (iii) GEO: 1. Kregel et al.<sup>53</sup>, microarray gene expression data, [GSE78201](https://doi.org/10.1038/s41467-024-44686-5). Data could be downloaded directly from GEO. 2. Kumar et al.<sup>69</sup>, microarray gene expression data, [GSE77930](https://doi.org/10.1038/s41467-024-44686-5). Data could be downloaded directly from GEO. (iv) Broad Single Cell Portal: 1. He et al.<sup>19</sup>, RNA-sequencing data [[https://singlecell.broadinstitute.org/single\\_cell/study/SCP1244/transcriptional-mediators-of-treatment-resistance-in-lethal-prostate-cancer](https://singlecell.broadinstitute.org/single_cell/study/SCP1244/transcriptional-mediators-of-treatment-resistance-in-lethal-prostate-cancer)]. Data could be downloaded directly from Broad Single Cell Portal. (v) Manuscript Supplemental Information 1. Alumkal et al.<sup>12</sup>, RNA-sequencing data: <https://doi.org/10.1073/pnas.1922207117>. Data could be obtained from the Supplementary Information in ref. 12. Hallmark and C2 pathway gene sets were obtained from the molecular signatures database (MSigDB 3.0) <https://www.gsea-msigdb.org/gsea/msigdb>. Source data are provided with this paper: experimental source data are provided as a Source Data file and computational source data are provided as a part of TR-2-PATH R software package (<https://github.com/mitrofanova-lab/TR2PATH>) under <https://doi.org/10.5281/zenodo.10368948><sup>58</sup>.

### Code availability

TR-2-PATH is released as an R software package in GitHub for community utilization (<https://github.com/mitrofanova-lab/TR2PATH>) under <https://doi.org/10.5281/zenodo.10368948><sup>58</sup>.

### References

- Huggins, C. & Hodges, C. V. Studies on prostatic cancer. I. The effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate. *Cancer Res.* **1**, 293–297 (1941).
- Perlmutter, M. A. & Lepor, H. Androgen deprivation therapy in the treatment of advanced prostate cancer. *Rev. Urol.* **9**, S3–S8 (2007).
- Shen, M. M. & Abate-Shen, C. Molecular genetics of prostate cancer: new prospects for old challenges. *Genes Dev.* **24**, 1967–2000 (2010).
- Salonen, A. J. et al. Finnish multicenter study comparing intermittent to continuous androgen deprivation for advanced prostate cancer: interim analysis of prognostic markers affecting initial response to androgen deprivation. *J. Urol.* **180**, 915–919 (2008). discussion 919–920.
- Waltering, K. K., Urbanucci, A. & Visakorpi, T. Androgen receptor (AR) aberrations in castration-resistant prostate cancer. *Mol. Cell. Endocrinol.* **360**, 38–43 (2012).
- Schweizer, M. T. & Yu, E. Y. Persistent androgen receptor addiction in castration-resistant prostate cancer. *J. Hematol. Oncol.* **8**, 128 (2015).
- Scher, H. I. & Sawyers, C. L. Biology of progressive, castration-resistant prostate cancer: directed therapies targeting the androgen-receptor signaling axis. *J. Clin. Oncol.* **23**, 8253–8261 (2005).
- de Bono, J. S. et al. Abiraterone and increased survival in metastatic prostate cancer. *N. Engl. J. Med.* **364**, 1995–2005 (2011).



9. Tran, C. et al. Development of a second-generation antiandrogen for treatment of advanced prostate cancer. *Science* **324**, 787–790 (2009).
10. Clegg, N. J. et al. ARN-509: a novel antiandrogen for prostate cancer treatment. *Cancer Res.* **72**, 1494–1503 (2012).
11. Tagawa, S. T. & Beltran, H. Abiraterone plus prednisone improves survival in metastatic castration-resistant prostate cancer. *Asian J. Androl.* **13**, 785–786 (2011).
12. Alumkal, J. J. et al. Transcriptional profiling identifies an androgen receptor activity-low, stemness program associated with enzalutamide resistance. *Proc. Natl Acad. Sci.* **117**, 12315 (2020).
13. Tagawa, S. T. et al. Survival outcomes in patients with chemotherapy-naïve metastatic castration-resistant prostate cancer treated with enzalutamide or abiraterone acetate. *Prostate Cancer Prostatic Dis.* (2021). <https://doi.org/10.1038/s41391-021-00318-3>
14. Vander Ark, A., Cao, J. & Li, X. Mechanisms and approaches for overcoming enzalutamide resistance in prostate cancer. *Front Oncol.* **8**, 180–180 (2018).
15. Beer, T. M. et al. Enzalutamide in metastatic prostate cancer before chemotherapy. *N. Engl. J. Med.* **371**, 424–433 (2014).
16. Scher, H. I. et al. Increased survival with enzalutamide in prostate cancer after chemotherapy. *N. Engl. J. Med.* **367**, 1187–1197 (2012).
17. Blatt, E. B. & Raj, G. V. Molecular mechanisms of enzalutamide resistance in prostate cancer. *Cancer Drug Resist.* **2**, 189–197 (2019).
18. Zhang, Z. et al. Loss of CHD1 promotes heterogeneous mechanisms of resistance to AR-targeted therapy via chromatin dysregulation. *Cancer Cell* **37**, 584–598.e511 (2020).
19. He, M. X. et al. Transcriptional mediators of treatment resistance in lethal prostate cancer. *Nat. Med.* **27**, 426–433 (2021).
20. Taavitsainen, S. et al. Single-cell ATAC and RNA sequencing reveal pre-existing and persistent cells associated with prostate cancer relapse. *Nat. Commun.* **12**, 5307 (2021).
21. Chen, W. S. et al. Genomic drivers of poor prognosis and enzalutamide resistance in metastatic castration-resistant prostate cancer. *Eur. Urol.* **76**, 562–571 (2019).
22. Pellakuru, L. G. et al. Global levels of H3K27me3 track with differentiation in vivo and are deregulated by MYC in prostate cancer. *Am. J. Pathol.* **181**, 560–569 (2012).
23. Tomlins, S. A. et al. Integrative molecular concept modeling of prostate cancer progression. *Nat. Genet.* **39**, 41–51 (2007).
24. Koh, C. M. et al. MYC and prostate cancer. *Genes Cancer* **1**, 617–628 (2010).
25. Gurel, B. et al. Nuclear MYC protein overexpression is an early alteration in human prostate carcinogenesis. *Mod. Pathol.* **21**, 1156–1167 (2008).
26. Bai, S. et al. A positive role of c-Myc in regulating androgen receptor and its splice variants in prostate cancer. *Oncogene* **38**, 4977–4989 (2019).
27. Arriaga, J. M. et al. A MYC and RAS co-activation signature in localized prostate cancer drives bone metastasis and castration resistance. *Nat. Cancer* **1**, 1082–1096 (2020).
28. Qiu, X. et al. MYC drives aggressive prostate cancer by disrupting transcriptional pause release at androgen receptor targets. *Nat. Commun.* **13**, 2559 (2022).
29. Rebello, R. J., Pearson, R. B., Hannan, R. D. & Furic, L. Therapeutic approaches targeting MYC-driven prostate cancer. *Genes* **8**, 71 (2017).
30. Han, H. et al. Small-molecule MYC inhibitors suppress tumor growth and enhance immunotherapy. *Cancer Cell* **36**, 483–497.e415 (2019).
31. Llombart, V. & Mansour, M. R. Therapeutic targeting of “undruggable” MYC. *eBioMedicine* **75** (2022). <https://doi.org/10.1016/j.ebiom.2021.103756>
32. Robinson, D. et al. Integrative clinical genomics of advanced prostate cancer. *Cell* **161**, 1215–1228 (2015).
33. Abida, W. et al. Genomic correlates of clinical outcome in advanced prostate cancer. *Proc. Natl Acad. Sci.* **116**, 11428 (2019).
34. Wang, L. et al. A prospective genome-wide study of prostate cancer metastases reveals association of wnt pathway activation and increased cell cycle proliferation with primary resistance to abiraterone acetate-prednisone. *Ann. Oncol.* **29**, 352–360 (2018).
35. Abate-Shen, C. & Nunes de Almeida, F. Establishment of the LNCaP cell line – the dawn of an era for prostate cancer research. *Cancer Res.* **82**, 1689–1691 (2022).
36. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci.* **102**, 15545 (2005).
37. Kaplan, E. L. & Meier, P. Nonparametric Estimation from Incomplete Observations. *J. Am. Stat. Assoc.* **53**, 457–481 (1958).
38. Cox, D. R. Regression models and life-tables. *J. R. Stat. Soc.: Ser. B (Methodol.)* **34**, 187–202 (1972).
39. Aytes, A. et al. Cross-species regulatory network analysis identifies a synergistic interaction between FOXM1 and CENPF that drives prostate cancer malignancy. *Cancer Cell* **25**, 638–651 (2014).
40. Murphy, A. H. The Finley affair: a signal event in the history of forecast verification. *Weather Forecast.* **11**, 3–20 (1996).
41. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).
42. Nishimura, D. BioCarta. *Biotech. Softw. Internet Rep. Comput. Softw. J. Sci.* **2**, 117–120 (2001).
43. Fabregat, A. et al. The Reactome Pathway knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).
44. Liberzon, A. et al. The molecular signatures database hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
45. Alvarez, M. J. et al. Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nat. Genet.* **48**, 838–847 (2016).
46. Chatterjee, S., Hadi, A. & Price, B. Simple linear regression. *Regression analysis by example*, 21–51 (2006).
47. Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
48. Chow, M. T. & Luster, A. D. Chemokines in cancer. *Cancer Immunol. Res.* **2**, 1125–1131 (2014).
49. Dranoff, G. Cytokines in cancer pathogenesis and cancer therapy. *Nat. Rev. Cancer* **4**, 11–22 (2004).
50. Johnson, D. E., O’Keefe, R. A. & Grandis, J. R. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nat. Rev. Clin. Oncol.* **15**, 234–248 (2018).
51. Dechant, M. & Valerius, T. IgA antibodies for cancer therapy. *Crit. Rev. Oncol. Hematol.* **39**, 69–77 (2001).
52. Casey, S. C., Baylot, V. & Felsher, D. W. The MYC oncogene is a global regulator of the immune response. *Blood* **131**, 2007–2015 (2018).
53. Kregel, S. et al. Acquired resistance to the second-generation androgen receptor antagonist enzalutamide in castration-resistant prostate cancer. *Oncotarget* **7**, 26259–26274 (2016).
54. Arora, V. K. et al. Glucocorticoid receptor confers resistance to antiandrogens by bypassing androgen receptor blockade. *Cell* **155**, 1309–1322 (2013).
55. Smith, R. et al. Enzalutamide response in a panel of prostate cancer cell lines reveals a role for glucocorticoid receptor in enzalutamide resistant disease. *Sci. Rep.* **10**, 21750 (2020).
56. Culig, Z. & Santer, F. R. Androgen receptor signaling in prostate cancer. *Cancer Metastasis Rev.* **33**, 413–427 (2014).
57. Snee, R. D. Developments in linear regression methodology: 1959–1982: discussion. *Technometrics* **25**, 230–237 (1983).

58. Vatcheva, K. P., Lee, M., McCormick, J. B. & Rahbar, M. H. Multicollinearity in regression analyses conducted in epidemiologic studies. *Epidemiology* **6**, 227 (2016).
59. Kim, J. H. Multicollinearity and misleading statistical results. *Korean J. Anesthesiol.* **72**, 558–569 (2019).
60. Wold, H. Soft modeling: the basic design and some extensions. *Syst. Indirect Observ.* **2**, 343 (1982).
61. Mateos-Aparicio, G. Partial Least Squares (PLS) methods: origins, evolution, and application to social sciences. *Commun. Stat.—Theory Methods* **40**, 2305–2317 (2011).
62. Riou, J., Guyon, H. & Falissard, B. An introduction to the partial least squares approach to structural equation modelling: a method for exploratory psychiatric research. *Int. J. Methods Psychiatr. Res.* **25**, 220–231 (2016).
63. Lobaugh, N. J., West, R. & McIntosh, A. R. Spatiotemporal analysis of experimental differences in event-related potential data with partial least squares. *Psychophysiology* **38**, 517–530 (2001).
64. Worsley, K. J. An overview and some new developments in the statistical analysis of PET and fMRI data. *Hum. Brain Mapp.* **5**, 254–258 (1997).
65. Puts, G. S., Leonard, M. K., Pamidimukkala, N. V., Snyder, D. E. & Kaetzel, D. M. Nuclear functions of NME proteins. *Lab. Invest. J. Tech. Methods Pathol.* **98**, 211–218 (2018).
66. Thakur, R. K. et al. Metastases suppressor NM23-H2 interaction with G-quadruplex DNA within c-MYC promoter nucleic hyper-sensitive element induces c-MYC expression. *Nucleic Acids Res.* **37**, 172–183 (2009).
67. McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018).
68. Beltran, H. et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nat. Med.* **22**, 298–305 (2016).
69. Kumar, A. et al. Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. *Nat. Med.* **22**, 369–378 (2016).
70. Quigley, D. A. et al. Genomic hallmarks and structural variation in metastatic prostate cancer. *Cell* **174**, 758–769.e759 (2018).
71. Aggarwal, R. et al. Targeting adaptive pathways in metastatic treatment-resistant prostate cancer: update on the stand up 2 Cancer/Prostate Cancer Foundation-Supported West Coast Prostate Cancer Dream Team. *Eur. Urol. Focus* **2**, 469–471 (2016).
72. Lasko, T. A., Bhagwat, J. G., Zou, K. H. & Ohno-Machado, L. The use of receiver operating characteristic curves in biomedical informatics. *J. Biomed. Inf.* **38**, 404–415 (2005).
73. Rotinen, M. et al. ONECUT2 is a targetable master regulator of lethal prostate cancer that suppresses the androgen axis. *Nat. Med.* **24**, 1887–1898 (2018).
74. Tomlins, S. A. et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **310**, 644–648 (2005).
75. Cai, C., Wang, H., Xu, Y., Chen, S. & Balk, S. P. Reactivation of androgen receptor-regulated TMPRSS2:ERG gene expression in castration-resistant prostate cancer. *Cancer Res* **69**, 6027–6032 (2009).
76. Goel, S. et al. Transcriptional network involving ERG and AR orchestrates Distal-less homeobox-1 mediated prostate cancer progression. *Nat. Commun.* **12**, 5325–5325 (2021).
77. Li, Y. et al. Targeting cellular heterogeneity with CXCR2 blockade for the treatment of therapy-resistant prostate cancer. *Sci. Transl. Med.* **11**, eaax0428 (2019).
78. Xin, L. EZH2 accompanies prostate cancer progression. *Nat. Cell Biol.* **23**, 934–936 (2021).
79. Bowen, C. et al. Loss of NKX3.1 expression in human prostate cancers correlates with tumor progression. *Cancer Res.* **60**, 6111–6115 (2000).
80. Hagiwara, M. et al. MUC1-C activates the PBAF chromatin remodeling complex in integrating redox balance with progression of human prostate cancer stem cells. *Oncogene* **40**, 4930–4940 (2021).
81. Cyrta, J. et al. Role of specialized composition of SWI/SNF complexes in prostate cancer lineage plasticity. *Nat. Commun.* **11**, 5549 (2020).
82. Kim, M. Y. et al. High FOXM1 expression is a prognostic marker for poor clinical outcomes in prostate cancer. *J. Cancer* **10**, 749–756 (2019).
83. Brady, N. J. et al. Temporal evolution of cellular heterogeneity during the progression to advanced AR-negative prostate cancer. *Nat. Commun.* **12**, 3372 (2021).
84. Zhou, Q. et al. Key genes and pathways controlled by E2F1 in human castration-resistant prostate cancer cells. *Onco Targets Ther.* **12**, 8961–8976 (2019).
85. Papachristodoulou, A. et al. NKX3.1 localization to mitochondria suppresses prostate cancer initiation. *Cancer Discov.* **11**, 2316–2333 (2021).
86. Lachmann, A. et al. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* **26**, 2438–2444 (2010).
87. Chen, Y., Gutkind, J. S., Sreenath, T. L. & McLeod, D. G. TMPRSS2-ERG fusion, a common genomic alteration in prostate cancer activates C-MYC and abrogates prostate. *Oncogene* **27**, 5348–5353 (2008).
88. Kissick, H. T. et al. The transcription factor ERG increases expression of neurotransmitter receptors on prostate cancer cells. *BMC Cancer* **15**, 604 (2015).
89. Udager, A. M. et al. Concurrent nuclear ERG and MYC protein overexpression defines a subset of locally advanced prostate cancer: potential opportunities for synergistic targeted therapeutics. *Prostate* **76**, 845–853 (2016).
90. Armenia, J. et al. The long tail of oncogenic drivers in prostate cancer. *Nat. Genet.* **50**, 645–651 (2018).
91. Salmaninejad, A. et al. Molecular characterization of KRAS, BRAF, and EGFR genes in cases with prostatic adenocarcinoma; reporting bioinformatics description and recurrent mutations. *Clin. Lab.* **61**, 749–759 (2015).
92. Haddad, B. R. et al. Positive STAT5 protein and locus amplification status predicts recurrence after radical prostatectomy to assist clinical precision management of prostate cancer. *Cancer Epidemiol. Biomark. Prev.* **28**, 1642–1651 (2019).
93. Mollaoglu, G. et al. MYC drives progression of small cell lung cancer to a variant neuroendocrine subtype with vulnerability to aurora kinase inhibition. *Cancer Cell* **31**, 270–285 (2017).
94. Panja, S., Hayati, S., Epsi, N. J., Parrott, J. S. & Mitrofanova, A. Integrative (epi) genomic analysis to predict response to androgen-deprivation therapy in prostate cancer. *EBioMedicine* **31**, 110–121 (2018).
95. Hankey, W., Chen, Z. & Wang, Q. Shaping chromatin states in prostate cancer by pioneer transcription factors. *Cancer Res.* **80**, 2427–2436 (2020).
96. Zhang, Z. et al. Tumor microenvironment-derived NRG1 promotes antiandrogen resistance in prostate cancer. *Cancer Cell* **38**, 279–296 (2020). e279.
97. Kohrt, S. E. et al. Identification of genes required for enzalutamide resistance in castration-resistant prostate cancer cells in vitro. *Mol. Cancer Ther.* **20**, 398–409 (2021).
98. Verma, S. et al. Androgen deprivation induces transcriptional reprogramming in prostate cancer cells to develop stem cell-like characteristics. *Int. J. Mol. Sci.* **21**, <https://doi.org/10.3390/ijms21249568> (2020).
99. He, Y. et al. Androgen receptor splice variants bind to constitutively open chromatin and promote abiraterone-resistant

- growth of prostate cancer. *Nucleic Acids Res.* **46**, 1895–1911 (2018).
100. Korpala, M. et al. An F876L mutation in androgen receptor confers genetic and phenotypic resistance to MDV3100 (enzalutamide). *Cancer Discov.* **3**, 1030–1043 (2013).
101. Lee, G. T. et al. TCF4 induces enzalutamide resistance via neuroendocrine differentiation in prostate cancer. *PLoS One* **14**, e0213488 (2019).
102. Liang, Y. et al. Emergence of enzalutamide resistance in prostate cancer is associated with BCL-2 and IKKB dependencies. *Clin. Cancer Res.* **27**, 2340–2351 (2021).
103. Pilling, A. B. & Hwang, C. Targeting prosurvival BCL2 signaling through Akt blockade sensitizes castration-resistant prostate cancer cells to enzalutamide. *Prostate* **79**, 1347–1359 (2019).
104. Chen, X. et al. Inhibition of Rac1 reverses enzalutamide resistance in castration-resistant prostate cancer. *Oncol. Lett.* **20**, 2997–3005 (2020).
105. Lee, H. C. et al. YAP1 overexpression contributes to the development of enzalutamide resistance by induction of cancer stemness and lipid metabolism in prostate cancer. *Oncogene* **40**, 2407–2421 (2021).
106. Ning, S. et al. PD07-02 Targeting Wnt5a/Fzd2 signaling overcomes resistance to enzalutamide in advanced prostate cancer. *J. Urol.* **207**, e98 (2022).
107. Farah, E. et al. NOTCH signaling is activated in and contributes to resistance in enzalutamide-resistant prostate cancer cells. *J. Biol. Chem.* **294**, 8543–8554 (2019).
108. Liu, C. et al. Intracrine androgens and AKR1C3 activation confer resistance to enzalutamide in prostate cancer. *Cancer Res.* **75**, 1413–1422 (2015).
109. Guan, X. et al. Copy number loss of 17q22 is associated with enzalutamide resistance and poor prognosis in metastatic castration-resistant prostate cancer. *Clin. Cancer Res.* **26**, 4616–4624 (2020).
110. Nasejje, J. B., Mwambi, H., Dheda, K. & Lesosky, M. A comparison of the conditional inference survival forest model to random survival forests based on a simulation study as well as on two applications with time-to-event data. *BMC Med. Res. Methodol.* **17**, 115 (2017).
111. Drucker, H., Burges, C. J., Kaufman, L., Smola, A. & Vapnik, V. Support vector regression machines. *Adv. Neural Inf. Process. Syst.* **28**, 779–784 (1996).
112. Berberich, S. J. & Postel, E. H. PuF/NM23-H2/NDPK-B transactivates a human c-myc promoter-CAT gene via a functional nuclease hypersensitive element. *Oncogene* **10**, 2343–2347 (1995).
113. Postel, E. H., Berberich, S. J., Flint, S. J. & Ferrone, C. A. Human c-myc transcription factor PuF identified as nm23-H2 nucleoside diphosphate kinase, a candidate suppressor of tumor metastasis. *Science* **261**, 478–480 (1993).
114. Yao, Y. et al. PIWIL2 induces c-Myc expression by interacting with NME2 and regulates c-Myc-mediated tumor cell proliferation. *Oncotarget* **5**, 8466–8477 (2014).
115. Dang, ChiV. *Cell* **149**, 22–35 (2012). <em>MYC</em> on the Path to Cancer.
116. Dang, C. V. et al. The c-Myc target gene network. *Semin Cancer Biol.* **16**, 253–264 (2006).
117. Adhikary, S. & Eilers, M. Transcriptional regulation and transformation by Myc proteins. *Nat. Rev. Mol. Cell Biol.* **6**, 635–645 (2005).
118. Chen, H., Liu, H. & Qing, G. Targeting oncogenic Myc as a strategy for cancer treatment. *Signal Transduct. Target. Ther.* **3**, 5 (2018).
119. Castell, A. et al. A selective high affinity MYC-binding compound inhibits MYC:MAX interaction and MYC-dependent tumor cell proliferation. *Sci. Rep.* **8**, 10064 (2018).
120. Blackwell, T. K., Kretzner, L., Blackwood, E. M., Eisenman, R. N. & Weintraub, H. Sequence-specific DNA binding by the c-Myc protein. *Science* **250**, 1149–1151 (1990).
121. Evan, G. I. & Vousden, K. H. Proliferation, cell cycle and apoptosis in cancer. *Nature* **411**, 342–348 (2001).
122. Meyer, N. & Penn, L. Z. Reflecting on 25 years with MYC. *Nat. Rev. Cancer* **8**, 976–990 (2008).
123. Dong, Y., Tu, R., Liu, H. & Qing, G. Regulation of cancer cell metabolism: oncogenic MYC in the driver's seat. *Signal Transduct. Target. Ther.* **5**, 124 (2020).
124. Baudino, T. A. et al. c-Myc is essential for vasculogenesis and angiogenesis during development and tumor progression. *Genes Dev.* **16**, 2530–2543 (2002).
125. Soucek, L. et al. Modelling Myc inhibition as a cancer therapy. *Nature* **455**, 679–683 (2008).
126. Posternak, V. & Cole, M. D. Strategically targeting MYC in cancer. *F1000Reserach* **5**, F1000 (2016).
127. Holmes, A. G. et al. A MYC inhibitor selectively alters the MYC and MAX cistromes and modulates the epigenomic landscape to regulate target gene expression. *Sci. Adv.* **8**, eabh3635 (2022).
128. Guo, H. et al. Androgen receptor and MYC equilibration centralizes on developmental super-enhancer. *Nat. Commun.* **12**, 7308 (2021).
129. Bouvard, C. et al. Small molecule selectively suppresses MYC transcription in cancer cells. *Proc. Natl Acad. Sci.* **114**, 3497–3502 (2017).
130. Qi, Y., Wei, J. & Zhang, X. Requirement of transcription factor NME2 for the maintenance of the stemness of gastric cancer stem-like cells. *Cell Death Dis.* **12**, 924 (2021).
131. Gong, Y., Yang, G., Wang, Q., Wang, Y. & Zhang, X. NME2 is a master suppressor of apoptosis in gastric cancer cells via transcriptional regulation of miR-100 and other survival factors. *Mol. Cancer Res.* **18**, 287–299 (2020).
132. Wen, S. et al. Nucleoside diphosphate kinase 2 confers acquired 5-fluorouracil resistance in colorectal cancer cells. *Artif. Cells Nanomed. Biotechnol.* **46**, 896–905 (2018).
133. Leinonen, R., Sugawara, H. & Shumway, M., Collaboration, I. N. S. D. The sequence read archive. *Nucleic Acids Res.* **39**, D19–D21 (2010).
134. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinforma* **29**, 15–21 (2013).
135. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
136. Zhao, Y., Wong, L. & Goh, W. W. B. How to do quantile normalization correctly for gene expression data analyses. *Sci. Rep.* **10**, 15534 (2020).
137. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
138. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
139. Epsi, N. J., Panja, S., Pine, S. R. & Mitrofanova, A. pathCHEMO, a generalizable computational framework uncovers molecular pathways of chemoresistance in lung adenocarcinoma. *Commun. Biol.* **2**, 334 (2019).
140. Rahem, S. M., Epsi, N. J., Coffman, F. D. & Mitrofanova, A. Genome-wide analysis of therapeutic response uncovers molecular pathways governing tamoxifen resistance in ER+ breast cancer. *EBio-Medicine* **61**, 103047–103047 (2020).
141. Lefebvre, C. et al. A human B-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. *Mol. Syst. Biol.* **6**, 377–377 (2010).
142. Team, R. C. R. A language and environment for statistical computing. (2013).



143. Krijthe, J. H. & Van der Maaten, L. Rtsne: T-distributed stochastic neighbor embedding using Barnes-Hut implementation. *R package version 0.13*, <https://github.com/jkrijthe/Rtsne> (2015).
  144. Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K. & Toxopeus, A. G. Where is positional uncertainty a problem for species distribution modelling? *Ecography* **37**, 191–203 (2014).
  145. Sanchez, G. & Sanchez, M. G. Package ‘plsdepot’. *Partial Least Squares (PLS) Data Analysis Methods*, v. 0.1 **17** (2012).
  146. Moat, J. & Bachman, S. (See <https://CRAN.R-project.org/package=rCAT>, 2017).
  147. Revelle, W. & Revelle, M. W. Package ‘psych’. *Compr. R. Arch. Netw.* **337**, 338 (2015).
  148. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2019).
  149. Cook, R. D. Detection of influential observation in linear regression. *Technometrics* **19**, 15–18 (1977).
  150. Therneau, T. M. & Lumley, T. Package ‘survival’. *R. Top. Doc.* **128**, 28–33 (2015).
  151. Kassambara, A., Kosinski, M., Biecek, P. & Fabian, S. Package ‘survminer’. *Drawing Survival Curves Using ‘ggplot2’(R package version 03 1)* (2017).
  152. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **27**, 861–874 (2006).
  153. Robin, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinforma.* **12**, 1–8 (2011).
  154. Fisher, R. A. in *Breakthroughs in Statistics* 66–70 (Springer, 1992).
  155. Ishwaran, H. & Kogalur, U. B. Fast unified random forests for survival, regression, and classification (RF-SRC). *R package version 3.2.3* (manual, 2023) <https://cran.r-project.org/package=randomForestSRC>.
  156. Cortez, P. & Cortez, M. P. Package ‘rminer’. *Teach. Rep.* **59** (2016).
  157. Wickham, H. Elegant graphics for data analysis. *Media* **35**, 10.1007 (2009).
  158. Panja, S. et al. Mechanism-centric regulatory network identifies NME2 and MYC programs as markers of Enzalutamide resistance in CRPC. <https://github.com/mitrofanova-lab/TR2PATH> (2023). <https://doi.org/10.5281/zenodo.10368948>.
- package. C.Y.Y. and V.S. contributed equally as co-second authors. M.W.C. and K.W. performed a quality check of the R software package. M.V.T. analyzed treatment information for CRPC patients. A.A.-S. and Sh.J. advised on the parallelization of RNA-seq data mapping. R.V. consulted on the VIF analysis. S.G. advised on comparison to other computational methods. J.S.P. advised on statistical methods utilized in the manuscript. S.Q. performed control in vitro experiments. F.C., Sh.G., and I.K. advised on the clinical utilization of the findings. E.S. supervised in vitro experiments. V.K. designed and performed in vitro analysis. S.A.A. conceived, designed, and analyzed in vitro and in vivo studies. All authors edited and approved the final manuscript.

## Competing interests

S.A.A. is a coinventor on patents covering the methods and assays to identify and characterize MYC inhibitors and derivatives (US11420957B2, “Substituted heterocycles as c-MYC targeting agents”). A.M., S.P., S.A.A. E.M.S., and V.K. are coinventors on patent applications covering the methods and assays to identify Enzalutamide-responsive tumors (PCT Application No. PCT/US2023/065533, “MYC program as a marker of response to Enzalutamide in prostate cancer”). A.M. and S.P. are coinventors on patent applications covering the computational method TR-2-PATH (U.S. Patent Application No. 18/297,470, “Identifying treatment response signatures”). The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-44686-5>.

**Correspondence** and requests for materials should be addressed to Vishal Kothari, Sarki A. Abdulkadir or Antonina Mitrofanova.

**Peer review information** *Nature Communications* thanks Tanya Stoyanova, Ganesh Raj and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

## Acknowledgements

We thank Nusrat J. Epsi, PostDoctoral fellow at Henry M. Jackson Foundation for Military Medicine, for her advice on estimating the activity level of transcriptional regulatory programs. We thank Carrie Esopenko, for sharing materials on partial least square regression analysis. Grant support includes New Jersey Commission on Cancer Research Pre-Doctoral Fellowship DCHS20PPC028 (S.P), *2021CIF-Rutgers-15, by the National Science Foundation (NSF) under Grant # 2127309 to the Computing Research Association for the CI-Fellows Post-Doctoral Fellowship Award* (M.W.C.), NIH NLM R01LM013236 (A.M., I.K.), ACS RSG-21-023-01-TBG (A.M., S.A., E.S., V.K.), NJCCR COCR21RBG00 (A.M., I.K.), NIH P50 CA180995 (S.A.A), NIH R01 CA257258 (S.A.A.), a Prostate Cancer Foundation (PCF) TACTICAL Award (S.A.A.), Polsky Urologic Cancer Institute (S.A.A.), and NIH F30 CA250196 (M.T.).

## Author contributions

S.P., M.I.T., V.K., S.A.A., and A.M. conceived and designed the study. A.M. and S.P. performed the computational and statistical analysis, prepared the figures, and wrote the paper. M.I.T. designed, performed, and analyzed in vitro and in vivo studies and prepared associated figures. C.Y.Y. performed the single-cell analysis. V.S. built the R software