

9-25-2019

## High-fat diet fuels prostate cancer progression by rewiring the metabolome and amplifying the MYC program.

David P. Labbé

*Dana-Farber Cancer Institute; Harvard Medical School; McGill University; Research Institute of the McGill University Health Centre*

Giorgia Zadra

*Dana-Farber Cancer Institute; Brigham and Women's Hospital*

Meng Yang

*Harvard T.H. Chan School of Public Health*

Follow this and additional works at: <https://jdc.jefferson.edu/radoncfp>

Jaime M Reyes

 Part of the [Urology Commons](#), [Oncology Commons](#), and the [Urology Commons](#)

**Let us know how access to this document benefits you**

Charles Y Lin

*Baylor College of Medicine*

### Recommended Citation

Labbé, David P.; Zadra, Giorgia; Yang, Meng; Reyes, Jaime M; Lin, Charles Y; Cacciatore, Stefano; Ebot, Ericka M.; Creech, Amanda L.; Giunchi, Francesca; Fiorentino, Michelangelo; Elfandy, Habiba; Syamala, Sudeepa; Karoly, Edward D.; Alshalalfa, Mohammed; Erho, Nicholas; Ross, Ashley; Schaeffer, Edward M.; Gibb, Ewan A.; Takhar, Mandeep; Den, Robert B.; Lehrer, Jonathan; Karnes, R. Jeffrey; Freedland, Stephen J.; Davicioni, Elai; Spratt, Daniel E.; Ellis, Leigh; Jaffe, Jacob D.; D'Amico, Anthony V.; Kantoff, Philip W.; Bradner, James E.; Mucci, Lorelei A.; Chavarro, Jorge E.; Loda, Massimo; and Brown, Myles, "High-fat diet fuels prostate cancer progression by rewiring the metabolome and amplifying the MYC program." (2019). *Department of Radiation Oncology Faculty Papers*. Paper 123.

<https://jdc.jefferson.edu/radoncfp/123>

This Article is brought to you for free and open access by the Jefferson Digital Commons. The Jefferson Digital Commons is a service of Thomas Jefferson University's [Center for Teaching and Learning \(CTL\)](#). The Commons is a showcase for Jefferson books and journals, peer-reviewed scholarly publications, unique historical collections from the University archives, and teaching tools. The Jefferson Digital Commons allows researchers and interested readers anywhere in the world to learn about and keep up to date with Jefferson scholarship. This article has been accepted for inclusion in Department of Radiation Oncology Faculty Papers by an authorized administrator of the Jefferson Digital Commons. For more information, please contact: [JeffersonDigitalCommons@jefferson.edu](mailto:JeffersonDigitalCommons@jefferson.edu).

---

## Authors


David P. Labbé, Giorgia Zadra, Meng Yang, Jaime M Reyes, Charles Y Lin, Stefano Cacciatore, Ericka M. Ebot, Amanda L. Creech, Francesca Giunchi, Michelangelo Fiorentino, Habiba Elfandy, Sudeepa Syamala, Edward D. Karoly, Mohammed Alshalalfa, Nicholas Erho, Ashley Ross, Edward M. Schaeffer, Ewan A. Gibb, Mandeep Takhar, Robert B. Den, Jonathan Lehrer, R. Jeffrey Karnes, Stephen J. Freedland, Elai Davicioni, Daniel E. Spratt, Leigh Ellis, Jacob D. Jaffe, Anthony V. D'Amico, Philip W. Kantoff, James E. Bradner, Lorelei A. Mucci, Jorge E. Chavarro, Massimo Loda, and Myles Brown

ARTICLE

<https://doi.org/10.1038/s41467-019-12298-z>

OPEN

# High-fat diet fuels prostate cancer progression by rewiring the metabolome and amplifying the MYC program

David P. Labbé  et al.<sup>#</sup>

Systemic metabolic alterations associated with increased consumption of saturated fat and obesity are linked with increased risk of prostate cancer progression and mortality, but the molecular underpinnings of this association are poorly understood. Here, we demonstrate in a murine prostate cancer model, that high-fat diet (HFD) enhances the MYC transcriptional program through metabolic alterations that favour histone H4K20 hypomethylation at the promoter regions of MYC regulated genes, leading to increased cellular proliferation and tumour burden. Saturated fat intake (SFI) is also associated with an enhanced MYC transcriptional signature in prostate cancer patients. The SFI-induced MYC signature independently predicts prostate cancer progression and death. Finally, switching from a high-fat to a low-fat diet, attenuates the MYC transcriptional program in mice. Our findings suggest that in primary prostate cancer, dietary SFI contributes to tumour progression by mimicking MYC over expression, setting the stage for therapeutic approaches involving changes to the diet.

Correspondence and requests for materials should be addressed to M.L. (email: [mloda@med.cornell.edu](mailto:mloda@med.cornell.edu)) or to M.B. (email: [myles\\_brown@dfci.harvard.edu](mailto:myles_brown@dfci.harvard.edu)).

<sup>#</sup>A full list of authors and their affiliations appears at the end of the paper.

Prostate cancer is a leading cause of cancer-related lethality<sup>1</sup>. Systemic metabolic alterations can severely affect the course of the disease. Indeed, epidemiological studies have reported that saturated fat intake and obesity are associated with increased prostate cancer progression and mortality<sup>2–5</sup>. Considering the pandemic of obesity and diet-associated metabolic diseases<sup>6–8</sup> combined with the high frequency of newly diagnosed prostate cancers in developed countries, a better understanding of the mechanistic underpinnings of this link is of significant importance.

Preclinical and clinical studies have shown that systemic metabolic alterations associated with fat-enriched diets and obesity cooperate with tumour initiating genetic alterations to foster disease progression. Modulation of insulin/insulin-like growth factor 1 levels, phosphatidylinositol-3-kinase/mammalian target of rapamycin complex 1 pathway activation and pro-inflammatory stimuli have been implicated<sup>9–14</sup>. However, it is now clear that metabolic rewiring is tightly connected to changes at the epigenetic level as metabolites act as substrates or cofactors for epigenetic remodelling<sup>15,16</sup>.

In prostate cancer, the landscape of epigenetic alterations varies greatly as the disease progresses from a confined tumour to the incurable castration-resistant metastatic stage<sup>15,17</sup>. However, the influence of metabolic alterations triggered by increased fat intake and/or obesity on prostate cancer epigenome rewiring and disease progression is still unexplored.

The oncogene *c-MYC* (*MYC*) is a key driver of human prostate cancer tumorigenesis and progression. *MYC* protein is over-expressed at early stages of the disease<sup>18</sup>, whereas chromosome 8q gain, or focal amplification of 8q24.21, are associated with amplification of the *MYC* oncogene in primary prostate cancer, a feature exacerbated in metastatic disease and associated with poor disease-specific survival<sup>19,20</sup>. In the murine prostate, *MYC* over expression faithfully recapitulates the primary human disease<sup>21</sup>.

A hallmark of *MYC* over expression in tumours is the induction of a global metabolic reprogramming to support cancer cell survival and growth<sup>22–25</sup>. Previous studies have shown that increased dietary fat intake significantly alters the biological behaviours of prostate cancers driven by *MYC*<sup>10,11</sup> suggesting this preclinical model as ideal to investigate the interplay between HFD, oncogene-driven metabolic vulnerabilities, and epigenetic alterations in prostate cancer progression.

Here, we integrate metabolome, epigenome and transcriptome profiling to identify HFD-driven alterations that foster prostate cancer progression in vivo. We demonstrate that increased fat intake amplifies *MYC* hallmarks and further enhances *MYC*'s transcriptional program. Importantly, we identified a fat-induced *MYC* signature with clinical utility in identifying patients at higher risk of a more aggressive, lethal disease. Altogether, our findings suggest that a substantial subset of prostate cancer patients, including some without *MYC* amplification, may benefit from epigenetic therapies targeting *MYC* transcriptional activity or from dietary interventions targeting the metabolic dependencies regulated by *MYC*.

## Results

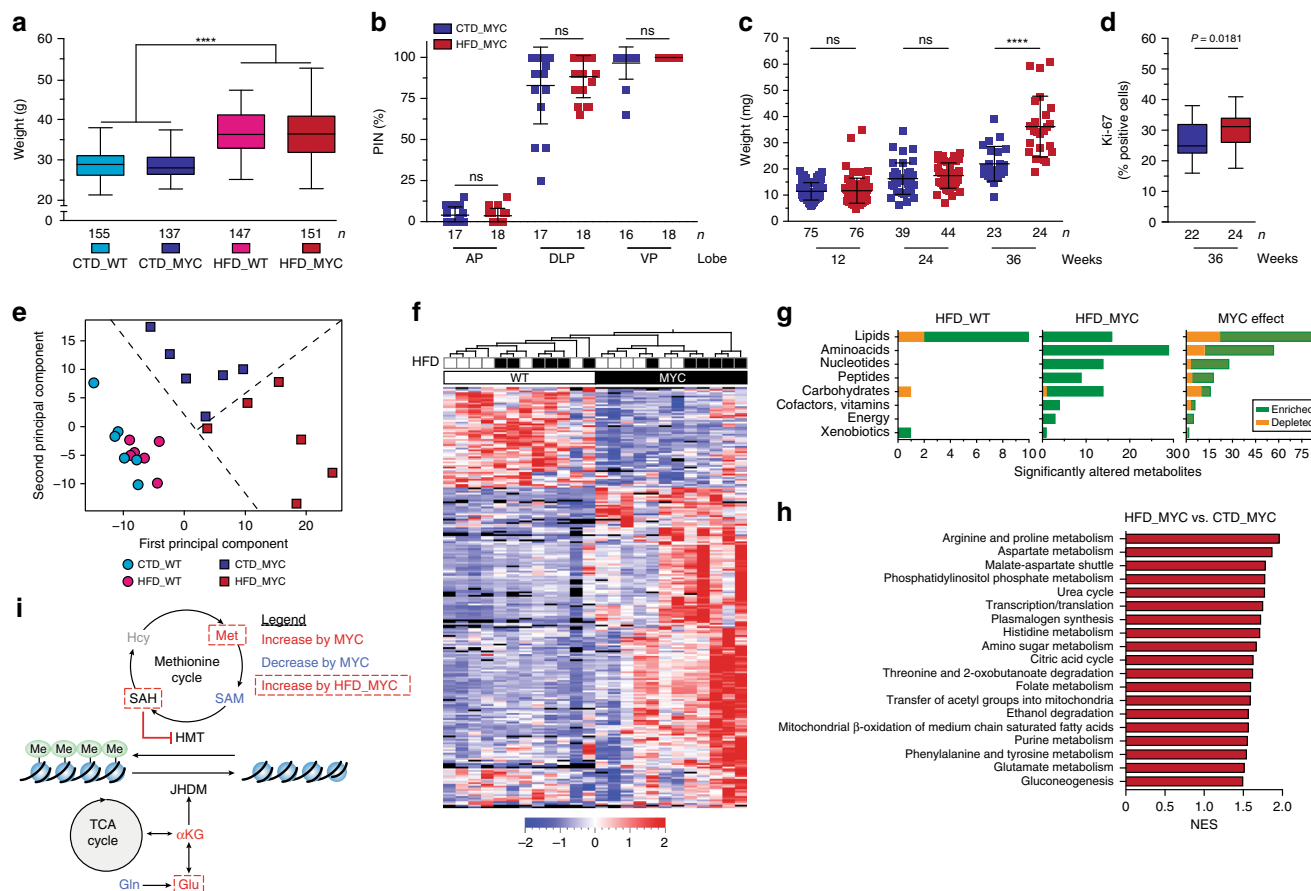
**HFD reprograms cancer metabolome and accelerates progression.** To examine the potential role of high-fat diet (HFD) in promoting metabolic rewiring of prostatic tissues, we compared mice that overexpress a human *c-MYC* transgene (*MYC*) in the prostate epithelium<sup>21</sup> to wild-type littermates (WT) that were fed either a HFD (60% kcal from fat; lard—rich in saturated fat) or a control diet (CTD; 10% kcal from fat; Supplementary Table 1). Irrespective of their genotype, mice that were fed with HFD developed the hallmarks of a diet-induced obesity phenotype,

including increased body weight, liver steatosis, hyperinsulinemia, hyperglycaemia and a decrease in circulating 1,5-anhydroglucitol (a marker of short-term hyperglycaemia) (Fig. 1a and Supplementary Fig. 1a–e). At 12 weeks of age, *MYC* over expression, irrespectively of HFD, resulted in extensive cellular epithelium transformation to prostatic intraepithelial neoplasia (PIN) in the dorsolateral (DLP) and ventral (VP) prostate lobes, the latter with almost complete penetrance. Conversely, the anterior prostate (AP) remained mostly unaffected (Fig. 1b and Supplementary Fig. 1f). No presence of PIN was detected in the prostate lobes of WT animals fed a HFD (Supplementary Data 1). Increased tumour weight (Fig. 1c) and cell proliferation (Ki-67; Fig. 1d) were evident by 36 weeks of age in the HFD-fed mice compared to the CTD group, confirming previous reports that HFD significantly enhances the progression of *MYC*-driven prostate cancer<sup>10,11</sup>.

The lack of a HFD-dependent phenotype at 12 weeks of age, combined with the robust and uniform transition to PIN triggered by *MYC* over expression observed in the VP (Fig. 1b, c and Supplementary Data 1), enabled us to investigate metabolic alterations driven by HFD before the appearance of a more aggressive, HFD-dependent phenotype. Untargeted metabolomics identified 414 metabolites in the prostate. As previously described<sup>26</sup>, we confirmed that *MYC* induces a profound metabolic reprogramming in the VP affecting more than half of the metabolites detected, including metabolites related to glutamine, glucose, lipid, nucleotide metabolism and protein synthesis (Fig. 1e–g and Supplementary Data 2). Importantly, we found that these *MYC*-driven metabolic vulnerabilities were enhanced by HFD. Indeed, HFD resulted in increased levels of metabolites from glycolysis (*i.e.* lactate), glutaminolysis (*i.e.* glutamate), glutamine-metabolism related pathways including substrates, intermediates and final products of the citric acid cycle, nucleotide synthesis, amino acid metabolism (*e.g.* arginine, proline, aspartate and histidine), urea cycle, lipid metabolism and hexosamine biosynthesis (Fig. 1g and Supplementary Data 3); those features were also supported by Metabolite Set Enrichment Analysis (MSEA; Fig. 1h and Supplementary Data 4). Conversely, HFD had little impact on the WT prostatic metabolome, affecting only a total of 12 metabolites, nine of which were glycerophospholipids, and lowering 1,5-anhydroglucitol levels, in line with HFD-driven increase in circulating glucose and reduction of serum 1,5-anhydroglucitol (Fig. 1g, Supplementary Fig. 1d, e, g and Supplementary Data 5).

Notably, *MYC* over expression led to a significant decrease in *s*-adenosylmethionine (SAM), a member of the methionine cycle and the ultimate methyl donor required for methylation reactions (Fig. 1i). The donation of a methyl group by SAM results in its conversion to *s*-adenosylhomocysteine (SAH), which if accumulated, is a potent inhibitor of methyltransferases<sup>27</sup>. *MYC* also enhanced the levels of alpha-ketoglutarate ( $\alpha$ KG), a critical cofactor for histone demethylation mediated by Jumonji Domain-containing Histone Demethylases (JHDM)<sup>28</sup>. Thus, these results suggest that histone methylation processes may be severely hindered during *MYC*-driven prostate cancer progression. Again, this feature was further exacerbated by diet since increased SAH levels (higher SAH/SAM ratio) were observed in the VP of HFD-fed mice (Fig. 1i and Supplementary Fig. 1h–i). Altogether, our data support the notion that HFD amplifies *MYC*-driven metabolic reprogramming.

**HFD enhances transcriptional changes at H4K20me1 dynamic genes.** To validate whether *MYC*/HFD affects histone methylation, we characterised 69 distinct combinations of histone modifications that span H2, H3, and H4 from all four genotype/diet

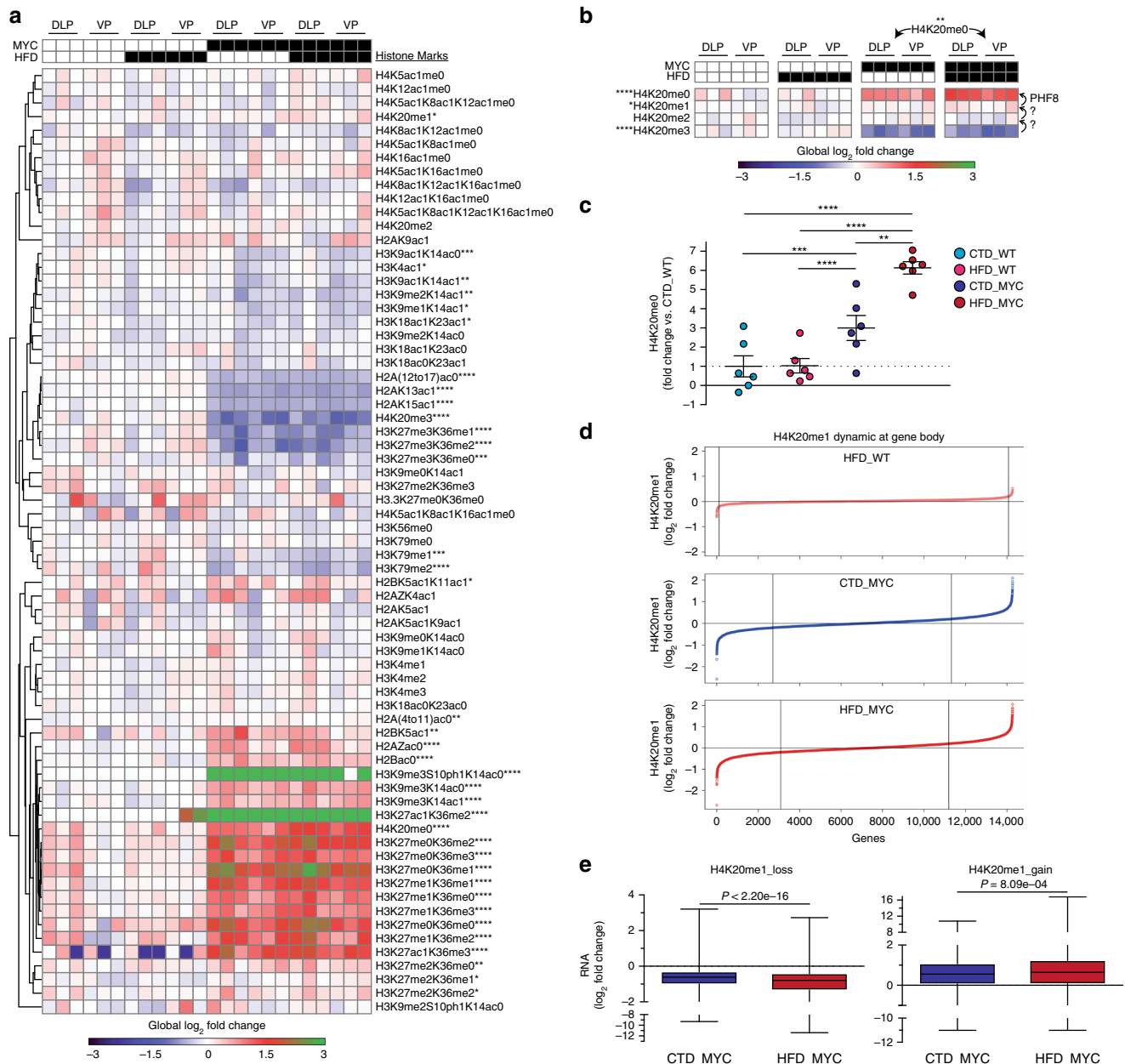


**Fig. 1** High-fat diet reprograms prostate cancer metabolome and accelerates disease progression. **a** Mice fed a high-fat diet (HFD) develop diet-induced obesity at 12 weeks of age ( $n =$  biologically independent animals; two-way ANOVA, median, whiskers  $\pm$  min/max; \*\*\*\* $P < 0.0001$ ). **b–d** HFD does not alter the penetrance of prostatic intraepithelial neoplasia (PIN) at 12 weeks of age (**b**,  $n =$  biologically independent lobes; unpaired  $t$  test, mean  $\pm$  s.d.; CTD: control diet; AP: anterior prostate; DLP: dorsolateral prostate; VP: ventral prostate; ns: non significant), but does lead to a greater tumour burden (**c**,  $n =$  biologically independent lobes; Welch’s  $t$  test, mean  $\pm$  s.d.; \*\*\*\* $P < 0.0001$ ) and to cell proliferation, as assessed by Ki-67 (**d**,  $n =$  biologically independent lobes; unpaired  $t$  test, median, whiskers  $\pm$  min/max) in the VP, at 36 weeks of age. **e** Principal component analysis identifies a distinct metabolic profile in the VP that is triggered by HFD, in a MYC context ( $n = 6$  biologically independent VP/condition, 414 metabolites detected). **f, g** Representation of all metabolites significantly altered by HFD in a WT ( $n = 12$ ) or a MYC ( $n = 89$ ) context, or by MYC overexpression irrespective of the diet ( $n = 214$ ) (**f**, unsupervised hierarchical clustering,  $P < 0.05$  and FDR  $< 0.15$ ); the breakdown of metabolite classes is shown **g**. **h** Metabolite Set Enrichment Analysis (MSEA) revealed metabolic pathways significantly enriched by HFD in MYC-transformed VP ( $P < 0.05$  and FDR  $< 0.15$ ). **i** Metabolic rewiring triggered by MYC and by HFD in a MYC context suggests dampened histone methylation. Hcy: homocysteine (undetected); Met: methionine; TCA: tricarboxylic acid (citric acid cycle); Gln: glutamine; Glu: glutamate. Source data are provided as a Source Data file

combinations in all murine prostatic lobes (DLP, VP, AP; Supplementary Data 6) by using a targeted mass spectrometry approach<sup>29</sup>. Unsupervised clustering of the different combinations of histone modifications revealed a strong MYC-driven signature in both DLP and VP (Fig. 2a). This was absent in the AP (Supplementary Fig. 2a) in line with the marginal PIN penetrance observed in this lobe (Fig. 1b). Among the histone peptides monitored, H3K27/K36 and H4K20 were significantly affected by MYC over expression. As previously described<sup>30</sup>, MYC over expression induced a steep decrease in H3K27me3 (corresponding to the H3K27me3K36meX peptides). In particular, the H3K27me3 mark was hypomethylated in a stepwise process that can be catalysed by multiple JHDM enzymes and culminates with the unmethylated/acetylated H3K27 mark (Supplementary Fig. 2b). A similar pattern was observed for the H4K20 mark, but in this case the effect of MYC was significantly enhanced by HFD, leading to greater levels of the unmethylated mark (Fig. 2b). Importantly, HFD had no effect on the H4K20 mark in the WT tissues (Fig. 2b, c). H4K20me0 can be generated

from H4K20me1, a mark that is associated with transcriptional elongation<sup>31</sup>, by the JHDM enzyme PHF8<sup>32</sup>. Chromatin immunoprecipitation followed by sequencing (ChIP-seq) of H4K20me1 revealed highly dynamic levels of this mark along each gene body upon MYC over expression with respect to the corresponding CTD\_WT reference (Fig. 2d). Interestingly, modulation of the H4K20me1 mark at the gene body dictates levels of gene expression: thus, loss of H4K20me1 is associated with a decrease, while gain of H4K20me1 is associated with an increase in gene expression (Supplementary Fig. 2c). When comparing the gene expression levels for shared H4K20me1 dynamic gene body-associated regions between CTD\_MYC and HFD\_MYC conditions (Supplementary Fig. 2d), we found that the MYC-effect was systematically enhanced by HFD (Fig. 2e). These results suggest that HFD further enhances MYC-driven H4K20 hypomethylation leading to transcriptional changes.

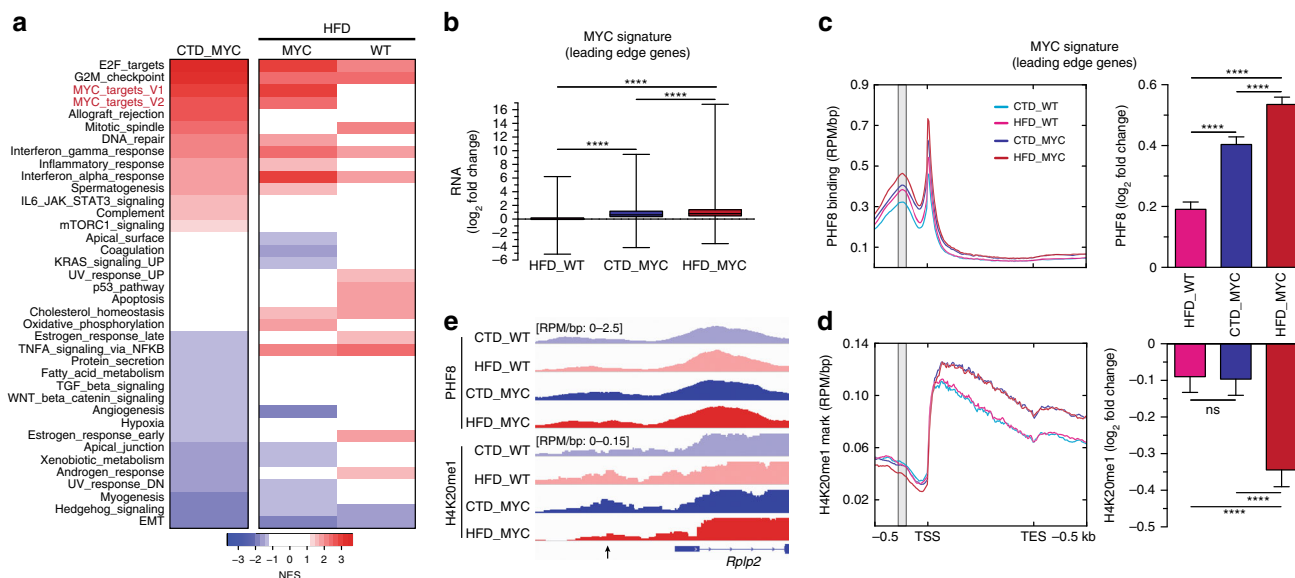
**High-fat diet enhances MYC transcriptional activity.** To determine the cellular program specifically enhanced by HFD



**Fig. 2** High-fat diet enhances MYC-driven transcriptional changes at H4K20me1 dynamic genes. **a** Global chromatin profiling identifies distinct chromatin-signature profiles in MYC-overexpressing DLP and VP lobes (histone marks levels relative to the DLP, VP and AP CTD\_WT median values; MYC vs. WT comparisons, unpaired *t* test; \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001, \*\*\*\**P* < 0.0001; Supplementary Data 6). **b** HFD enhances H4K20 hypomethylation triggered by MYC (MYC vs. WT and HFD\_MYC vs. CTD\_MYC comparisons, unpaired *t* test; \**P* < 0.05, \*\**P* < 0.01, \*\*\*\**P* < 0.0001). **c** H4K20me0 levels relative to the CTD\_WT condition (fold change; unpaired *t* test, mean ± SEM; \*\**P* < 0.01, \*\*\**P* < 0.001, \*\*\*\**P* < 0.0001). **d** H4K20me1 dynamic regions across all murine gene bodies relative to the CTD\_WT condition. **e** HFD enhances the effect of MYC on gene expression at H4K20me1 dynamic regions that were lost (top, *n* = 2508 genes) or gained (bottom, *n* = 3208 genes; paired *t* test, median, whiskers ± min/max)

within a MYC context, we performed Gene Sets Enrichment Analyses (GSEA) using the Hallmark gene sets (Supplementary Data 7)<sup>33</sup>. As expected, MYC over expression led to the enrichment of gene sets related to cell proliferation (E2F\_targets, G2M\_checkpoint), as well as MYC-transcriptional activity per se (Fig. 3a, left). Interestingly, HFD further enriched both gene sets related to MYC transcriptional activity (V1/V2), but only in MYC-transformed prostates (Fig. 3a, right). This feature was not linked to an increased expression of the MYC transgene (Supplementary Fig. 3a). Because the MYC transcriptional program is highly context-specific<sup>34</sup>, we generated a murine prostatic MYC

signature by including the leading edge genes (*n* = 610) of MYC-related gene sets that were significantly enriched by MYC and/or HFD feeding (Supplementary Fig. 3b and Supplementary Data 8–9). As expected, the expression levels of MYC signature genes were elevated following MYC over expression, and further increased by HFD (Fig. 3b). ChIP-seq of PHF8, the JHDM that mediates H4K20me1 demethylation<sup>32</sup> and a known MYC transcriptional coactivator and regulator of proliferation<sup>35,36</sup>, revealed that MYC over expression increases the recruitment of PHF8 to the promoter regions of MYC signature genes. Again, we observed that this effect was enhanced by HFD (Fig. 3c).



**Fig. 3** High-fat diet enhances MYC transcriptional activity. **a** Gene Set Enrichment Analysis (GSEA, Hallmark,  $P < 0.05$  and  $FDR < 0.1$ ) revealed enhanced expression of MYC target genes triggered by HFD in the MYC context (right column-left side: HFD\_MYC vs. CTD\_MYC; right column-right side: HFD\_WT vs. CTD\_WT comparisons; left column: CTD\_MYC vs. CTD\_WT comparison). **b** HFD boosts the expression of the murine prostatic-derived MYC signature ( $n = 610$  genes; fold-change relative to the CTD\_WT condition, paired  $t$  test, median, whiskers  $\pm$  min/max; \*\*\*\* $P < 0.0001$ ). **c** Increased PHF8 recruitment at the promoter (grey;  $-270$  to  $-210$  bp) of MYC signature genes, is triggered by MYC overexpression and boosted by HFD (fold-change relative to the CTD\_WT condition, paired  $t$  test, mean  $\pm$  SEM; \*\*\*\* $P < 0.0001$ ). **d** Depletion in H4K20me1 mark at the promoter of MYC signature genes (grey) requires increased PHF8 recruitment together with HFD (paired  $t$  test, mean  $\pm$  SEM; \*\*\*\* $P < 0.0001$ ). **e** Representative PHF8 and H4K20me1 tracks at *Rplp2* promoter (arrow), a MYC signature leading edge gene (RPM/bp: reads per million/base pair)

However, only when MYC over expression was combined with HFD, was a significant decrease in H4K20me1 observed at PHF8 recruitment sites (Fig. 3d, e). Taken together, these results suggest that the observed HFD-induced enhancement of MYC transcriptional program is, at least in part, mediated via an increased recruitment and activity of PHF8 toward the H4K20me1 mark at MYC signature genes. This program then culminates in augmented cell proliferation and tumour burden (Fig. 1c, d).

**A SFI MYC signature is associated with lethal disease.** Since our results in the preclinical model represent the combined effects of both increased dietary animal fat intake (AFI) and the diet-induced obesity phenotype, we next investigated whether dietary AFI, independently of obesity, could recapitulate the HFD-driven phenotype in humans. We used data on AFI, as documented in the Health Professionals Follow-up Study (HPFS) and Physicians' Health Study (PHS) cohorts, to stratify the 319 prostate cancer patients for whom tumour (genetic background uncharacterised) and adjacent-normal gene expression profiles were available (Table 1). Using GSEA analysis, we identified the MYC\_targets\_V1 among the three gene sets that were significantly enriched by high AFI, while no gene set was enriched in the adjacent-normal prostatic tissues (Fig. 4a and Supplementary Data 10). When the leading edge genes within the AFI-induced MYC\_targets\_V1 signature (Fig. 4b, Table 1 and Supplementary Data 11) were used to create a metagene score, we found that prostate cancer patients with greater AFI-dependent MYC transcriptional activation in their tumour tissues were more likely to die of their disease ( $n$  lethal = 31, Odds Ratio (OR) = 3.44, 95% CI = 1.69–7.38). This holds true after adjusting for gleason grade and body mass index (BMI; Table 2). Conversely, when we used the MYC signature genes not enriched by AFI (non-leading edges genes) to create a similar metagene score, this score was unable to predict lethal disease after adjusting for gleason grade and BMI

(Table 2). We next investigated which type of fatty acid contributes to the enrichment of the MYC transcriptional program. We identified the MYC\_targets\_V1 as the top gene set that was enriched by high saturated fat intake (SFI; Fig. 4a, and Supplementary Data 12), while neither monounsaturated nor polyunsaturated fat intake was associated with a positive enrichment of any given gene set (Fig. 4a and Supplementary Data 13–14). Importantly, the SFI-induced MYC\_targets\_V1 signature was more robustly associated with prostate cancer lethality ( $n$  lethal = 34, Odds Ratio (OR) = 4.02, 95% CI = 1.98–8.63; Tables 1 and 2), a feature that was not recapitulated when using a randomly picked MYC\_targets\_V1 signature (Supplementary Table 2). Furthermore, the metagene score was more strongly related to lethal disease in men with a high SFI than in men with a low SFI ( $P$  for interaction = 0.03; Fig. 4c). These results indicate that the MYC-transcriptional program specifically induced by SFI drives prostate cancer lethality.

We confirmed the prognostic value of the SFI-induced MYC signature in four independent clinical cohorts by analysing gene expression in the tumours from 631 prostate cancer patients. Strikingly, even in these cohorts lacking patient dietary information, the high expression of the SFI-induced MYC signature identified patients that were more likely to progress to a metastatic disease in the Thomas Jefferson University (TJU), Johns Hopkins Medical Institutions-I (JHMI-I), Mayo Clinic and Cedar-Sinai cohorts ( $P = 1.33e-04$ ), a feature that was much less pronounced when using the non-SFI-associated MYC signature ( $P = 1.26e-02$ ; Fig. 4d). Importantly, in patients from the TJU/JHMI-I/II cohorts, the SFI-induced MYC signature was not associated with BMI (Supplementary Fig. 4). Additional univariate and multivariate analyses confirmed the prognostic power of the SFI-induced MYC signature in predicting prostate cancer progression to a metastatic disease, even after adjusting for gleason grade or the Cell Cycle Progression score consisting of 31 cell cycle genes (Supplementary Tables 3–5)<sup>37</sup>. Altogether, these

**Table 1 Characteristics of 319 men diagnosed with prostate cancer from 1982 to 2005 in the Health Professionals Follow-up Study and the Physicians' Health Study according to fat intake MYC metagene scores**

Characteristic	All men (n = 319)	Animal fat MYC metagene score <sup>a</sup>			Saturated fat MYC metagene score <sup>a</sup>		
		Tertile 1 (low) (n = 107)	Tertile 2 (n = 106)	Tertile 3 (high) (n = 106)	Tertile 1 (low) (n = 106)	Tertile 2 (n = 107)	Tertile 3 (high) (n = 106)
Age at diagnosis, years, mean (SD)	65.0 (6.3)	65.6 (6.2)	64.9 (6.6)	64.5 (6.1)	65.6 (6.2)	64.7 (6.3)	64.7 (6.3)
Year of diagnosis, n (%)							
Before 1990 (pre-PSA era)	28 (8.8)	9 (8.4)	10 (9.4)	9 (8.5)	8 (7.4)	12 (11.3)	8 (7.5)
1990-1993 (peri-PSA era)	83 (26.0)	32 (29.9)	27 (25.5)	24 (22.6)	32 (29.9)	27 (25.5)	24 (22.6)
After 1993 (PSA era)	208 (65.2)	66 (61.7)	69 (65.1)	73 (68.9)	67 (62.6)	67 (63.2)	74 (69.8)
BMI at diagnosis, kg/m <sup>2</sup> , mean (SD)	25.2 (2.9)	24.7 (2.6)	25.4 (2.9)	25.5 (3.1)	24.7 (2.6)	25.4 (2.8)	25.6 (3.2)
PSA at diagnosis, ng/ml, median [25th - 75th percentile] <sup>b</sup>	7.4 [5.3-11.6]	7.0 [5.0, 11.9]	7.6 [5.5, 12.9]	8.0 [5.6, 11.1]	6.9 [5.0-12.0]	7.6 [5.4-13.0]	7.9 [5.6-11.1]
Pathologic TNM stage, n (%) <sup>c</sup>							
T2 N0 M0	192 (61.9)	73 (69.5)	60 (58.3)	59 (57.8)	72 (68.6)	60 (58.3)	60 (58.8)
T3 N0 M0	107 (34.5)	29 (27.6)	40 (38.8)	38 (37.3)	30 (28.6)	40 (38.8)	37 (36.3)
T4/N1/M1	11 (3.5)	3 (2.9)	3 (2.9)	5 (4.9)	3 (2.9)	3 (2.9)	5 (4.9)
Clinical TNM stage, n (%) <sup>d</sup>							
T1/T2 N0 M0	297 (93.4)	103 (96.3)	98 (92.5)	96 (91.4)	102 (95.3)	99 (93.4)	96 (91.4)
T3 N0 M0	21 (6.6)	4 (3.7)	8 (7.5)	9 (8.6)	5 (4.7)	7 (6.6)	9 (8.6)
Gleason grade, n (%)							
<7	51 (16.0)	24 (22.4)	16 (15.1)	11 (10.4)	22 (20.6)	17 (16.0)	12 (11.3)
3+4	124 (38.9)	49 (45.8)	39 (36.8)	36 (34.0)	49 (45.8)	40 (37.7)	35 (33.0)
4+3	81 (25.4)	19 (17.8)	32 (30.2)	30 (28.3)	20 (18.7)	32 (30.2)	29 (27.4)
>7	63 (19.7)	15 (14.0)	19 (17.9)	29 (27.4)	16 (15.0)	17 (16.0)	30 (28.3)
Tissue type, n (%)							
RP	311 (97.4)	105 (98.1)	103 (97.2)	103 (97.2)	105 (98.1)	103 (97.2)	103 (97.2)
TURP	8 (2.5)	2 (1.9)	3 (2.8)	3 (2.8)	2 (1.9)	3 (2.8)	3 (2.8)
Cohort, n (%)							
HPFS	213 (66.8)	61 (57.0)	73 (68.9)	79 (74.5)	65 (60.7)	68 (64.2)	80 (75.5)
PHS	106 (33.2)	46 (43.0)	33 (31.1)	27 (25.5)	42 (39.3)	38 (35.8)	26 (24.5)

SD standard deviation, PSA prostate-specific antigen, BMI body mass index, TNM tumour, lymph node, metastasis, RP radical prostatectomy, TURP transurethral resection of the prostate, HPFS Health Professionals Follow-up Study, PHS Physicians' Health Study

<sup>a</sup>The genes identified in the enrichment analysis of MYC\_targets\_V1 pathway in tumour tissues were used to create a metagene score. A score was computed for each sample by averaging the normalised (mean-centered and variance scaled) expression values of all member genes. The score was divided into tertiles

<sup>b</sup>29 men missing PSA at diagnosis

<sup>c</sup>Nine men missing pathologic TNM stage

<sup>d</sup>One man missing clinical TNM stage

results demonstrate that high SFI, independent of obesity or features of it, fosters a MYC-driven cellular program, promoting the progression to a metastatic and lethal disease.

Finally, we investigated whether a dietary intervention could reverse the HFD-induced MYC transcriptional program. While the HFD robustly enhanced the MYC transcriptional program induced by MYC over expression in the murine prostate, switching to a CTD at 10 weeks of age was sufficient to dampen the MYC\_targets\_V1 signature observed in 12-week-old mice (Fig. 4e). This suggests that a dietary intervention aimed at lowering AFI and potentially more importantly SFI in patients might be able to directly impact the MYC transcriptional program, thereby reducing or delaying the progression to a lethal, metastatic disease.

## Discussion

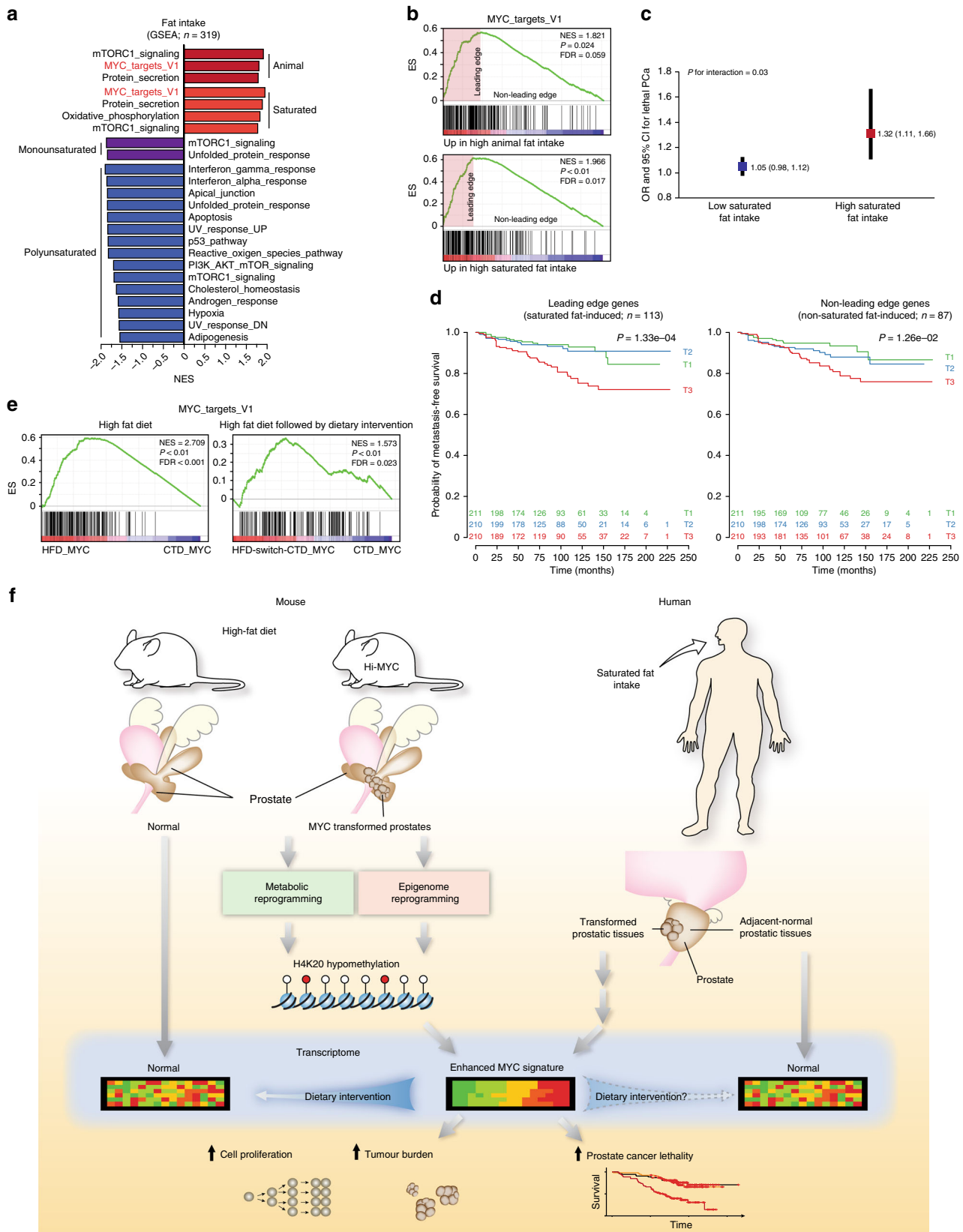
In this study, we report the effect of HFD-mediated systemic alterations on prostate cancer progression. Our data demonstrate that HFD synergises with oncogenic transformation of the prostate to promote a MYC-driven program and disease progression. In the normal prostate, HFD impacts metabolites that are primarily restricted to membrane lipid remodelling, has little influence on histone modifications, and results in a distinct transcriptional program compared to that induced by HFD in the transformed prostate. Conversely, HFD profoundly alters an early

stage of MYC-induced prostate transformation characterised by PIN, resulting in the enhancement of MYC-driven metabolic, epigenetic, and transcriptional programs (Fig. 4f). These data suggest that a premalignant condition such as PIN, which often precedes the onset of invasive adenocarcinoma in humans<sup>38</sup>, is required for HFD to exert its MYC-amplifying effects in the prostate.

A substantial body of literature supports the notion that cellular metabolism has a profound influence on epigenetic modifications, which rely on metabolites as substrates or cofactors<sup>39-42</sup>. Here, we provide the evidence that HFD acts as a master effector of prostate cancer metabolism, creating an environment that favours histone hypomethylation and results in an enhanced MYC-driven transcriptional program. Notably, we observed a decrease in the H4K20me1 mark at the promoter region of MYC signature genes, a feature that was associated with both an increased recruitment and activity of PHF8, a JHDM and the only enzyme known to demethylate the H4K20me1 mark<sup>32</sup>. Along this line, PHF8 has been documented in cell culture systems as a MYC transcriptional coactivator<sup>35</sup> and a regulator of prostate cancer cell proliferation, migration and invasion<sup>36,43</sup>, supporting the idea that an incremental gain in PHF8 activity by HFD at H4K20me1 might provide increased tumour fitness over the course of prostate cancer development.

The unifying aspect that translates from mice to humans is the fat-induced MYC signature, a feature that persists despite the





**Fig. 4** A saturated fat-induced MYC signature is associated with lethal prostate cancer. **a, b** GSEA analysis (Hallmark) revealed that high animal fat and high saturated fat intake enriches for the MYC\_targets\_V1 gene set (**a**,  $P < 0.05$  and FDR < 0.1), as represented by the enrichment plot (**b**) in the HSPH/PHS cohorts. **c** The lethality for every 0.1 unit increase of MYC score was significantly elevated among patients with high saturated fat intake compared with those with low saturated fat intake. **d** High expression of the saturated fat-induced MYC signature is significantly associated with reduced metastatic-free survival (T3) in four independent cohorts (TJU/JHMI-/Mayo Clinic/Cedar-Sinai cohorts,  $n = 631$ ). **e** Short-term dietary intervention (HFD switch to CTD) dampens the HFD-induced MYC transcriptional activity in MYC-driven murine prostate cancer. **f** Graphical summary

**Table 2 Fat-induced and non-fat-induced MYC signature score in relation to risk of prostate cancer death among men diagnosed with non-metastatic prostate cancer**

MYC score	n	Leading edge genes (fat-induced) <sup>a</sup>			n	Non-leading edge genes (non-fat-induced) <sup>b</sup>		
		OR (95% CI) <sup>c</sup>	OR (95% CI) <sup>d</sup>	OR (95% CI) <sup>e</sup>		OR (95% CI) <sup>c</sup>	OR (95% CI) <sup>d</sup>	OR (95% CI) <sup>e</sup>
Animal fat								
Tertile 1 (low)	13	1.00	1.00	1.00	17	1.00	1.00	1.00
Tertile 2	18	1.58 (0.73, 3.53)	1.31 (0.57, 3.08)	1.27 (0.55, 2.99)	19	1.17 (0.57, 2.44)	1.03 (0.47, 2.30)	0.96 (0.43, 2.16)
Tertile 3 (high)	31	3.44 (1.69, 7.38)	2.50 (1.14, 5.70)	2.37 (1.07, 5.43)	26	1.79 (0.90, 3.64)	1.07 (0.81, 3.70)	1.66 (0.78, 3.61)
P, linear trend <sup>f</sup>		0.001	0.019	0.03		0.09	0.15	0.17
Saturated fat								
Tertile 1 (low)	13	1.00	1.00	1.00	16	1.00	1.00	1.00
Tertile 2	15	1.23 (0.55, 2.80)	1.07 (0.45, 2.59)	1.05 (0.44, 2.54)	18	1.24 (0.59, 2.64)	1.17 (0.52, 2.65)	1.09 (0.48, 2.48)
Tertile 3 (high)	34	4.02 (1.98, 8.63)	3.21 (1.47, 7.35)	3.04 (1.38, 7.01)	28	2.34 (1.18, 4.82)	1.93 (0.90, 4.23)	1.86 (0.87, 4.08)
P, linear trend <sup>f</sup>		0.0001	0.002	0.004		0.015	0.085	0.107

n lethal events, OR odds ratio; CI confidence interval

<sup>a</sup>Animal fat: n = 122; Saturated fat: n = 113

<sup>b</sup>Animal fat: n = 78; Saturated fat: n = 87

<sup>c</sup>Logistic regression model adjusted for age and year at diagnosis (continuous)

<sup>d</sup>Logistic regression model adjusted for age and year at diagnosis (continuous), and gleason grade (continuous: <7, 3 + 4, 4 + 3 >7)

<sup>e</sup>Logistic regression model adjusted for age and year at diagnosis (continuous), gleason grade (continuous: <7, 3 + 4, 4 + 3, >7), and BMI at diagnosis (continuous)

<sup>f</sup>Estimated by modeling tertiles of MYC score as continuous variable (tertile 1 = 0, tertile 2 = 1, tertile 3 = 2)

genetic heterogeneity of human prostate cancers<sup>19,44</sup>. Indeed, dietary fat intake does not only amplify the MYC transcriptional program in MYC-driven prostate cancers, but can enrich for it, even in cancers lacking MYC over expression. This suggests that the enhancement of MYC-driven metabolic and epigenetic reprogramming may be a general mechanism that underlies the influence of dietary fat intake on prostate cancer progression although this hypothesis remains to be tested across prostate cancer molecular subtypes.

Finally, we show that the saturated fat-induced MYC transcriptional signature is not only a tumour biomarker for the patients' diet, but it is prognostic for progression to metastatic, lethal prostate cancer. Indeed, the SFI-induced MYC signature is able to predict prostate cancer lethality, independently of the degree of tumour differentiation and patient obesity status and the robustness of this finding was validated in four independent cohorts. Importantly, the non-fat-induced MYC signature was unable to predict lethal disease in the HPFS/PHS cohorts and was only marginally significant as prognostic marker in the validation cohorts. This highlights the fact that saturated fat intake not only enriches the expression of MYC-regulated genes but does so especially for the most predictive subset of genes, possibly reflecting the convergence of oncogenic signalling pathways on dysregulated lipid metabolism, a key feature for prostate cancer development and progression to a metastatic disease<sup>45</sup>.

Several in vitro studies showed that MYC drives the de novo synthesis of fatty acids and increases the expression of the key lipogenic enzymes such as the ATP Citrate Lyase (ACLY), Acetyl-CoA carboxylase (ACC) and fatty acid synthase (FASN)<sup>46–48</sup>. Moreover, previous metabolic profiling performed by our group on MYC engineered prostate cells, in vivo models, and human prostate cancers showed that MYC overexpression specifically enhances lipid metabolism underlying lipid deregulation as key metabolic feature of MYC oncogenic activity in prostate cancer<sup>26</sup>. Recently HFD has been shown to induce an intra-tumoural lipogenic signature driving metastasis formation in *Pten* deficient mouse model<sup>49</sup>. We are currently performing lipidomics analysis to investigate the interplay between diet-derived fats and MYC-driven de novo lipogenesis in shaping the tumour lipidome and promoting a more aggressive phenotype.

Collectively, our results suggest that extrinsic risk factors—such as saturated fat intake—contribute to prostate cancer lethality by enriching for a MYC-transcriptional program; and either

synergise with MYC over expression, which is observed in 37% of metastatic prostate cancers<sup>50</sup>, or phenocopy MYC amplification (Fig. 4f). While neither MYC protein nor MYC mRNA over-expression measured in primary tumours from patients subjected to radical prostatectomy are strong prognostic markers<sup>51</sup>, our findings suggest that a SFI-dependent MYC signature can be used in the clinical setting to identify patients with a worse prognosis. Finally, our study raises the possibility that a nutritional program such as that involving the reduction of animal fat and specifically saturated fat consumption in men with early-stage cancer may dampen the MYC transcriptional program and diminish or delay the risk of disease progression.

## Methods

**Animal husbandry.** FVB Hi-MYC mice (strain number 01XK8), expressing the human c-MYC transgene in prostatic epithelium, were obtained from the National Cancer Institute Mouse Repository at Frederick National Laboratory for Cancer Research<sup>21</sup>. Upon weaning (3 weeks), male mice heterozygous for the transgene (MYC), together with their wild type littermates (WT), were fed a purified control diet (CTD; Harlan Laboratories, TD.130838) consisting of 10% fat, or a high-fat diet (HFD; Harlan Laboratories, TD. 06414) consisting of 60% fat (Supplementary Table 1) until 12, 24 or 36 weeks of age; ingredients were adjusted on a kcal basis (Supplementary Table 6). For dietary intervention experiments, mice assigned an HFD were switched to a CTD at 10 weeks of age for the following 2 weeks until the experimental endpoint. Litters were randomly assigned to each diet. Group allocation was performed in a non-blinded fashion. Food was changed on a weekly basis, and mice were weighed every three weeks, starting at weaning. Animals were kept on a 12-h light/12-h dark cycle, and allowed free access to food and water at the Dana-Farber Cancer Institute (DFCI) Animal Resources Facility. The animal protocol was reviewed and approved by the DFCI Institutional Care and Use Committee (IACUC), and was in accordance with the Animal Welfare Act. Mice sample size estimate for analyses was based on published literature.

**Tissue collection.** At defined time points, mice were weighed and euthanized by CO<sub>2</sub>, followed by cervical dislocation; blood was collected by cardiac puncture, and serum was collected using serum-separating tubes (#41.1378.005, Sarstedt), aliquoted, and stored at –80 °C. Urogenital apparatus and liver tissues were fixed in 10% buffered formalin and processed for paraffin embedding. Alternatively, mouse prostate lobes (anterior prostate, AP; dorsolateral prostate, DLP; ventral prostate, VP) were immediately dissected, weighed and flash-frozen in liquid nitrogen. Serum and tissues were consistently collected during the same periods to minimise inter-samples and circadian rhythm variability.

**Histopathologic and immunohistochemical analyses.** Formalin-fixed, paraffin-embedded mouse urogenital apparatus and liver tissues were sectioned (5 µm) and stained with hematoxylin and eosin (H&E). Histopathological slides were analysed by expert murine uropathologist, who were blind to the experimental conditions. Hepatic steatosis was also assessed for liver tissues (M.L.). The presence and extent

of PIN in 12-week-old mice (AP, DLP, VP) was estimated for each mouse, by evaluating the percentage of the gland affected for each prostate lobe and reported in Supplementary Data 1 (M.L.). For Ki-67 staining, slides were baked for 60 min in an oven set to 60 °C. They were then loaded into the Bond III staining platform with appropriate labels. Slides were antigen retrieved in Bond Epitope Retrieval 2 for 20 min, and incubated with rabbit monoclonal anti-Ki-67 antibody (#VP-RM04 (clone SP6), Vectors Laboratories) at dilution 1:250 for 30 min, room temperature. Primary antibody was detected using Bond Polymer Refine Detection kit. Slides were developed in 3,3'-diaminobenzidine (DAB), dehydrated, and coverslipped. The percentage of Ki-67 positive cells was evaluated by counting the number of cell that expresses nuclear Ki-67 as a function of the total number of cells per high power field. Whenever possible, up to 10 high power fields for each VP lobe were counted, averaged, and counts were reported as each sample's score (F.G. and M. F.). Sample size for histological evaluation was estimated based on previous literature data, using the same model<sup>10</sup>. For Ki-67 analysis, we performed sample size calculation using the software G\*power version 3.1, extrapolating the effect size ( $d = \text{around } 0.87$ ) from the data of Kobayashi et al.<sup>11</sup> in MYC mice fed with HFD. Based on this assumption, we calculated that at least 22 mice/group should be used to detect a significant difference in Ki-67 positivity using a two-sided *t*-test for change in mean between two independent groups, with an alpha-error of 0.05 and a priori power of 0.8.

**Insulin ELISA.** Serum insulin levels were measured using an insulin-1 ELISA kit from Sigma-Aldrich (#RAB0817). Briefly, samples were diluted 1:3 or 1:5 in diluent buffer C (provided in the kit) and the assay was performed according to the manufacturer's instructions. Each sample was measured twice (technical duplicate). Outliers (identified using the ROUT method,  $Q = 0.1\%$ ), and samples in which insulin levels were under the detection limit of the assay, were removed from the analysis. Statistical analysis and graphical representation were performed with use of GraphPad Prism version 7.0.

**Metabolic profiling.** For Metabolic profiling of serum and prostatic tissues (VP), we used the platform from Metabolon Inc. (Durham, NC, USA). Mice sample size to ensure adequate power for metabolomics analysis was based on previous literature data using a similar model<sup>26</sup>. Information regarding sample preparation, quality assurance (QA) and control (QC), and metabolite quantification was provided by the company as follows:

**Sample preparation:** Biological samples were stored at  $-80\text{ }^{\circ}\text{C}$  and then thawed on ice just prior to extraction. Tissue samples were weighed at Metabolon on a 4-position analytical scale (1/10th mg) and then soaked overnight in 80% methanol/20% deionized water with recovery standards at a 60  $\mu\text{L}$ : 1 mg ratio. The methanol contained four recovery standards (DL-2-fluorophenylglycine, tridecanoic acid, d6-cholesterol and 4-chlorophenylalanine) to allow confirmation of extraction efficiency. For serum, 100  $\mu\text{L}$  sample volume was extracted with 500  $\mu\text{L}$  of methanol containing recovery standards. All extracts were divided into four fractions: one for Ultra-performance liquid chromatography tandem mass-spectrometry (UPLC-MS/MS) with positive ion mode electrospray ionisation (IMEI); one for (UPLC-MS/MS) with negative IMEI; one for liquid chromatography (LC) polar platform; the final fraction was reserved as a backup. Aliquots were dried and then the first aliquot was reconstituted in 80  $\mu\text{L}$  of 6.5 mM ammonium bicarbonate in water (pH 8) for the negative ion analysis, the second aliquot was reconstituted using 80  $\mu\text{L}$  0.1% formic acid in water (pH  $\sim 3.5$ ) for the positive ion method, while the third aliquot was reconstituted in 80  $\mu\text{L}$  of hydrophilic interaction liquid chromatography (HILIC) solvent (15%  $\text{H}_2\text{O}$ : 5% MeOH: 80% ACN) with 10 mM ammonium formate (pH $\sim 10$ ) for the HILIC method.

**QA/QC:** Several types of controls were analysed together with the experimental samples: (1) a pooled matrix sample specific for each sample type (i.e. prostate and serum) was generated by combining 20  $\mu\text{L}$  of each experimental sample and injecting the pooled sample six times for each data set to serve as a technical replicate to assess process variability; (2) five water aliquots were extracted and analysed to serve as process blanks for artifact determination; (3) a cocktail of internal standards, carefully chosen to not interfere with the measurement of endogenous compounds, was spiked into every analysed sample to monitor instrument performance and serve as retention markers for chromatographic alignment. The list of internal standards is provided in Supplementary Table 7. Instrument variability was evaluated during the entire procedure. Experimental samples were randomised across the platform run.

**UPLC Method:** Separations were performed using a Waters Acquity UPLC (Waters, Milford, MA). Reverse-phase (RP) positive ion method analysis used mobile phase consisting of 0.1% formic acid in water (A) and 0.1% formic acid in methanol (B). Reverse-phase negative ion analysis used mobile phase consisting of 6.5 mM ammonium bicarbonate in water, pH 8 (A) and 6.5 mM ammonium bicarbonate in 95% methanol/5% water (B). The sample injection volume was 5  $\mu\text{L}$  and a 2x needle loop overflow was used. Separations utilised separate acid and base-dedicated 2.1 mm  $\times$  100 mm Waters BEH C18 1.7  $\mu\text{m}$  columns held at 40 °C. HILIC used mobile phase consisting of 10 mM ammonium formate in 15% water, 5% methanol, 80% acetonitrile (effective pH 10.16 with  $\text{NH}_4\text{OH}$ ) (A) and 10 mM ammonium formate in 50% water, 50% acetonitrile (effective pH 10.60 with  $\text{NH}_4\text{OH}$ ) (B). The sample injection volume was identical to RP method. The stationary phase consisted of a 2.1 mm  $\times$  150 mm Waters BEH Amide 1.7  $\mu\text{m}$

column held at 40 °C. The gradient profiles for RP and HILIC methods can be found in Supplementary Table 8.

**High Resolution Accurate Mass (HRAM) method:** A ThermoFisher Scientific (Waltham, MA) Q-Exactive was the HRAM instrument used<sup>52</sup>. Detailed source and MS settings can be found in Supplementary Table 9 (conditions are also described in supplementary information from Evans et al.)<sup>53</sup>. The scan range was 80–1000  $m/z$  with a scan speed of  $\sim 9$  scans per second (alternating between MS and MS/MS scans), and the resolution was set to 35,000 (measured at 200  $m/z$ ). Mass calibration was performed as needed to maintain  $<5$  ppm mass error for all standards monitored.

**Biological sample analysis:** Metabolon has developed a chemocentric approach that was used in peak detection and integration, and is described in detail elsewhere<sup>54–56</sup>. This in-house peak detection and integration software was used, the data output of which was a list of  $m/z$  ratios, retention indices (RI) and area under the curve (AUC) values. User specified criteria for peak detection included thresholds for signal to noise ratio, area and width. Relative standard deviations (RSDs) of peak area were determined for each internal and recovery standard to confirm extraction efficiency, instrument performance, column integrity, chromatography and mass calibration. The biological data sets, including QC samples, were chromatographically aligned based on a retention index that utilised internal standards assigned a fixed RI value. The RI of the experimental peak was determined by assuming a linear fit between flanking RI markers whose RI values are set. Peaks were matched against an in-house library of authentic standards and routinely detected unknown compounds specific to the respective method. The library consisted of 3200 endogenous and exogenous metabolites for which super and subpathway designations were provided. Identifications were based on retention index values, experimental precursor mass match to the library authentic standard within 10 ppm, and quality of MS/MS match. MS/MS forward and reverse match scores were based on a comparison of the ions present in the experimental spectrum to the ions present in the library spectrum. A forward score of 100 would mean all the ions present in the experimental spectrum were present in the library at the correct ratios. Any deviations in ion ratios or additional experimental ions not present in the library reduced the forward score, thus the forward score is a good indication of the purity of the compound being detected. Co-elution with another molecule with the same mass add ions to the experimental spectrum and reduce the forward score. Similarly, a reverse score of 100 indicated that all ions present in the library were present in the experimental spectrum at the correct ratios and deviations in ion ratios or ions in the library not present in the experimental spectrum reduced the reverse score. Identification was automatically approved if all the above criteria were met and the MS/MS forward and reverse scores were above 80. Compounds which met the above criteria but had low MS/MS scores, below 35 for both forward and reverse, were automatically rejected. Compounds with intermediate MS/MS forward and reverse scores, 36–79, were marked for manual review. If an MS/MS spectrum was not obtained for a given ion, the identification was based on retention and parent mass alone and marked for analyst reviews. In this case, identification can still be confirmed if it has historical precedent in the specific matrix. Further details can be found in Evans et al.<sup>56</sup>

**Metabolite quantification and data normalisation:** Peaks were quantified using area-under-the-curve. Data was normalised, to correct variations that resulted from differences in the inter-day tuning of the instruments. Essentially, each compound was corrected in run-day blocks, by registering the medians to equal one, and normalising each data point proportionately. Each biochemical in OrigScale data was then rescaled, to set the median equal to 1. Compounds in which more than 50% of values were missing were not included in the statistical analyses. Scaled data are provided in Supplementary Data 2 and 15. Raw and OrigScale data for VP are provided in Supplementary Data 16 and 17. Raw serum data are provided in Supplementary Data 18. These tables include RI, accurate mass values, mean differences in the detected metabolite, and conversion to parts per million (PPM). Metabolomic data were log-transformed (applying the natural logarithm to the data plus one) before data analysis.

**Data analysis:** Principal Component Analysis (PCA) using R software was used to visualise the metabolomic data. Before PCA, data were imputed using a *k*-nearest neighbour (*k*NN) algorithm<sup>57</sup> (with  $k = 5$ ); they were then mean-centered and scaled to unit variance. Two-way ANOVA was used to compare the diets (irrespective of genotypes) or genotypes (irrespective of diets) and a *t* test was used for two groups' comparison (Supplementary Data 2 and Supplementary Data 15). Differences were considered significant if the *P* was  $<0.05$ ; and to account for multiple testing, a FDR<sup>58</sup> of  $<0.15$ . Qlucore Omics Explorer (<http://www.qlucore.com>; version 3.1) was used for heatmap representation and unsupervised clustering of metabolites that were significantly altered by HFD in a WT or a MYC context, or by MYC overexpression irrespective of the diet. Metabolites were grouped into 8 different classes (lipids, aminoacids, nucleotides, peptides, carbohydrates, cofactors and vitamins, energy, or xenobiotics), according to Metabolon's classification. Biochemical annotations were assigned by PhD level biochemists at Metabolon, integrating information from literature and public databases (e.g. HMDB). Metabolite Set Enrichment Analysis (MSEA) was performed using a hand-curated metabolite set (Supplementary Data 19) and run using the Gene Set Enrichment Analysis platform (GSEA; Broad Institute)<sup>33</sup> using 1000 permutations. Metabolite sets including fewer than three metabolites were excluded from the analysis. Metabolite sets were considered significantly enriched at  $P < 0.05$  and FDR  $< 0.15$ .

**Global chromatin profiling.** The global chromatin profiling assay was performed as described in Creech et al.<sup>59</sup> with the following modifications:

**Cell lysis, tissue lysis, and histone extraction:** Flash-frozen tissue samples, 10–40 mg in mass, were thawed on ice and resuspended in 200  $\mu$ L ice-cold PBS. Samples were homogenised for about 2 min using a motorised pestle (VWR, 47747–370), and were spun down at 4 °C, at 1500 g for 5 min. Supernatant was removed and 0.5 mL ice-cold nucleus buffer was added to the resultant pellet. Nuclei were centrifuged at 4 °C, at 10,000 g for 1 min and supernatant was removed. The nucleus isolation procedure was repeated twice, removing supernatant each time. Histones were extracted from the remaining pellet, with 400  $\mu$ L 0.4 N H<sub>2</sub>SO<sub>4</sub> at room temperature for 16 h, while shaking; at this point, histone isolation proceeded using the same protocol as described<sup>59</sup>. In addition to the flash-frozen tissue, histones were extracted from one 25 million cell pellet each of Arg-<sup>15</sup>N<sub>6</sub>,<sup>13</sup>C<sub>4</sub> SILAC-labeled HeLa, K562, and 293 T (as in Jaffe et al.<sup>29</sup>), following the protocol described by Creech et al.<sup>59</sup>.

**Histone derivatization:** The sample set used SILAC standardisation, with histones extracted from HeLa, K562 and 293T cell lines, as described above. In this workflow, input amount was reduced to 10  $\mu$ g per sample (5  $\mu$ g sample and 5  $\mu$ g SILAC heavy standards), based on the protocol. Samples were adjusted to 100 mM sodium phosphate, pH 8.0, by adding 3  $\mu$ L 500 mM sodium phosphate, pH 8.0; the total volume of the sample was brought up to 15  $\mu$ L with HPLC-grade water. Phosphate-buffered samples were reacted with 60  $\mu$ L of 400 mM NHS propionate in anhydrous methanol at room temperature, with shaking. Three hundred microliters of 0.1% trifluoroacetic acid (TFA) was added, to bring samples to a volumetric concentration of 20% organic solvent. Samples were desalted on a 96-well Oasis HLB 5 mg/cc plate (Waters, 186000309). Activation, equilibration, and wash volumes were 200  $\mu$ L for each step, and sample elution volume was 100  $\mu$ L. For the trypsin digestion, 1  $\mu$ g trypsin was used in 10  $\mu$ L of 50 mM ammonium bicarbonate, pH 8.0, while all other conditions were as described<sup>59</sup>. After digestion and lyophilization, new N-termini were derivatized, by resuspending peptides in 40  $\mu$ L of 400 mM NHS propionate/anhydrous methanol, and adjusting to 18 mM sodium phosphate, pH 8.0, with 10  $\mu$ L 100 mM sodium phosphate, pH 8.0. The reaction was quenched with 10  $\mu$ L 15% hydroxylamine solution and incubated for 30 min at room temperature with shaking. Samples were brought up to a total volume of 260  $\mu$ L with HPLC-grade water, frozen, and lyophilised via vacuum concentrator. Samples were resuspended in 200  $\mu$ L 0.1% TFA, and desalted on a SepPac tC18 96-well  $\mu$ Elution plate (Waters, 186002318). All activation and wash volumes were 200  $\mu$ L. Elution volume was 100  $\mu$ L. Desalted peptides were lyophilised via vacuum concentrator, and were brought up to a volume of 10  $\mu$ L with 3% acetonitrile (ACN)/5% formic acid (FA). Samples were further diluted 1:10 with 3% ACN/5% FA, before introducing them into the mass spectrometer.

**LC-MS/MS assay parameters:** The gradient was modified so that peptides were separated at a flow rate of 200 nL/minute, with a 60 min linear gradient from 97% solvent A (3% ACN/ 0.1% FA) to 33% solvent B (90% ACN/ 0.1% FA). This gradient was followed by a 15 min linear gradient, from 33% solvent B to 65% solvent B. This gradient was followed by a 5 min linear gradient from 65% solvent B to 90% solvent B, at which point the 90% solvent B was held for an additional 5 min. Including sample loading and column equilibration times, each sample took 120 min to completion, 90 min of which was taken up by active data acquisition.

**Scheduling for H3, H4, H2A, H2A.Z and H2B targets:** To determine each peptide's retention time, we employed a scheduling sample, comprising three samples in a 1:1:1 ratio instead of a synthetic peptide mix. Most method parameters were the same as in Creech et al.<sup>59</sup>, except that peptides were scheduled within a 23-min window, based on hypothesised elution time; also, the total run time for each scan was 0–90 min. The isolation width for MS1 and MS2 scans were narrowed to 1.7 m/z with a 0.3 m/z offset; these data were acquired on a Q-Exactive Plus (Thermo Scientific) mass spectrometer. A list of peptides targeted in addition to published histone marks in Creech et al.<sup>59</sup> is presented in Supplementary Data 20.

**Scheduled data acquisition:** After determining retention times, 1  $\mu$ L of sample was injected onto the same column that was utilised for scheduling, using the same gradients with previously described modifications. MS1 and MS2 scans used the same parameters as described in Creech et al.<sup>59</sup>, with the same scan run time and isolation width modifications as described above. The inclusion list was turned on for each MS2 scan, and included heavy as well as light versions of each peptide to be observed, its charge state, new acquisition windows based on the scheduling runs, and optimal collision energies.

**Heatmap generation:** GENE-E (<http://www.broadinstitute.org/cancer/software/GENE-E/>) was used for heatmap representation as well as statistical analysis of the data, using the comparative marker selection suite<sup>60</sup>. Differences were considered significant if the *p*-value was <0.05, and FDR was <0.1. Unsupervised clustering of histone marks (one minus Pearson correlation) was done on normalised values, based on the median level of each mark in the three WT prostate lobes (VP, DLP and AP).

**ChIP-sequencing.** The ChIP-sequencing was performed as described in Ku et al.<sup>61</sup>, with the following modifications. Fresh-frozen VP tissues from 12-week-old mice were pulverised (Cryoprep Impactor, Covaris), resuspended in PBS + 1% formaldehyde, and incubated at room temperature for 20 min. Fixation was stopped by the addition of 0.125 M glycine (final concentration) for 15 min at room

temperature, then washing in ice-cold PBS + EDTA-free protease inhibitor cocktail (PIC; #4693132001, Roche). Multiple biological replicates were combined for each condition in two distinct pools (replicates). Chromatin was isolated by the addition of lysis buffer (0.1% SDS, 1% Triton X-100, 10 mM Tris-HCl (pH 7.4), 1 mM EDTA (pH 8.0), 0.1% NaDOC, 0.13 M NaCl, 1X PIC) + sonication buffer (0.25% sarkosyl, 1 mM DTT) to the samples, which were maintained on ice for 30 min. Lysates were sonicated (E210 Focused-ultrasonicator, Covaris) and the DNA was sheared to an average length of ~200–500 bp. Genomic DNA (input) was isolated by treating sheared chromatin samples with RNase (30 min at 37 °C), proteinase K (30 min at 55 °C), de-crosslinking buffer (1% SDS, 100 mM NaHCO<sub>3</sub> (final concentration), 6–16 h at 65 °C), followed by purification (#28008, Qiagen). DNA was quantified on a NanoDrop spectrophotometer, using the Quant-iT High-Sensitivity dsDNA Assay Kit (#Q33120, Thermo Fisher Scientific). On ice, ChIP-validated H4K20me1 (2  $\mu$ g, #ab9051, Abcam) or PHF8 (5  $\mu$ g, #A301-772A, Bethyl Laboratories) antibodies<sup>62</sup> were conjugated to a mix of washed Dynalbeads protein A and G (Thermo Fisher Scientific), and incubated on a rotator (overnight at 4 °C) with 1.5  $\mu$ g (H4K20me1) or 5  $\mu$ g (PHF8) of chromatin. ChIP'ed complexes were washed, sequentially treated with RNase (30 min at 37 °C), proteinase K (30 min at 55 °C), de-crosslinking buffer (1% SDS, 100 mM NaHCO<sub>3</sub> (final concentration), 6–16 h at 65 °C), and purified (#28008, Qiagen). The concentration and size distribution of the immunoprecipitated DNA was measured using the Bioanalyzer High Sensitivity DNA kit (#5067-4626, Agilent). Dana-Farber Cancer Institute Molecular Biology Core Facilities prepared libraries from 2 ng of DNA, using the ThruPLEX DNA-seq kit (#R400427, Rubicon Genomics), according to the manufacturer's protocol; finished libraries were quantified by the Qubit dsDNA High-Sensitivity Assay Kit (#32854, Thermo Fisher Scientific), by an Agilent TapeStation 2200 system using D1000 ScreenTape (# 5067-5582, Agilent), and by RT-qPCR using the KAPA library quantification kit (# KK4835, Kapa Biosystems), according to the manufacturers' protocols; ChIP-seq libraries were uniquely indexed in equimolar ratios, and sequenced to a target depth of 40 M reads on an Illumina NextSeq500 run, with single-end 75 bp reads; Bowtie2 (version 2.2.1) was used to align the ChIP-seq datasets to build version NCB37/MM9 of the mouse genome<sup>63</sup>. Alignments were performed using default parameters that preserved reads mapping uniquely to the genome without mismatches.

**H4K20me1.** H4K20me1 read density between transcriptional start site (TSS) and transcriptional end site (TES) was averaged for each gene, and reported against the CTD\_WT (reference) for the HFD\_WT, CTD\_MYC or HFD\_MYC conditions. Waterfall plots of rank-ordered log<sub>2</sub>-fold changes were used to visualise H4K20me1 dynamic changes. Genes with a loss (<1.15 fold-change) or a gain (>1.15 fold-change) of the H4K20me1 mark between TSS and TES relative to the CTD\_WT (reference) were identified for the HFD\_WT, CTD\_MYC and HFD\_MYC conditions, and were associated with their corresponding transcript abundance. Venn diagrams were generated using the 'VennDiagram' R package (version 1.6.9).

**RNA-sequencing.** Fresh VP tissues from 12-week-old mice were dissociated to form a single cell suspension. RNA from a similar number of cells was extracted using the miRNeasy Micro Kit (#217084, Qiagen) coupled with on-column DNase treatment (#79254, Qiagen). RNA sample concentration was measured and subjected to quality evaluation, using a Bioanalyzer RNA 6000 Nano kit (#5067-1511, Agilent). The Dana-Farber Cancer Institute Molecular Biology Core Facilities prepared libraries from 500 ng of purified total RNA, using TruSeq Stranded mRNA sample preparation kits (#RS-122-2101, Illumina) according to the manufacturer's protocol; submitted the finished libraries to quality control analyses as described in the ChIP-seq Methods section, pooled uniquely indexed RNA-seq libraries in equimolar ratios, and sequenced these to a target depth of 40M reads on an Illumina NextSeq500 run with single-end 75 bp reads. Fastq files were aligned to the mm9 genome using tophat with default parameters (version 2.0.11). Transcript abundances were calculated using the cuffquant module of Cufflinks (version 2.2.0). FPKM values were calculated and normalised using the cuffnorm module of Cufflinks (version 2.2.0). Paired *t*-test was calculated using the *t.test* function in R (version 3.3.2).

**Murine gene set enrichment analysis and MYC signature.** Gene expression values from biological triplicates were input for Gene Set Enrichment Analysis (GSEA)<sup>33</sup> using the Hallmark (H, v5.01; Supplementary Data 7) or the Chemical and Genetic Perturbations (C2.cgp, v5.1; Supplementary Data 8) Molecular Signature Databases (MSigDB) with 10,000 permutations. The Normalised Enrichment Score (NES)—associated with gene sets that were significantly enriched or depleted (*p* < 0.05 and FDR < 0.1)—was used for heatmap generation, using a custom-made R script. A murine prostatic MYC signature was obtained by combining leading edge genes from all MYC-related gene sets that were significantly enriched (*P* < 0.05 and FDR < 0.1) in the H and C2.cgp MSigDB (Supplementary Data 9). Aggregate read density profiles of PHF8 and H4K20me1, and their quantification around MYC signature genes, were generated using deepTools<sup>64</sup>. Mapped regions were visualised using the Integrated Genomics Viewer (IGV, version 2.3.68)<sup>65</sup>.

**Protein analysis.** Fresh-frozen VP tissues from 12-week-old mice were pulverised (Cryoprep Pulvizer, Covaris) and lysed on ice in RIPA buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% NP-40) with the addition of phosphatases and protease inhibitor cocktail tablets (Complete Mini, EDTA-free, Roche). MYC-CaP cells (kindly provided by Dr. Charles Sawyers, Memorial Sloan Kettering Cancer Center, New York, NY<sup>66</sup>) were rinsed on ice with PBS and lysed as for the mouse prostates. Cells were authenticated via STR profiling (DDC Medical, 16 January 2015). Cells were tested negative for mycoplasma contamination using MycoAlert™ Mycoplasma Detection Kit (Lonza). Equal amounts of protein (15–20 µg; Bradford protein assay, Bio-Rad) were resolved on precast 4–12 or 4–20% Tris-glycine SDS-polyacrylamide gels (Invitrogen), and transferred to Nitrocellulose Blotting membranes (Amersham), following standard procedures. Membranes were probed with the following antibodies according to the manufacturer's instructions: rabbit monoclonal [Y69] anti-c-MYC (#ab32072, Abcam), or rabbit polyclonal anti-β-Actin (#4967, Cell Signaling Technology). Densitometry analyses were made with ImageJ (U.S. NIH, Bethesda, MD; <http://imagej.nih.gov/ij/>). Results were normalised to β-actin and expressed as arbitrary units.

**Epidemiological studies.** *Study population:* We tested our hypothesis among prostate cancer patients who were enrolled in two prospective studies: the Physicians' Health Study (PHS) and the Health Professionals Follow-up Study (HPFS). PHS I and II began in 1982 and 1997, respectively, as randomised trials of aspirin (PHS I) and dietary supplements (PHS II), and enrolled 29,067 male U.S. physicians for the primary prevention of cardiovascular disease and cancer<sup>67–70</sup>. The HPFS was initiated in 1986, when 51,529 U.S. men, 40–75 years of age and working in health professions, completed a biennial questionnaire mailed to them<sup>71</sup>. In both studies, participants were followed by means of regular questionnaires, and self-reported data on diet, lifestyle behaviours, medical history, and disease outcomes were collected. We confirmed the incidence of prostate cancer cases in this population by reviewing medical records and pathology reports. Following the confirmation of diagnosis, we retrieved archival formalin-fixed paraffin-embedded (FFPE) prostate tissue specimens, collected during radical prostatectomy or transurethral resection of the prostate. Pathologists undertook a standardised histopathologic review, including Gleason grading<sup>72</sup>, and standardised clinical data were abstracted from medical records. Deaths were ascertained via mail, telephone, and through periodic systematic searches of the National Death Index. Lethal prostate cancer was defined as the occurrence of distant metastases, or death due to prostate cancer. Men were followed through March 2011 for PHS and through December 2011 for HPFS. We obtained written informed consent from all participants, and the study was approved by institutional review boards at the Harvard T.H. Chan School of Public Health and Partners Health Care.

*Whole-transcriptome expression profiling:* In the current study, we undertook gene expression profiling of archival tumour tissue among 402 men with prostate cancer in the cohorts using an extreme case control design. Cases were men with lethal prostate cancer (developed metastatic disease or died from prostate cancer) and controls were men with indolent cancer (those survived at least 8 years after prostate cancer diagnosis, without any evidence of metastases). In total, there were 113 lethal cases and 289 indolent cases. We also included adjacent normal tissue for a subset of these tumour tissues ( $n = 200$ ). Gene expression profiling of archival FFPE tissue was performed as described<sup>73</sup>. Briefly, two to three 0.6-mm cores were sampled from regions of high-density tumour, and from adjacent normal prostate tissue. RNA was extracted with the Agencourt FormaPure kit (Beckman Coulter), with use of the Biomek FX<sup>P</sup> automated platform. Whole-transcriptome amplification was performed using WT-Ovation FFPE System V2 (NuGEN) and the amplified cDNA was hybridised to a GeneChip Human Gene 1.0 ST microarray (Affymetrix). For the expression profiles generated, we regressed out technical variables and then shifted the residuals to derive the original mean expression values, and normalised these using the robust multi-array average method<sup>74,75</sup>. NetAffx annotations were used to map gene names to Affymetrix transcript cluster IDs, as implemented in the Bioconductor annotation package `pd.hugene.1.0.st.v1`; this resulted in 20,254 unique gene names.

*Diet assessment:* Self-administered semi-quantitative food frequency questionnaires (FFQs) were collected every four years from 1986 for the HPFS, and were administered once between 1999 and 2002 for the PHS. The FFQs asked men to report their usual intake of approximately 130 foods and beverages during the previous year, and also their fried food consumption, the type of cooking fat they used, and whether they consumed the visible fat on meat. Fat intake levels were estimated by multiplying the frequency of intake by the amount of the fat in the specific portion of each food (based on nutrient composition data from the US Department of Agriculture, supplemented with food manufacturer data), and were summed across all foods. The FFQ was validated among 127 men in the HPFS. The correlations between the FFQ and four prospectively collected one-week weighed diet records were 0.67 for total fat, and 0.75 for saturated fat<sup>76</sup>. Because FFQ was mainly administered after the diagnosis of prostate cancer for the PHS participants, we estimated post-diagnostic fat intakes in both HPFS and PHS, to maintain a bigger sampler size and harmonize the two cohorts. In HPFS, we calculated cumulative average post-diagnostic intake from the FFQ preceding diagnosis until the end of the follow-up in HPFS<sup>77</sup>. Fat intake (g/d) was multiplied by 9 kcal and divided by total calories per day to calculate the percent of daily calories from each fat of interest.

*Statistical analysis:* Fat intake after diagnosis was estimated in 4577 men enrolled in the HPFS and in 926 men from the PHS, all of whom had non-metastatic prostate cancer. Cohort-specific quintiles were determined based on fat intake distributions for each cohort, with the highest quintile denoted as the high-fat group and the lower four quintiles grouped as the low-fat group (Supplementary Data 21). The categorised fat intake groups were then integrated with gene expression data in tumour or in adjacent normal tissues. Finally, we had 319 tumour tissues from patients (213 from the HPFS and 106 from the PHS) for whom we had complete fat intake estimation (animal fat: high-fat group  $n = 65$  vs. low-fat group  $n = 254$ ; saturated fat: high-fat group  $n = 62$  vs. low-fat group  $n = 257$ ; monounsaturated fat: high-fat group  $n = 66$  vs. low-fat group  $n = 253$ ; polyunsaturated fat: high-fat group  $n = 55$  vs. low-fat group  $n = 264$ ) and a total of 157 adjacent normal tissues after merging with fat intake data (animal fat: high-fat group  $n = 33$  vs. low-fat group  $n = 124$ ; saturated fat: high-fat group  $n = 29$  vs. low-fat group  $n = 128$ ; monounsaturated fat: high-fat group  $n = 33$  vs. low-fat group  $n = 124$ ; polyunsaturated fat: high-fat group  $n = 24$  vs. low-fat group  $n = 133$ ).

*Gene set enrichment analysis:* Gene expression profiles of tumour and adjacent normal prostate tissues were input for GSEA<sup>33</sup>, with use of the Hallmark (H, v4.0) MSigDB with 10,000 phenotype-based permutations, to identify predefined sets of functionally related genes correlated with specific fat intakes (Supplementary Data 10, 12–14). Gene sets with  $P < 0.05$  and  $FDR < 0.1$  were considered for subsequent analyses. Animal fat and saturated fat intake-dependent MYC signatures were obtained by combining either the leading edge or the non-leading edge genes from the MYC\_targets\_V1 gene set from the H MSigDB in tumour tissues (Supplementary Data 11), to create a metagene score as previously described<sup>78</sup>. This was computed for each sample by averaging the normalised (mean-centered and variance scaled) expression values of all member genes. An additional signature was derived from 113 randomly selected genes from the MYC\_targets\_V1 gene set (Supplementary Data 11). Odds ratios and 95% confidence intervals were obtained by logistic regression for the association between the metagene score and lethal prostate cancer. The score was modelled as categorical (tertiles). We tested for linear trend across score categories by modelling the tertiles as a continuous variable. All models were adjusted for age and year at diagnosis. We further adjusted for Gleason grade to test whether the score is an independent predictor of lethal prostate cancer and BMI at diagnosis, to differentiate the effect from overweight/obesity. To assess whether the association between the score and lethal prostate cancer was modified by saturated fat intake, we obtained  $P$  for interaction by including an interaction term (saturated fat intake  $\times$  MYC score) in the multivariable model using a Wald test. All analyses were conducted using SAS version 9.3 and R version 3.1.0.

*Validation cohorts:* To investigate the power of SFI-induced and non-SFI-induced MYC signatures to predict metastatic disease, we utilised genome-wide expression profiles of 751 patients with metastatic outcome follow-up from the Decipher Genomic Resource Information Database (GRID; NCT02609269). These patients were pooled from four studies of either case-cohort or cohort design. Patients for these studies came from four institutes: Thomas Jefferson University (TJU;  $n = 139$ )<sup>79</sup>, Johns Hopkins Medical Institutions-I (JHMI-I;  $n = 260$ )<sup>80</sup>, Mayo Clinic ( $n = 235$ )<sup>81</sup>, Cedars-Sinai ( $n = 117$ )<sup>82</sup>. A total of 120 non-randomly selected patients from case-cohort studies were removed before pooling the studies to avoid bias in estimating the hazard ratio. 631 patients were thus eligible for analysis, 70 of which developed metastasis. Median follow-up time for censored patients was 8 years and the median age at radical prostatectomy was 61 years.

The fat-induced MYC signature (113 genes) and non-fat-induced MYC signature (87 genes) were used to calculate pathway expression scores for each patient, using a z-score scaled, mean gene expression. Based on the tertiles of these scores, patients were divided into three groups with T1 being the lowest and T3 the highest. Kaplan–Meier curves and Cox proportional hazard regression were used to evaluate the metastatic prognosis. To test associations between signatures and BMI, we extracted BMI data from 494 patients pooled from three cohorts (TJU,  $n = 139$ ; JHMI-I only,  $n = 144$ ; JHMI-II<sup>83</sup> only,  $n = 95$ ; JHMI-I/II,  $n = 116$ ). Correlation analysis using Pearson's correlation was used to measure the association between MYC signatures score and BMI. JHMI-II was excluded from the survival analysis because only patients that developed biochemical recurrence were selected for this study, hence it was statistically inappropriate to pool the JHMI-II cohort with the others lacking this inclusion criteria as it would inflate the event rate. We also conducted univariate and multivariate analyses to associate the SFI-induced MYC signature with clinical outcome after adjusting for other clinicopathologic variables including pre-operative prostate-specific antigen (PSA) levels, seminal vesicle invasion, surgical margins, extracapsular extension, lymph node invasion, gleason grade or the Cell Cycle Progression score in the pooled cohort from which we utilised genome-wide expression profiles of 631 patients (deidentified and aggregated from routine clinical use of the Decipher prostate cancer classifier test; Decipher Biosciences Laboratory, San Diego, CA) with metastatic outcome follow-up from the Decipher GRID.

**Adequacy of statistical analyses.** All the statistical tests were justified as appropriate. Assumption criteria were met, analysis of variance was performed. When variance was not equal, Welch's t-test (unequal variance t-test) was applied. Data are reported including estimation of variation within each group. Two-sided tests were used. Measurements were taken from distinct samples.

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

## Data availability

Data are available from the corresponding authors upon request. Raw data underlying reported averages in graphs and uncropped versions of blots are provided in the source data file or supplementary tables. Raw metabolomics data generated by Metabolon were deposited on MetaboLights and are available through the study identifier [MTBL135](https://www.ebi.ac.uk/metabolights/study/MTBL135). Raw, scaled metabolomics data, and statistics were also provided as supplementary tables. The sequencing data reported in this paper (ChIP-seq and RNA-seq) were deposited on NCBI Gene Expression Omnibus (GEO) and are accessible through GEO Series accession number [GSE90912](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi). Human gene expression data is available through [GSE62872](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi) [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi>].

Received: 6 September 2018 Accepted: 23 August 2019

Published online: 25 September 2019

## References

- Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2017. *CA Cancer J. Clin.* **67**, 7–30 (2017).
- Van Blarigan, E. L. et al. Fat intake after prostate cancer diagnosis and mortality in the Physicians' Health Study. *Cancer Causes Control* **26**, 1117–1126 (2015).
- Cao, Y. & Ma, J. Body mass index, prostate cancer-specific mortality, and biochemical recurrence: a systematic review and meta-analysis. *Cancer Prev. Res. (Phila)* **4**, 486–501 (2011).
- Strom, S. S. et al. Saturated fat intake predicts biochemical failure after prostatectomy. *Int J. Cancer* **122**, 2581–2585 (2008).
- Meyer, F., Bairati, I., Shadmani, R., Fradet, Y. & Moore, L. Dietary fat and prostate cancer survival. *Cancer Causes Control* **10**, 245–251 (1999).
- Ng, M. et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **384**, 766–781 (2014).
- Danaei, G. et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *Lancet* **378**, 31–40 (2011).
- Farzadfar, F. et al. National, regional, and global trends in serum total cholesterol since 1980: systematic analysis of health examination surveys and epidemiological studies with 321 country-years and 3.0 million participants. *Lancet* **377**, 578–586 (2011).
- Labbé, D. P. et al. PTP1B deficiency enables the ability of a high-fat diet to drive the invasive character of PTEN-deficient prostate cancers. *Cancer Res* **76**, 3130–3135 (2016).
- Blando, J. et al. Dietary energy balance modulates prostate cancer progression in Hi-Myc mice. *Cancer Prev. Res. (Phila)* **4**, 2002–2014 (2011).
- Kobayashi, N. et al. Effect of low-fat diet on development of prostate cancer and Akt phosphorylation in the Hi-Myc transgenic mouse model. *Cancer Res* **68**, 3066–3073 (2008).
- Zeigler-Johnson, C., Morales, K. H., Lal, P. & Feldman, M. The Relationship between obesity, prostate tumor infiltrating lymphocytes and macrophages, and biochemical failure. *PLoS One* **11**, e0159109 (2016).
- Allott, E. H. & Hursting, S. D. Obesity and cancer: mechanistic insights from transdisciplinary studies. *Endocr. Relat. Cancer* **22**, R365–R386 (2015).
- Liu, J. et al. High-calorie diet exacerbates prostate neoplasia in mice with haploinsufficiency of Pten tumor suppressor gene. *Mol. Metab.* **4**, 186–198 (2015).
- Labbé, D. P. et al. Role of diet in prostate cancer: the epigenetic link. *Oncogene* **34**, 4683–4691 (2015).
- Kinnaird, A., Zhao, S., Wellen, K. E. & Michelakis, E. D. Metabolic control of epigenetics in cancer. *Nat. Rev. Cancer* **16**, 694–707 (2016).
- Labbé, D. P., Brown, M. Transcriptional regulation in prostate cancer. *Cold Spring Harb. Perspect. Med.* **8**, a030437 (2018).
- Gurel, B. et al. Nuclear MYC protein overexpression is an early alteration in human prostate carcinogenesis. *Mod. Pathol.* **21**, 1156–1167 (2008).
- Cancer Genome Atlas Research N. The molecular taxonomy of primary prostate. *Cell* **163**, 1011–1025 (2015).
- Ribeiro, F. R. et al. 8q gain is an independent predictor of poor survival in diagnostic needle biopsies from prostate cancer suspects. *Clin. Cancer Res* **12**, 3961–3970 (2006).
- Ellwood-Yen, K. et al. Myc-driven murine prostate cancer shares molecular features with human prostate tumors. *Cancer Cell* **4**, 223–238 (2003).
- Eilers, M. & Eisenman, R. N. Myc's broad reach. *Genes Dev.* **22**, 2755–2766 (2008).
- Dang C. V. MYC, metabolism, cell growth, and tumorigenesis. *Cold Spring Harb. Perspect. Med.* **3**, a014217 (2013).
- Hsieh, A. L., Walton, Z. E., Altman, B. J., Stine, Z. E. & Dang, C. V. MYC and metabolism on the path to cancer. *Semin Cell Dev. Biol.* **43**, 11–21 (2015).
- Stine, Z. E., Walton, Z. E., Altman, B. J., Hsieh, A. L. & Dang, C. V. MYC, metabolism, and cancer. *Cancer Discov* **5**, 1024–1039 (2015).
- Priolo, C. et al. AKT1 and MYC induce distinctive metabolic fingerprints in human prostate cancer. *Cancer Res* **74**, 7198–7204 (2014).
- Lu, C. & Thompson, C. B. Metabolic regulation of epigenetics. *Cell Metab.* **16**, 9–17 (2012).
- Klose, R. J., Kallin, E. M. & Zhang, Y. JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet* **7**, 715–727 (2006).
- Jaffe, J. D. et al. Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. *Nat. Genet* **45**, 1386–1391 (2013).
- 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).
- Vavouri, T. & Lehner, B. Human genes with CpG island promoters have a distinct transcription-associated chromatin organization. *Genome Biol.* **13**, R110 (2012).
- Qi, H. H. et al. Histone H4K20/H3K9 demethylase PHF8 regulates zebrafish brain and craniofacial development. *Nature* **466**, 503–507 (2010).
- Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
- Walz, S. et al. Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles. *Nature* **511**, 483–487 (2014).
- Fortschegger, K. et al. PHF8 targets histone methylation and RNA polymerase II to activate transcription. *Mol. Cell Biol.* **30**, 3286–3298 (2010).
- Bjorkman, M. et al. Systematic knockdown of epigenetic enzymes identifies a novel histone demethylase PHF8 overexpressed in prostate cancer with an impact on cell proliferation, migration and invasion. *Oncogene* **31**, 3444–3456 (2012).
- Cuzick, J. et al. Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study. *Lancet Oncol.* **12**, 245–255 (2011).
- Bostwick, D. G., Liu, L., Brawer, M. K. & Qian, J. High-grade prostatic intraepithelial neoplasia. *Rev. Urol.* **6**, 171–179 (2004).
- Carrer, A. & Wellen, K. E. Metabolism and epigenetics: a link cancer cells exploit. *Curr. Opin. Biotechnol.* **34**, 23–29 (2015).
- Etchegaray, J. P. & Mostoslavsky, R. Interplay between metabolism and epigenetics: a nuclear adaptation to environmental changes. *Mol. Cell* **62**, 695–711 (2016).
- Su, X., Wellen, K. E. & Rabinowitz, J. D. Metabolic control of methylation and acetylation. *Curr. Opin. Chem. Biol.* **30**, 52–60 (2016).
- Gut, P. & Verdin, E. The nexus of chromatin regulation and intermediary metabolism. *Nature* **502**, 489–498 (2013).
- Ma, Q. et al. The histone demethylase PHF8 promotes prostate cancer cell growth by activating the oncomiR miR-125b. *Onco Targets Ther.* **8**, 1979–1988 (2015).
- Taylor, B. S. et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell* **18**, 11–22 (2010).
- Zadra G., Loda M. Metabolic vulnerabilities of prostate cancer: diagnostic and therapeutic opportunities. *Cold Spring Harb. Perspect. Med.* **8**, a030569 (2018).
- Morrish, F., Isern, N., Sadilek, M., Jeffrey, M. & Hockenbery, D. M. c-Myc activates multiple metabolic networks to generate substrates for cell-cycle entry. *Oncogene* **28**, 2485–2491 (2009).
- Morrish, F. et al. Myc-dependent mitochondrial generation of acetyl-CoA contributes to fatty acid biosynthesis and histone acetylation during cell cycle entry. *J. Biol. Chem.* **285**, 36267–36274 (2010).
- Edmunds, L. R. et al. c-Myc programs fatty acid metabolism and dictates acetyl-CoA abundance and fate. *J. Biol. Chem.* **289**, 25382–25392 (2014).
- Chen, M. et al. An aberrant SREBP-dependent lipogenic program promotes metastatic prostate cancer. *Nat. Genet* **50**, 206–218 (2018).
- 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).
- Pettersson, A. et al. MYC overexpression at the protein and mRNA level and cancer outcomes among men treated with radical prostatectomy for prostate cancer. *Cancer Epidemiol. Biomark. Prev.* **27**, 201–207 (2018).
- Zubarev, R. A. & Makarov, A. Orbitrap mass spectrometry. *Anal. Chem.* **85**, 5288–5296 (2013).
- Evans, A. M. et al. High resolution mass spectrometry improves data quantity and quality as compared to unit mass resolution mass spectrometry in high-throughput profiling metabolomics. *Metabolomics* **4**, 132 (2014).
- Dehaven, C. D., Evans, A. M., Dai, H. & Lawton, K. A. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *J. Cheminform* **2**, 9 (2010).
- Dehaven C. D., Evans A. M., Dai H., Lawton K. A. Software Techniques for Enabling High-Throughput Analysis of Metabolomic Datasets. in *Metabolomics* (eds Roessner, U.). (IntechOpen, 2012)

56. Evans, A. M., DeHaven, C. D., Barrett, T., Mitchell, M. & Milgram, E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal. Chem.* **81**, 6656–6667 (2009).
57. Troyanskaya, O. et al. Missing value estimation methods for DNA microarrays. *Bioinformatics* **17**, 520–525 (2001).
58. Storey, J. D. A direct approach to false discovery rates. *J. R. Stat. Soc.: Ser. B (Stat. Methodol.)* **64**, 479–498 (2002).
59. Creech, A. L. et al. Building the connectivity map of epigenetics: chromatin profiling by quantitative targeted mass spectrometry. *Methods* **72**, 57–64 (2015).
60. Gould, J., Getz, G., Monti, S., Reich, M. & Mesirov, J. P. Comparative gene marker selection suite. *Bioinformatics* **22**, 1924–1925 (2006).
61. Ku, S. Y. et al. Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance. *Science* **355**, 78–83 (2017).
62. Ram, O. et al. Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell* **147**, 1628–1639 (2011).
63. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
64. Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187–W191 (2014).
65. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
66. Watson, P. A. et al. Context-dependent hormone-refractory progression revealed through characterization of a novel murine prostate cancer cell line. *Cancer Res* **65**, 11565–11571 (2005).
67. Steering Committee of the Physicians' Health Study Research Group. Final report on the aspirin component of the ongoing Physicians' Health Study. *New Engl. J. Med.* **321**, 129–135 (1989).
68. Hennekens, C. H. et al. Lack of effect of long-term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease. *New Engl. J. Med.* **334**, 1145–1149 (1996).
69. Gaziano, J. M. et al. Vitamins E and C in the prevention of prostate and total cancer in men: the Physicians' Health Study II randomized controlled trial. *Jama* **301**, 52–62 (2009).
70. Gaziano, J. M. et al. Multivitamins in the prevention of cancer in men: the Physicians' Health Study II randomized controlled trial. *Jama* **308**, 1871–1880 (2012).
71. Giovannucci, E., Liu, Y., Platz, E. A., Stampfer, M. J. & Willett, W. C. Risk factors for prostate cancer incidence and progression in the health professionals follow-up study. *Int. J. Cancer J. Int. du cancer* **121**, 1571–1578 (2007).
72. Stark, J. R. et al. Gleason score and lethal prostate cancer: does 3 + 4 = 4 + 3? *J. Clin. Oncol.: Off. J. Am. Soc. Clin. Oncol.* **27**, 3459–3464 (2009).
73. Penney, K. L. et al. Association of prostate cancer risk variants with gene expression in normal and tumor tissue. *Cancer Epidemiol. Biomark. Prev.* **24**, 255–260 (2015).
74. Irizarry, R. A. et al. Summaries of Affymetrix GeneChip probe level data. *Nucleic acids Res.* **31**, e15 (2003).
75. Irizarry, R. A. et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).
76. Rimm, E. B. et al. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am. J. Epidemiol.* **135**, 1114–1126 (1992). discussion 1127–1136.
77. Richman, E. L. et al. Fat intake after diagnosis and risk of lethal prostate cancer and all-cause mortality. *JAMA Intern. Med.* **173**, 1318–1326 (2013).
78. Ebot, E. M. et al. Gene expression profiling of prostate tissue identifies chromatin regulation as a potential link between obesity and lethal prostate cancer. *Cancer* **123**, 4130–4138 (2017).
79. Den, R. B. et al. Genomic prostate cancer classifier predicts biochemical failure and metastases in patients after postoperative radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **89**, 1038–1046 (2014).
80. Ross, A. E. et al. Tissue-based genomics augments post-prostatectomy risk stratification in a natural history cohort of intermediate- and high-risk men. *Eur. Urol.* **69**, 157–165 (2016).
81. Karnes, R. J. et al. Validation of a genomic classifier that predicts metastasis following radical prostatectomy in an at risk patient population. *J. Urol.* **190**, 2047–2053 (2013).
82. Freedland, S. J. et al. Utilization of a genomic classifier for prediction of metastasis following salvage radiation therapy after radical prostatectomy. *Eur. Urol.* **70**, 588–596 (2016).
83. Benzon, B. et al. Correlation of B7-H3 with androgen receptor, immune pathways and poor outcome in prostate cancer: an expression-based analysis. *Prostate Cancer Prostatic Dis.* **20**, 28–35 (2017).

## Acknowledgements

We thank Zach Herbert, Julie Batista, Noriko Uetani, Cornelia Photopoulos, Radha Kalekar, Maura Cotter, Natalia Scaglia, Teresa Bowman, Nadia Boufaeid (technical assistance and/or helpful discussions), and Sonal Jhaveri (critical review of the manuscript). D.P.L. is a Lewis Katz—Young Investigator of the Prostate Cancer Foundation, and is the recipient of a Scholarship for the Next Generation of Scientists from the Cancer Research Society, and is also recipient of a Canadian Institute of Health Research Fellowship. G.Z. is a recipient of the Idea Development Award from the U.S. Department of Defense (DoD; PC150263) and the Barr Award from the Dana-Farber Cancer Institute. C.Y.L. is a CPRIT Scholar in Cancer Research (RR150093), a Pew-Stewart Scholar for Cancer Research (Alexander and Margaret Stewart Trust), and is funded by the NIH and NCI (1R01CA215452-01). E.M.E. is supported by the DoD Prostate Cancer Research Program Postdoctoral Training Award (W81XWH-14-1-0250). L.E., D.E.S. and L.A.M. are Young Investigators of the Prostate Cancer Foundation. Support for HPFS/PHS cohorts was provided by grants from the DoD (W81XWH-11-1-0529) and grants from the National Institute of Health (NIH; CA42182, CA58684, CA90598, CA141298, CA97193, CA34944, CA40360, CA131945, CA167552, P50CA090381, 1U54CA155626-01, P30DK046200, HL26490 and HL34595). We would like to thank the participants and staff of the HPFS/PHS for their valuable contributions as well as the following state cancer registries for their help: A.L.Z., A.Z., A.R., C.A., C.O., C.T., D.E., F.L., G.A., I.D., I. L., I.N., I.A., K.Y., L.A., M.E., M.D., M.A., M.L., N.E., N.H., N.J., N.Y., N.C., N.D., O.H., O.K., O.R., P.A., R.I., S.C., T.N., T.X., V.A., W.A., W.Y. We assume full responsibility for analyses and interpretation of these data. The work reported here was supported by grants from the NIH (R01CA131945, R01CA187918 to M.L. and P50CA090381 to P.W.K., M.L. and M.B.), the NCI (1P01CA163227 to M.B.) and the Prostate Cancer Foundation to M.L. and M.B.

## Author contributions

D.P.L., G.Z., M.L. and M.B. conceived the study and designed the experiments. Wet lab experiments were performed by D.P.L., G.Z. and A.L.C. and supported by H.E., S.S., L.E., J.D.J., M.L. and M.B. E.D.K. performed MS-based metabolomics analysis. Computational analyses were performed by D.P.L., M.Y., J.M.R., C.Y.L., S.C., M.A. and N.E. and supported by E.M.E., E.A.G., M.T., J.L., E.D., A.V.D., P.W.K., J.E.B., L.A.M., J.E.C., M.L. and M.B. Access to validation cohorts data was provided by A.R., E.M.S., R.B.D., R.J.K., S.J.F. and D.E.S. Pathological analyses were performed by F.G., M.F. & M.L. D.P.L., G.Z., M.Y., S.C., A.L.C., M.L. and M.B. interpreted the data and drafted the paper considering inputs from all co-authors.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-12298-z>.

**Competing interests:** C.Y.L. receives sponsored research from and consults for Kronos Bio, is a shareholder and inventor of IP licensed to Syros Pharmaceuticals, is a shareholder of Amgen, and is an equity partner of Cambridge Science Corporation. M.A., N.E., M.T., J.L., E.A.G. and E.D. are employees of Decipher Biosciences. E.D.K. is currently employed of Metabolon. M.B. receives sponsored research support from Novartis. M.B. is a consultant to Aleta Biotherapeutics and H3 Biomedicine and serves on the SAB of Kronos Bio. The remaining authors declare no competing interests.

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

**Peer review information** *Nature Communications* thanks Ian Mills, Zheng-Jiang Zhu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, 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) 2019

David P. Labbé<sup>1,2,23,25</sup>, Giorgia Zadra<sup>3,4,25</sup>, Meng Yang<sup>5</sup>, Jaime M. Reyes<sup>1</sup>, Charles Y. Lin<sup>6</sup>, Stefano Cacciatore<sup>7</sup>, Ericka M. Ebot<sup>8</sup>, Amanda L. Creech<sup>9</sup>, Francesca Giunchi<sup>10</sup>, Michelangelo Fiorentino<sup>10</sup>, Habiba Elfandy<sup>3</sup>, Sudeepa Syamala<sup>3</sup>, Edward D. Karoly<sup>11</sup>, Mohammed Alshalalfa<sup>12</sup>, Nicholas Erho<sup>12</sup>, Ashley Ross<sup>13</sup>, Edward M. Schaeffer<sup>14</sup>, Ewan A. Gibb<sup>12</sup>, Mandeep Takhar<sup>12</sup>, Robert B. Den<sup>15</sup>, Jonathan Lehrer<sup>12</sup>, R. Jeffrey Karnes<sup>16</sup>, Stephen J. Freedland<sup>17,18</sup>, Elai Davicioni<sup>12</sup>, Daniel E. Spratt<sup>19</sup>, Leigh Ellis<sup>3,4,9</sup>, Jacob D. Jaffe<sup>9</sup>, Anthony V. D'Amico<sup>20</sup>, Philip W. Kantoff<sup>1,21</sup>, James E. Bradner<sup>1</sup>, Lorelei A. Mucci<sup>8,22</sup>, Jorge E. Chavarro<sup>5,8,22</sup>, Massimo Loda<sup>3,4,9,24</sup> & Myles Brown<sup>1,2</sup>

<sup>1</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA. <sup>2</sup>Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA, USA. <sup>3</sup>Department of Oncologic Pathology, Dana-Farber Cancer Institute, Boston, MA, USA. <sup>4</sup>Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA. <sup>5</sup>Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA. <sup>6</sup>Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA. <sup>7</sup>Cancer Genomics Group, International Centre for Genetic Engineering and Biotechnology, Cape Town, South Africa. <sup>8</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. <sup>9</sup>The Broad Institute of MIT and Harvard University, Cambridge, MA, USA. <sup>10</sup>Pathology Service, Addarii Institute of Oncology, S-Orsola-Malpighi Hospital, Bologna, IT, Italy. <sup>11</sup>Metabolon, Morrisville, NC, USA. <sup>12</sup>Decipher Biosciences, Vancouver, BC, Canada. <sup>13</sup>James Buchanan Brady Urological Institute, Johns Hopkins Medical Institutions, Baltimore, MD, USA. <sup>14</sup>Northwestern University Feinberg School of Medicine, Chicago, IL, USA. <sup>15</sup>Department of Radiation Oncology, Sidney Kimmel Medical College at Thomas Jefferson University, Philadelphia, PA, USA. <sup>16</sup>Department of Urology, Mayo Clinic Rochester, Rochester, MN, USA. <sup>17</sup>Department of Surgery, Division of Urology, Center for Integrated Research on Cancer and Lifestyle, Samuel Oschin Comprehensive Cancer Center, Cedars-Sinai Medical Center, Los Angeles, CA, USA. <sup>18</sup>Surgery Section, Durham Veteran Affairs Medical Center, Durham, NC, USA. <sup>19</sup>Department of Radiation Oncology, University of Michigan, Ann Arbor, MI, USA. <sup>20</sup>Department of Radiation Oncology, Dana-Farber Cancer Institute and Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. <sup>21</sup>Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>22</sup>Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. <sup>23</sup>Present address: Division of Urology, Department of Surgery, McGill University and Research Institute of the McGill University Health Centre, Montréal, QC, Canada. <sup>24</sup>Present address: Department of Pathology and Laboratory Medicine, Weil Cornell Medicine, New York Presbyterian-Weill Cornell Campus, New York, NY, USA. <sup>25</sup>These authors contributed equally: David P. Labbé, Giorgia Zadra.