EBioMedicine xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

Integrative (*epi*) Genomic Analysis to Predict Response to Androgen-Deprivation Therapy in Prostate Cancer

Sukanya Panja [a], Sheida Hayati [a], Nusrat J. Epsi [a], James Scott Parrott [b], Antonina Mitrofanova [a,c,*]

[a] Department of Health Informatics, Rutgers School of Health Professions, Rutgers Biomedical and Health Sciences, Newark, NJ 07107, USA
[b] Department of Interdisciplinary Studies, Rutgers School of Health Professions, Rutgers Biomedical and Health Sciences, Newark, NJ 07107, USA
[c] Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901, USA

ARTICLE INFO

Article history:
Received 12 January 2018
Received in revised form 24 March 2018
Accepted 5 April 2018
Available online xxxx

Keywords:
Therapeutic resistance
DNA methylation
mRNA expression
Epigenomics
Androgen-deprivation
Prostate cancer

ABSTRACT

Therapeutic resistance is a central problem in clinical oncology. We have developed a systematic genome-wide computational methodology to allow prioritization of patients with favorable and poor therapeutic response. Our method, which integrates DNA methylation and mRNA expression data, uncovered a panel of 5 differentially methylated sites, which explain expression changes in their site-harboring genes, and demonstrated their ability to predict primary resistance to androgen-deprivation therapy (ADT) in the *TCGA* prostate cancer patient cohort (hazard ratio = 4.37). Furthermore, this panel was able to accurately predict response to ADT across independent prostate cancer cohorts and demonstrated that it was not affected by Gleason, age, or therapy subtypes. We propose that this panel could be utilized to prioritize patients who would benefit from ADT and patients at risk of resistance that should be offered an alternative regimen. Such approach holds a long-term objective to build an adaptable accurate platform for precision therapeutics.

1. Introduction

Prostate cancer is the most common malignancy and one of the leading causes of death in American men [1–3]. Since prostate cancer initiation and progression depend on androgens [4–6], androgen-deprivation has been the mainstay of treatment for patients with advanced disease. Even though majority of patients initially respond to androgen-deprivation therapy (ADT), remission lasts 2–3 years on average, with eventual relapse and progression to castration-resistant disease, which is nearly always metastatic and lethal [7,8]. Resistance to ADT and the paucity of the therapeutic options for patients with castration-resistant disease are among major clinical challenges in prostate cancer management [9–11].

While multifaceted and heterogeneous, prostate cancer is characterized by the scarcity of genomic mutations [12] and absence of well-defined subtypes [13–15], thus making therapeutic management challenging and suggesting that more complex mechanisms (e.g., interplay of epigenomic and genomic mechanisms) might play a role in treatment response. In the last decade, epigenomics has been at the center of scientific interest, including recognition of its role in cancer initiation and progression [16–20]. In recent years, one of the most commonly observed epigenomic means, chromatin accessibility (i.e., DNA methylation), has received significant attention due to its role in cell development [21], genomic imprinting (i.e., biological process through which a gene carries information about its ancestor) [22], aging [23] and carcinogenesis [24,25]. DNA methylation (Fig. 1) is defined by an addition of methyl group to the fifth position of cytosine (converting it to 5-methylcytosine). In mammals, methylation of cytosine often occurs in regions where cytosine is followed by guanine (connected through phosphate molecule), named a CpG site [26,27]. A DNA region with frequent occurrences of CpG sites is commonly known as a CpG island or CGI [27,28]. Interestingly, 70% of gene promoter regions are associated with the CGIs, which can alter gene regulation [26]. In fact, if CGI within the promoter region is methylated, it becomes occupied by the Methylated DNA Binding Protein (MDBP) [29], which competes with transcription factor binding. MDBP can act as a transcription repressor or enhancer [30,31], depending on the transcription process it interferes with. In cancer, importance of the CGIs was initially observed among retinoblastoma patients, where CGI hyper-methylation led to silencing of *Rb* gene [32]; since then numerous groups have demonstrated the significant role of DNA methylation in oncogenesis [25,33–35].

In recent years, studies started to link aberrant level of DNA methylation to cellular transformation and clonal expansion [36,37], often implicated in therapeutic response and resistance. For instance, hyper-methylation of *MLH1* has been shown to be associated with increased resistance to cisplatin in ovarian cancer [38]; hyper-methylation of *HOXC10* has been found to influence resistance to anti-estrogen therapy

* Corresponding author at: Rutgers, The State University of New Jersey, Rutgers School of Health Professions, 65 Bergen Street, Rm 923B Newark, NJ 07101, USA.
E-mail address: amitrofa@shp.rutgers.edu. (A. Mitrofanova).

https://doi.org/10.1016/j.ebiom.2018.04.007

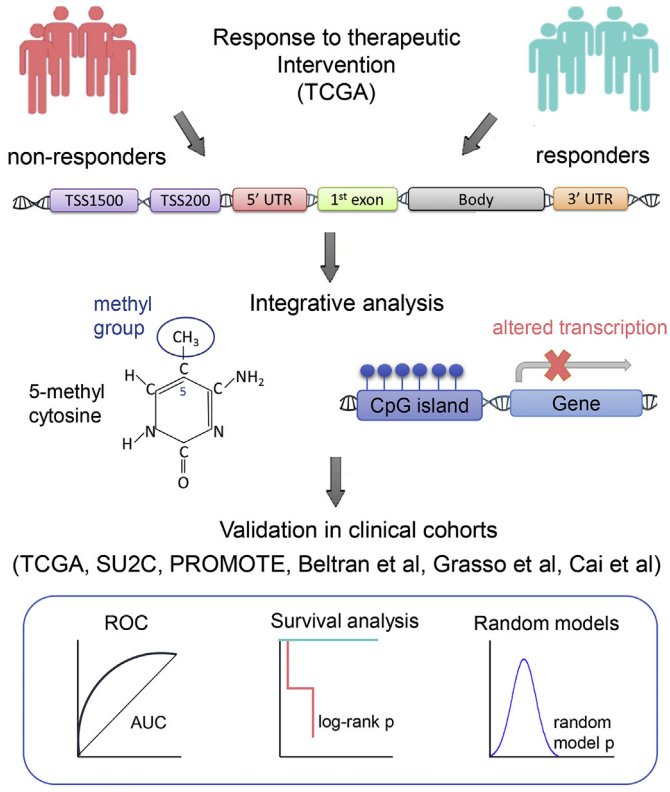

Fig. 1. Schematic representation of the systematic integrative approach. (Top) Non-responder and responder groups are compared for differentially methylated events/sites. (Middle) Differential methylation is integrated with expression of site-harboring genes. (Bottom) Candidate site-gene panel is evaluated for clinical significance.

in ER+ breast cancer [39]; hypo-methylation of *ABCB1* had been associated to paclitaxel-resistant ovarian cancer [40], etc. Further, recent studies have demonstrated that integrative analysis is crucial for in-depth understanding of molecular mechanisms involved in therapeutic response, for example *(i)* correlation between DNA methylation and mRNA expression of *FHIT* has been suggested as a marker for risk management in non-small cell lung and breast cancer [41]; *(ii)* aberrant frequencies of genes correlated between DNA methylation (as well as copy number variation) and expression levels could identify molecular subtypes in hepatocellular carcinoma patients [42]; *(iii)* correlation between DNA methylation and gene expression defined transcriptional patterns in molecular subtypes of breast cancer [43], etc. Thus, a systematic investigation of the effect of DNA methylation on therapeutic response and analysis of its functional effect on the expression of the harboring genes might enhance our understanding of the mechanisms implicated in resistance and provide valuable predictive markers of predisposition to therapeutic failure.

In this study, we have developed a systematic genome-wide integrative approach to analyze DNA methylation and its causal effect on mRNA gene expression to predict response to therapeutic intervention in cancer patients (see schematics in Fig. 1). We have named this approach Epi2GenR - Epigenomic and Genomic mechanisms of treatment Resistance. We have compared (*epi*) genomic profiles from primary tumors of prostate cancer patients with poor (i.e., non-responders) and favorable (i.e., responders) response to androgen-deprivation therapy and identified a panel of 5 differentially methylated sites, whose methylation changes explain expression variation in their site-harboring genes. We further tested the ability of this panel to predict therapeutic response in non-overlapping independent patient cohorts. In fact, the 5 site-gene panel was able to differentiate patients with predisposition to ADT failure from patients with favorable treatment response in *TCGA-PRAD* [13] (log-rank p = 0.0191, hazard ratio = 4.37) and other [44–48]patient cohorts (sensitivity = 100%, AUROC = 0.83, AUROC = 0.98). We have confirmed significant non-random predictive ability of the identified 5 site-gene panel and its robustness to increased false positive (FP) and false negative (FN) rates through random modeling and robustness analysis, respectively. Furthermore, we have demonstrated that the ability of this panel to predict therapeutic response does not depend on commonly used prognostic variables, such as pathological and clinical T-stage, Gleason score (i.e., pathology-based grading system of prostate tissues), age, and therapy subtype. We propose that this panel can potentially be used to pre-screen patients to prioritize those who would benefit from ADT and patients at risk of developing resistance. Our method holds a long-term potential to improve therapeutic management of cancer patients and builds a foundation for personalized therapeutic advice for patients with advanced malignancies.

2. Materials and Methods

2.1. DNA Methylation and mRNA Expression Resources

Prostate cancer patient cohorts utilized in this study come from the publicly available data sources, including *The Cancer Genome Atlas - Prostate Adenocarcinoma (TCGA-PRAD)*, *Stand up to Cancer (SU2C)*, *Grasso* et al. (GSE35988), *Cai* et al. (GSE32269), *Sboner* et al. (GSE16560), *Beltran* et al., and *Prostate Cancer Medically Optimized Genome-Enhanced Therapy (PROMOTE)* datasets (Table 1): *(i) TCGA-PRAD* [13] cohort was downloaded from Genomics Data Commons (GDC, https://gdc.nci.nih.gov/) on November 15, 2016. Information about type and time of treatment was obtained and synthesized from the clinical, follow-up, and the treatment data files, obtained from the TCGA GDC legacy archive (https://portal.gdc.cancer.gov/legacy-archive). For the purpose of this study we selected patients with primary tumors (obtained after radical prostatectomy), which were treated with adjuvant androgen deprivation therapy (ADT) and further monitored for disease progression (n = 66), which were suited to study primary ADT resistance. *TCGA-PRAD* DNA methylation was profiled on Illumina Infinium Human Methylation (HM450) array and RNA-seq was profiled on Illumina HiSeq 2000; *(ii) Stand up to Cancer (SU2C)* [48] contained tumors from metastatic castration-resistant prostate cancer (CRPC, n = 51, raw sequencing data for 51 patients were available for download from dbGaP phs000915.v1.p1) obtained from bone or soft tissue biopsies, profiled on Illumina HiSeq 2500 platform; *(iii) Grasso* et al. [46] dataset was obtained from Gene Expression Omnibus (GEO) GSE35988 and contained prostatectomy samples of primary tumors from patients with hormone-naïve prostate cancer (n = 58) and metastatic CRPC samples at rapid autopsy (n = 33), profiled on Agilent-014850 Whole Human Genome Microarray 4x44K G4112F; *(iv) Cai* et al. [45] dataset was obtained from GEO GSE32269 and contained primary tumors from patients with hormone-naïve prostate cancer isolated by laser capture microdissection (LCM) from frozen biopsies (n = 21) and CRPC bone metastasis obtained through CT guided bone marrow biopsies (n = 19), profiled on Affymetrix Human Genome U133A array; *(v) Beltran* et al. [44]: data were downloaded from dbGaP (phs000909.v1.p1) and contained tumors from metastatic castration-resistant prostate cancer (CRPC, neuroendocrine samples excluded, n = 34) obtained from lung, soft tissue and spinal cord biopsies, profiled on Illumina Genome Analyzer II; *(vi)* Prostate Cancer Medically Optimized Genome-Enhanced Therapy (PROMOTE) [47]: data were downloaded from dbGaP (phs001141.v1.p1) and contained tumors from metastatic CRPC that received 12 weeks of Abiraterone acetate and failed this treatment (n = 29), obtained from tissue biopsies and profiled on Illumina HiSeq 2500 platform; and as a negative control, we utilized *(vii) Sboner* et al. [49] dataset obtained from GEO GSE16560, which consisted of not treated primary tumors obtained from transurethral resection of the prostate (TURP) (n = 281) and profiled on 6 k

Please cite this article as: Panja, S., et al., Integrative (*epi*) Genomic Analysis to Predict Response to Androgen-Deprivation Therapy in Prostate Cancer, EBioMedicine (2018), https://doi.org/10.1016/j.ebiom.2018.04.007