

RESEARCH

Open Access

# Genome-wide methylation patterns predict clinical benefit of immunotherapy in lung cancer



Jeong Yeon Kim<sup>1</sup>, Jung Kyoong Choi<sup>1,2\*</sup> and Hyunchul Jung<sup>1,3\*</sup>

## Abstract

**Background:** It is crucial to unravel molecular determinants of responses to immune checkpoint blockade (ICB) therapy because only a small subset of advanced non-small cell lung cancer (NSCLC) patients responds to ICB therapy. Previous studies were concentrated on genomic and transcriptomic markers (e.g., mutation burden and immune gene expression). However, these markers are not sufficient to accurately predict a response to ICB therapy.

**Results:** Here, we analyzed DNA methylomes of 141 advanced NSCLC samples subjected to ICB therapy (i.e., anti-programmed death-1) from two independent cohorts (60 and 81 patients from our and IDIBELL cohorts). Integrative analysis of patients with matched transcriptome data in our cohort ( $n = 28$ ) at pathway level revealed significant overlaps between promoter hypermethylation and transcriptional repression in nonresponders relative to responders. Fifteen immune-related pathways, including interferon signaling, were identified to be enriched for both hypermethylation and repression. We built a reliable prognostic risk model based on eight genes using LASSO model and successfully validated the model in independent cohorts. Furthermore, we found 30 survival-associated molecular interaction networks, in which two or three hypermethylated genes showed significant mutual exclusion across nonresponders.

**Conclusions:** Our study demonstrates that methylation patterns can provide insight into molecular determinants underlying the clinical benefit of ICB therapy.

**Keywords:** Immunotherapy, Lung cancer, Methylation

## Introduction

Over the past several years, the number of non-small-cell lung cancer (NSCLC) patients treated with immune checkpoint blockade (ICB) therapy has increased at a fast rate due to its proven efficacy in treating NSCLC [1]. In particular, ICB therapy targeting programmed death-1 (PD-1) and programmed death-ligand 1 (PD-L1) has demonstrated effectiveness in dramatically improving NSCLC patient survival [2]. However, ICB therapy has varying degrees of effect in

patients. Previous studies showed that only a small subset of NSCLC (< 20%) patients can benefit from these novel agents [3, 4]. Thus, it is crucial to unravel the molecular determinants of the response to ICB therapy. To select patients who are likely to respond to this therapy, many genomic and transcriptomic biomarkers have been proposed, such as tumor mutational burden (TMB) [5] and the expression of key genes (e.g., *PD-L1*) [6, 7]. However, these biomarkers are not sufficient to accurately predict the response to ICB therapy. Beyond genomic and transcriptomic biomarkers, epigenetic aberrations have been reported to be associated with response to ICB therapy [8].

\* Correspondence: [jungkyoon@gmail.com](mailto:jungkyoon@gmail.com); [hastj7373@gmail.com](mailto:hastj7373@gmail.com)

<sup>1</sup>Department of Bio and Brain Engineering, KAIST, Daejeon 34141, Republic of Korea

Full list of author information is available at the end of the article



© The Author(s). 2020 **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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Aberrant DNA methylation has been recognized as a crucial factor in lung carcinogenesis [9]. In particular, the inhibition of tumor suppressor expression by promoter hypermethylation is a common event in NSCLC [10]. Thus, DNA methylation-based biomarkers have been extensively studied for predicting prognosis and response to conventional therapy [11, 12]. Silenced immune-related genes by promoter hypermethylation have been recently found to be predictive of ICB therapy response [13–15]. Given that the loss of functions in immune-related genes by genomic mutations (e.g., truncating mutations and copy number deletion) account for primary and acquired resistance to ICB therapy [16–18], we hypothesize that the silencing of immune genes through promoter hypermethylation could be one of the key mechanisms resulting in immune evasion. To comprehensively explore DNA methylation aberrations associated with the response to ICB therapy, we systematically analyzed DNA methylation markers predictive of ICB benefit by investigating genome-wide methylation data (Illumina 850K/EPIC platform) from a total of 141 NSCLC patients receiving ICB therapy. We discovered diverse predictive DNA methylation markers by building a lasso regression prediction model and investigating the molecular network of mutually exclusive promoter-hypermethylated genes. As a result, we suggest that DNA methylation is a promising candidate for biomarkers predicting the clinical response of ICB therapy.

## Results

### Differential DNA methylation pattern between nonresponders and responders

To comprehensively assess the genomic characteristics of our ICB therapy cohort ( $n = 60$ ; Supplementary Table 1), we performed genome-wide DNA methylation (850K Infinium Methylation EPIC Array) and exome-seq assays. First, we evaluated the association of patient survival and known predictive genomic biomarkers, such as TMB, neo-antigen load, aneuploidy level, and PD-L1 expression (Supplementary Figure 1). Remarkably, our cohort's clinical benefit was not predictable with these biomarkers; for example, our cohort did not show a correlation between mutation burden and progression-free survival (PFS). This independent pattern was preserved even after excluding a patient with an outlier TMB. Moreover, a high aneuploidy level was not significantly correlated with shorter PFS. We also examined known mutation signatures in NSCLC with respect to patient's PFS because APOBEC mutational signatures are reported to be enriched in NSCLC patients with durable clinical benefit [19, 20]. None of the mutation signatures were related with patient survival in our cohort.

Unlike these known biomarkers, the DNA methylation level illustrated its power to distinguish clinical response

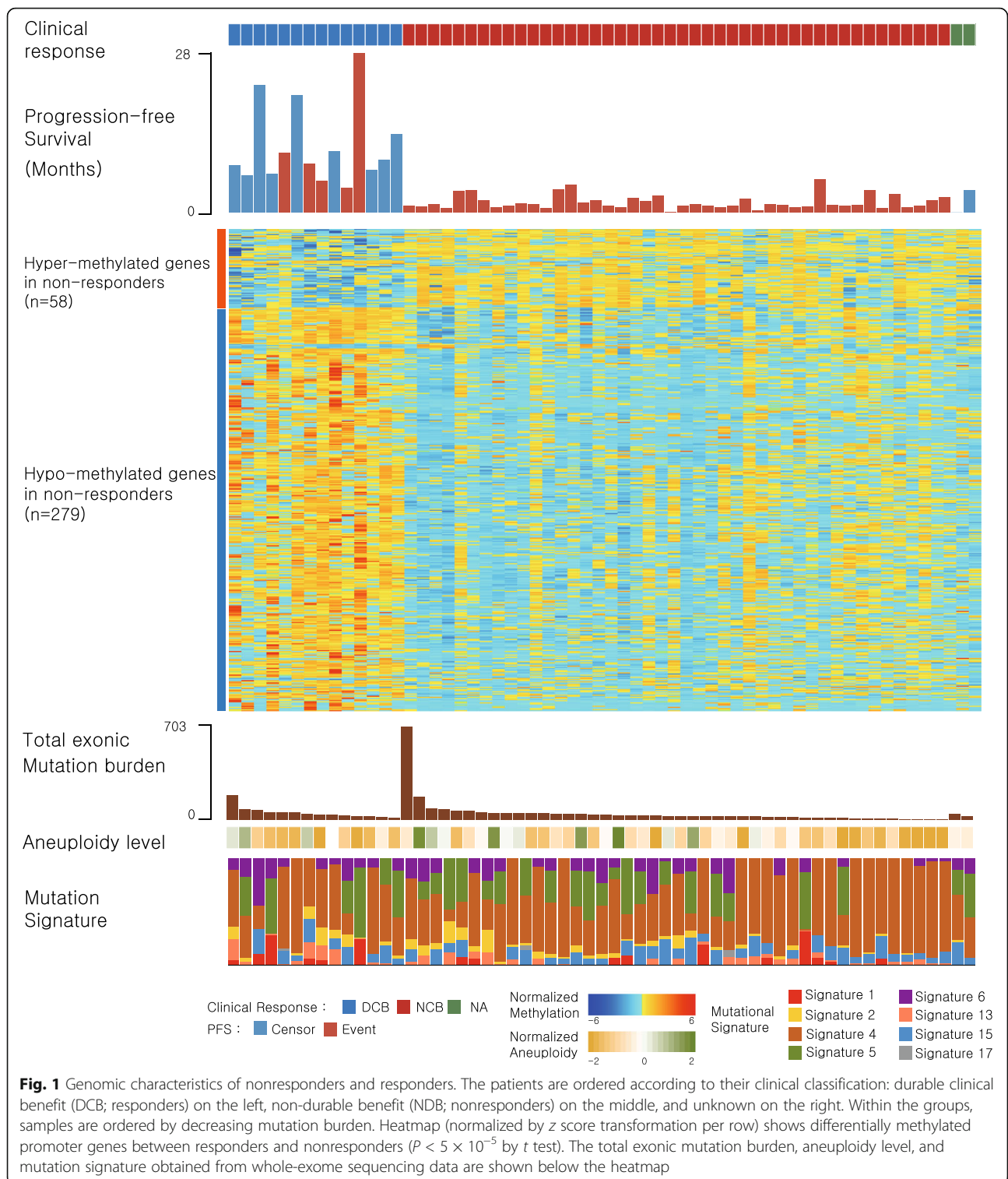
(Fig. 1). We found 65 and 377 differentially methylated regions (DMRs) and probes, respectively, between responders and nonresponders. A majority of DMRs (97%;  $n = 63$ ) is hypermethylated in nonresponders, and they are concentrated in promoter regions (Supplementary Figure 2), suggesting the potential of a strong local hypermethylation pattern regulating biological functions. On the contrary, the hypermethylated probes ( $n = 64$ ) in nonresponders tend to reside more in gene body regions whereas the hypomethylated probes ( $n = 313$ ) are concentrated in promoter regions. To analyze differential methylation patterns at the gene level, we first assigned the most differently methylated promoter probes to their corresponding genes (see “Methods” section). As a result, we found 337 differentially methylated genes, of which 58 and 279 genes are hyper- and hypomethylated in nonresponders, respectively. Interestingly, the hypermethylated genes in nonresponders were significantly enriched in the positive regulation of interferon-gamma secretion and the positive regulation of apoptotic process, whereas the hypomethylated genes were enriched in transcription regulation and DNA-directed RNA polymerase II activity (Supplementary Table 2).

### Promoter hypermethylation leads to the transcriptional silencing of immune pathways

DNA methylation in promoters is a well-established epigenetic mechanism for the inhibition of gene transcription [21]. To examine the extent to which promoter methylation affects the transcriptional silencing of immune genes, we further analyzed the RNA-seq data on a subset of our cohort ( $n = 28$ ). We then examined promoter hypermethylated and under-expressed pathways between nonresponders and responders using Gene Set Enrichment Analysis (GSEA) [22]. Interestingly, nonresponders had both enriched promoter hypermethylated ( $n = 24$ , 13.3%; Supplementary Table 3) and under-expressed immune pathways ( $n = 71$ , 39.2%; Supplementary Table 4), whereas responders did not have any significantly enriched immune pathways, suggesting the down-regulation of immune system in nonresponders (Fig. 2a). We further examined relationship between promoter hypermethylated and under-expressed pathways in nonresponders and found a statistically significant overlap between them (Fig. 2b,  $P = 3.4 \times 10^{-3}$ ). Representative GSEA plots of the overlapped pathway, namely interferon signaling, are illustrated in Fig. 2c. These results highlight the importance of the transcriptional silencing of immune pathways by promoter-hypermethylation in nonresponders.

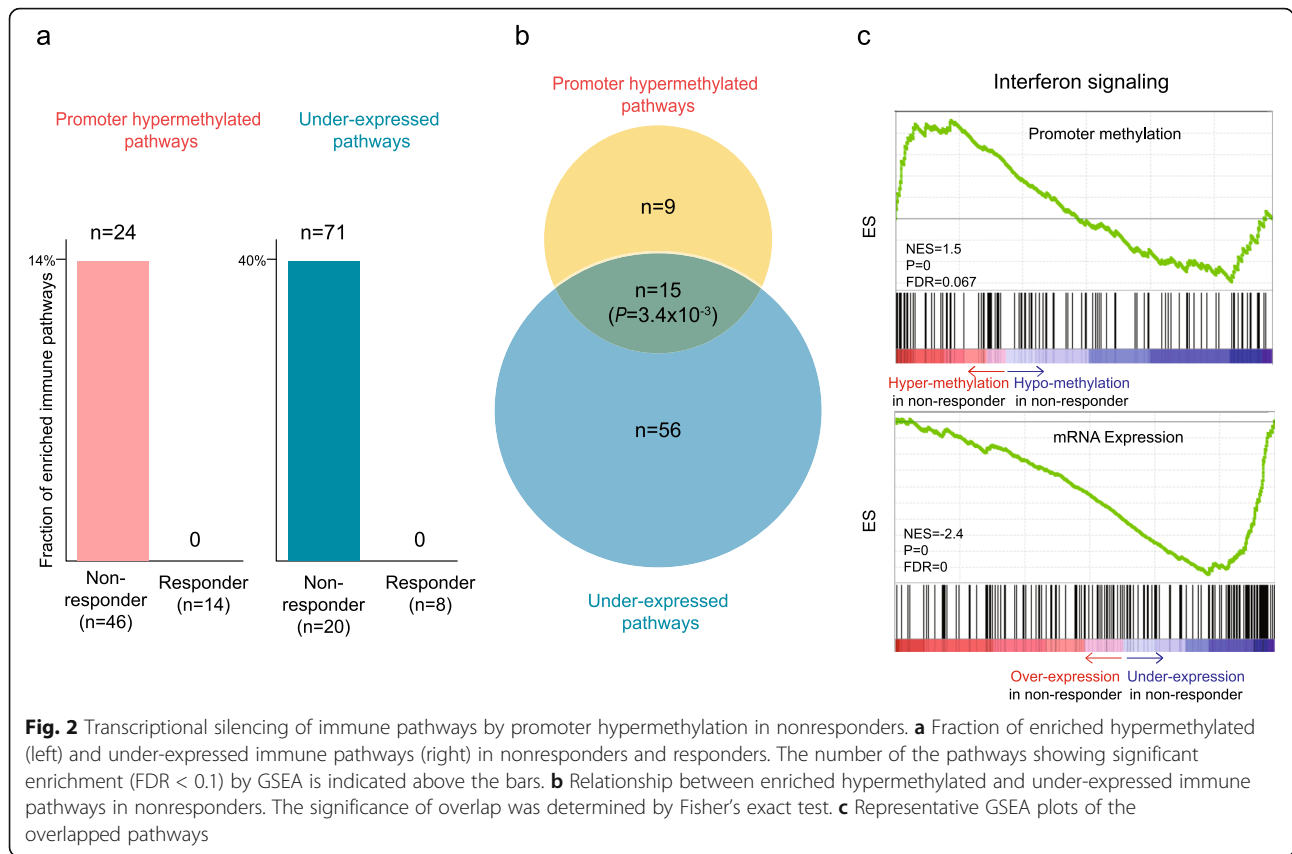
### Establishing a methylation-based prognostic prediction model

Next, we built a generalized linear Cox model for the accurate prediction of survival using the methylation



profile. Using our cohort as a training set ( $n = 60$ ), we built a survival model based on eight genes (Table 1, Fig. 3a, and Supplementary Figure 3). The coefficients of our model contribute to the risk score where negative and positive coefficients in the Cox

regression model signify factors that contribute to favorable and unfavorable survival, respectively. In other words, high methylation level of genes with a negative coefficient reduces risk factors. For example, *LTBR*, a gene with a negative coefficient in our



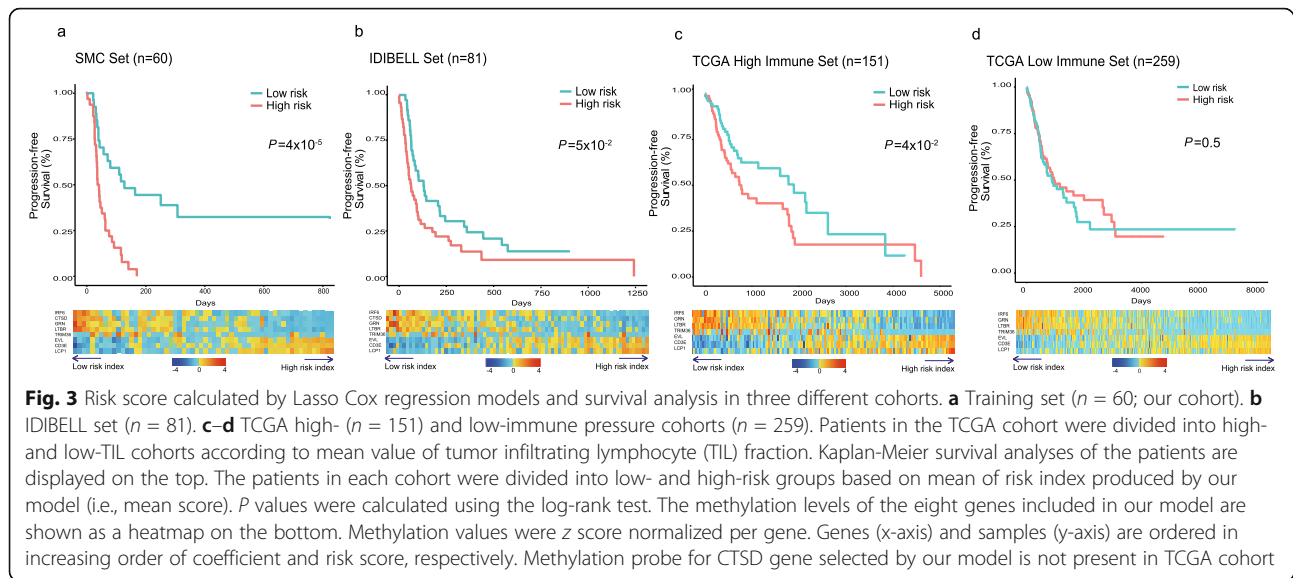
model, is associated with tumor necrosis and its activation is linked to carcinogenesis [23]. Thus, the promoter hypermethylation-derived transcription silencing of this gene would be beneficial for PFS. In contrast, a high methylation level of genes with a positive coefficient increases the risk score. For example, *CD3E* is involved in T cell development and activation [24]. The hypermethylation of this gene

will result in the negative regulation of T cell-mediated cytotoxicity, which can be expected to be unfavorable for survival.

We validated our model using two independent cohorts including TCGA NSCLC data (Fig. 3b–d). Although the TCGA samples in general are not appropriate to evaluate immunotherapy markers, some tumors under strong immune pressure can serve as

**Table 1** Genes used in methylation-based prognostic prediction model

Probe	Coefficient	Gene symbol	Gene function	Cancer relatedness
cg22029157	- 1.52	IRF6	Interferon regulatory factor associated with cytokine signaling	Tumor suppressor activity
cg12007048	- 1.096	CTSD	Proteolytic activation of hormones and growth factors (i.e., EGFR)	Linked with poor prognosis in NSCLC
cg07935119	- 1.04	GRN	Granulin coding gene; growth factor involved in inflammation and cell proliferation	Regulate tumorigenesis; immune evasion, proliferation
cg23079808	- 0.989	LTBR	Tumor necrosis factor receptor; signaling immune response and programmed cell death	Activation linked to carcinogenesis
cg04450862	- 0.907	TRIM36	Mediate ubiquitination and proteasomal degradation of target protein	Known to be upregulated in cancer
cg19918549	0.347	EVL	Actin-associated proteins involved in axon guidance	Upregulated in breast cancer
cg24612198	0.385	CD3E	Part of CD3 complex that facilitate T cell development	Down regulation of this gene results in T cell apoptosis
cg17771150	0.969	LCP1	Actin-binding protein that are involved in T cell activation	Upregulated in cancer



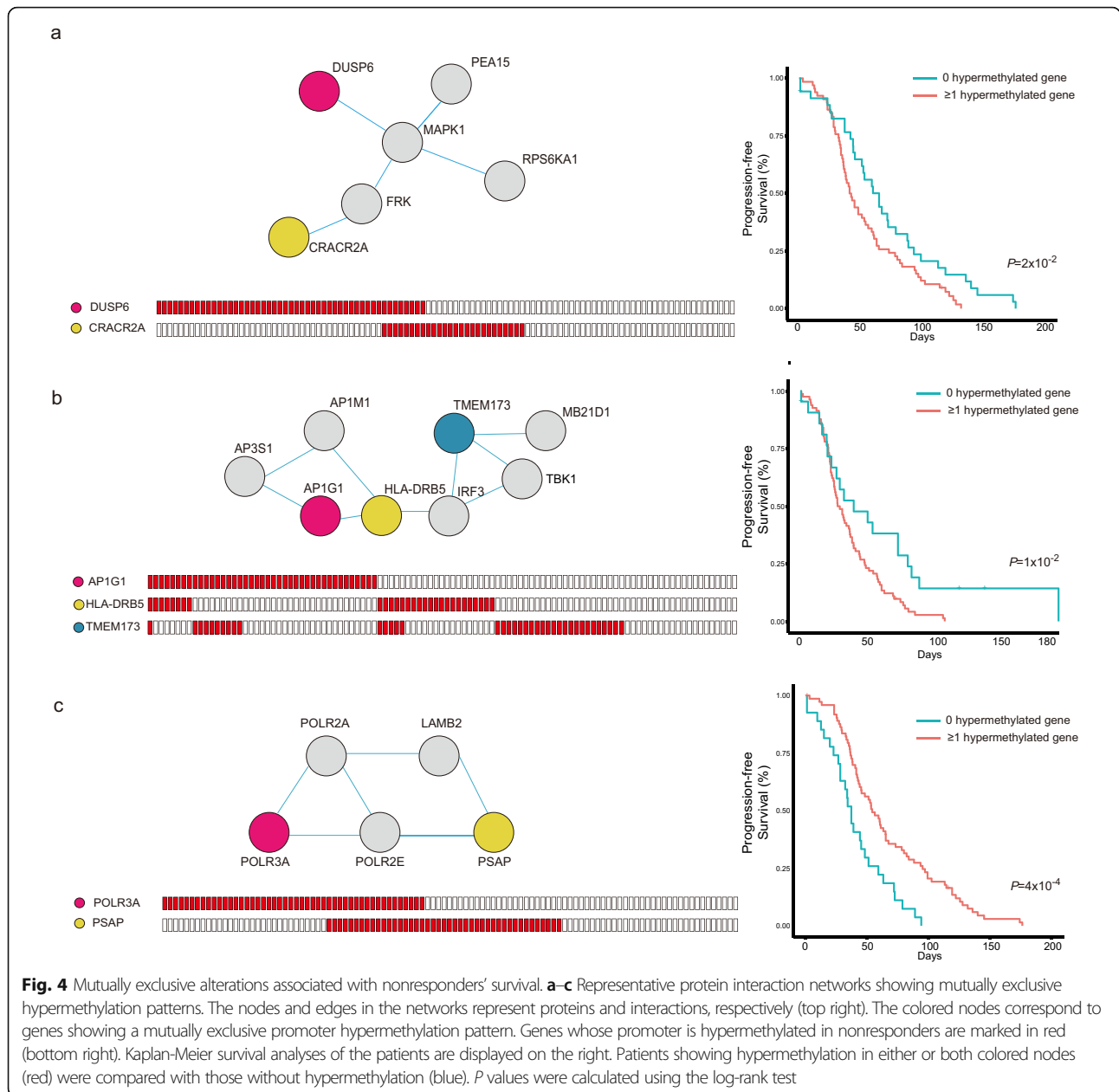
surrogates for samples treated with checkpoint blockade. Therefore, we selected samples with high immune cell infiltration using tumor-infiltrating lymphocyte (TIL) fraction and tested whether our methylation markers are associated with the survival of these samples [25]. As a result, our model demonstrates its predictive power to accurately distinguish survival outcomes in immunotherapy and TCGA high-TIL cohorts (Fig. 3a–c) but not in TCGA low-TIL cohort (Fig. 3d). For patients with a low risk score and better survival, genes with a negative coefficient had a high level of methylation and those with positive coefficient have low level of methylation (heatmaps in Fig. 3). This trend is opposite for patients with high risk scores. We also performed receiver operating characteristic (ROC) curve analysis of the risk score for predicting progression-free survival (Supplementary Figure 4). The area under curve (AUC) values at 6 and 12 months are consistently higher in immunotherapy (SMC and IDIBELL) and TCGA high-TIL cohorts than TCGA low-TIL cohort. The AUC values in immunotherapy and TCGA high-TIL cohorts ranged from 0.60 to 0.87, while the values in TCGA low-TIL cohort were close to 0.5, indicating no predictive effect of our risk score in this low-immune pressure cohort (i.e., random performance). Collectively, our results reinforce the role of the risk score as a potential biomarker of response to immunotherapy. Interestingly, each gene used in our model can also distinguish survival outcomes in immunotherapy cohorts (Supplementary Figure 5). Consistent with our regression model, genes with a negative coefficient show an inverse relationship between methylation and survival whereas those with a positive coefficient show a direct association. Thus, differential patterns of these genes contribute to the accurate classification of patients' survival outcome prediction.

### Mutually exclusive hypermethylated genes are associated with patient survival

Previous studies have revealed mutually exclusive patterns of genomic alterations across cancer patients, including mutations in driver genes [26]. Genes exhibiting mutually exclusive alteration pattern are generally involved in a common biological pathway due to their functional redundancy [27, 28]. Knowledge about these patterns can provide important insight into novel cancerous network and potential therapeutic targets [29]. To identify such mutually exclusive patterns associated with response to ICB therapy at the methylation level, we first integrated two different cohorts receiving immunotherapy (i.e., the IDIBELL cohort and SMC cohort;  $n = 141$ ) and identified genes for which promoter hypermethylation is specifically observed in nonresponders ( $n = 101$ ; see “Methods” section). We then searched for protein interaction networks whose member genes show mutually exclusive alteration patterns. We identified 474 protein interaction networks in which two or three hypermethylated genes showed significantly mutually exclusive patterns across nonresponders. Importantly, we further conducted survival analysis to assess the ultimate effect of exclusivity and identified 30 protein networks associated with progression-free survival (Supplementary Table 5).

Representative networks associated with patient survival are shown in Fig. 4. As shown in Fig. 4a, *DUSP6* and *CRACR2A* showed mutually exclusive patterns ( $P = 3.3 \times 10^{-6}$ ). In other words, there is a significantly low portion of samples that harbor promoter hypermethylation in both genes. Furthermore, we performed survival analysis by comparing patients with promoter hypermethylation in either or both genes with those without hypermethylation, resulting in an unfavorable prognosis





of patients with hypermethylated genes. *DUSP6* is a tumor suppressor gene known to negatively regulate MAPK pathways [30]. Furthermore, recent reports highlight its role in differentiation and apoptosis in cancer. *CRACR2A* is part of the calcium channel in T cells and is pivotal in T cell cytotoxic activity [31]. This result highlights that the disruption of the network caused by the silencing of either gene confers unfavorable prognosis in NSCLC patients.

This pattern can also be found in the network with three mutually exclusive genes (Fig. 4b; AP1G1-TMEM173:  $P = 1.06 \times 10^{-5}$ , AP1G1-HLA-DRB5:  $P = 3.4 \times 10^{-5}$ , TMEM173-HLA-DRB5:  $P = 1.98 \times 10^{-5}$ ).

*TMEM173* is known to be associated with the apoptotic signaling pathway related to MHC type II, in which *HLA-DRB5* is a subunit [32]. Liu et al. reported that the transcriptional silencing of *APIG1* can activate the PIK3/AKT pathway, which induces tumor proliferation and invasion [33]. Thus, the promoter hypermethylation of any of the three genes can lead to reduced survival by disrupting the network.

In contrast, the representative network in Fig. 4c shows the opposite survival trend; samples with hypermethylation in either gene have a longer PFS. This finding can also be explained by the gene function of the two mutually exclusive genes, *POLR3A* and *PSAP* ( $P =$

$1.3 \times 10^{-5}$ ). *POLR3A* is a RNA polymerase III subunit and is known as an unfavorable marker in cancer [34]. *PSAP* produces prosaposin, whose amplification is reported to cause carcinogenesis or tumor progression [35]. Thus, the disruption of this network by transcriptional silencing can be expected to yield better survival ( $P = 4 \times 10^{-4}$ ).

## Discussion

DNA methylation aberrations are closely related to cancer since they regulate spatiotemporal gene expression in cell type specific manner. In particular, the downregulation of tumor suppressors by promoter hypermethylation is a widely accepted mechanism in cancer progression [36, 37]. Due to its distinct pattern in cancer, DNA methylation has been used as a diagnostic and predictive biomarker. For example, methylation profiles obtained from blood liquid biopsy have the potential to detect early stage colorectal cancer (CRC) with high efficacy [38]. In addition, a five-CpG-based classifier improved risk stratification for patients with clear cell renal cell carcinomas (ccRCCs) [39]. In this study, we showed that DNA methylation alterations are also a promising biomarker for predicting the clinical response of ICB therapy. We first characterized methylation differences between responders and nonresponders on a genome-wide scale and demonstrated that gene promoter hypermethylation significantly correlates with the mRNA expression silencing of immune-related pathways in nonresponders. We built and validated an eight-gene based lasso regression prediction model to accurately classify patient survival outcomes. The established model can be used to stratify patients who are likely to benefit from ICB therapy. Furthermore, through mutually exclusive hypermethylation pattern analysis, we found protein interaction networks in which exclusive disruption leads to different survival outcomes. The wide range of genes identified in this study has broaden knowledge about potential immunotherapy resistance mechanism. More investigation is warranted in order to understand the mechanisms underlying the associations between genes.

A growing body of evidence suggests that DNA methyltransferase inhibitors (DMTIs), such as 5-azacytidine (AZA) and 5-aza-2'-deoxycytidine (DAC), have potential to improve the therapeutic efficacy of ICB [40, 41]. DMTI treatment can promote anti-tumor immunity by upregulating immune-related genes [42] or endogenous retrovirus (ERV) [43]. Our results warrant future research on the effectiveness of the combination of ICB therapy with DNMTi because a significant proportion of immunomodulatory pathways might be potentially downregulated in nonresponders by promoter hypermethylation. The identified methylome-based markers

could be further used to guide the selection of patients for this combined therapy in clinical trials.

## Conclusions

Our study demonstrates that methylation patterns can provide insight into molecular determinants underlying the clinical benefit of ICB therapy.

## Methods

### Data collection

Advanced NSCLC patients who received anti-PD-1/PD-L1 therapy ( $n = 60$ ) were registered for this study at Samsung Medical Center. This study was approved by Samsung Medical Center's institutional review board (2018-03-130 and 2013-10-112). The clinical benefit was determined according to RECIST v1.1 [44]. If the patient experienced a partial response (PR) or stable disease (SD) of more than 6 months, he or she was categorized as responder receiving durable clinical benefit (DCB). Conversely, if a patient experienced progressive disease (PD) or SD of less than 6 months, he or she was categorized as non-responder with non-durable benefit (NDB). Of the 60 patients, 14 were responders (DCB), 44 were nonresponders (NDB), and 2 were not determined (NA). The processing of DNA methylation (850K Infinium Methylation EPIC Array), exome- and RNA-seq ( $n = 27$ ) data was illustrated in our previous study [45]. Neo-antigen load was calculated using DeepNeo [46]. Aneuploidy level was retrieved from our previous research [45].

Methylation profiles (850K Infinium Methylation EPIC Array) and matching clinical information of 81 patients with stage IV NSCLC treated with anti-PD-1 medication were retrieved from the research performed by the Esteller et al. group (IDIBELL cohort) [47]. The methylation (Illumina Infinium HumanMethylation 450K) and the clinical profiles of 479 lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) patients from The Cancer Genome Atlas (TCGA) were obtained from the publication page (<https://gdc.cancer.gov/about-data/publications/pancanatlas>).

### Identification of differentially methylated/expressed pathways between nonresponders and responders

We first assigned the most differently methylated promoter probes between responders and nonresponders into a corresponding gene. Promoter probes were defined as those that reside in either the transcription start site (TSS), 1<sup>st</sup> exon, or 5'UTR region. Ranked genes by the t-static (methylation level difference between nonresponders and responders) were used for input into the preranked module of the GSEA software with Reactome pathways (provided at <https://reactome.org/>) belonging to the Immune System category ( $n = 181$ ). We also ranked genes by mRNA expression difference between

nonresponders and responders and then used the same procedure to identify differentially expressed pathways. DMRs were calculated as regions with differential methylation patterns with three minimum probes and 500 maximum gap using the ChAMP package [48].

### Building a prediction model of clinical benefit

We fitted the Cox model based on elastic-net to predict clinical benefit using the glmnet R package [49]. For efficient computation, we narrowed down the methylation probes with two criteria. First, we selected probes related to known immune genes curated by Reactome version 69. Second, probes showing significant differences ( $P < 5 \times 10^{-4}$  by  $t$  test) between responders and nonresponders were selected. After filtering, 18 probes were given as input to the lasso-regularized Cox proportional hazards model, resulting in eight probes in different genes in our final model (Table 1). For PSW-1711, whose survival information was not available, we treated this case with a modest standard as having no progression for 1 day.

The model's prediction power was validated on two independent data sets: IDIBELL and TCGA NSCLC cohorts. For 410 TCGA samples whose tumor-infiltrating lymphocyte (TIL) fraction is available, we divided them into high- ( $n = 151$ ) and low-immune pressure groups ( $n = 259$ ) according to mean value of TIL fraction [25]. The patients in each cohort were divided into low- and high-risk groups based on the risk index produced by our model (i.e., mean score). The risk index in the validation cohorts was calculated using the predict function of glmnet. A heatmap of eight genes (Fig. 4) was drawn using the heatmap.2 function in the gplots package [50].

### Mutually exclusive hypermethylation pattern analysis

A total of 141 samples from our cohort and IDIBELL cohort were combined using ComBat, a tool that adjusts batch effects when combining two different datasets [51]. In the combined dataset, 40 patients were responder, and 101 were nonresponders. We first identified hypermethylated genes in nonresponders for which the methylation level was greater than two standard deviations from the mean in responders. The data were further reduced by several factors. First, hypermethylated genes with more than 10 occurrences (approximately 10% of nonresponders) were selected to see general patterns among nonresponders. Second, immune related genes in the Reactome were selected to focus on immunologic patterns. Using WExT, a tool that searches for mutually exclusive patterns, we found 472 mutually exclusive gene sets comprised of two or three genes [52]. Then, nonresponders were divided into two groups: one group is comprised of samples with no hypermethylation in either gene and the other with one or more hypermethylated genes. Using the survival R package, we

performed survival analysis on each set and discovered 30 pairs with differential survival between samples with and without hypermethylation [53]. Networks of mutually exclusive genes were created using STRING and GeneMANIA, tools that predict functional interaction networks using multiple databases [54, 55].

### Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s13148-020-00907-4>.

**Additional file 1: Figure S1.** Survival plots of previously known biomarkers. (a) tumor mutation burden (b) neo-antigen load (c) aneuploidy level (d) PD-L1 expression. Mutation burden, neo-antigen load, and aneuploidy level were calculated with exome-sequencing data ( $n = 60$ ) and PD-L1 expression was obtained from RNA-sequencing data ( $n = 28$ ).

**Additional file 2: Figure S2.** Summary statistics of differentially hyper- (a), and hypomethylated (b) probes, and hypermethylated regions (c) in nonresponders according to their location. Differentially hypomethylated regions are now shown due to the lack of them.

**Additional file 3: Figure S3.** Tuning parameter log (lambda) selection during ten-fold cross validation (a) and L1-norm value (b) used for the regression model.

**Additional file 4: Figure S4.** ROC curve analysis of the risk score for predicting progression-free survival greater than six months (a) and one year (b).

**Additional file 5: Figure S5.** X-tile plots of the methylation patterns of the eight genes used in the regression model for a total of 141 samples across immunotherapy (SMC and IDIBELL) [56]. Each plot represents an association between individual gene methylation level and survival. The cut point, which separates high- and low-risk groups, represents the highest  $\chi^2$ -value obtained from the Kaplan-Meier survival analysis and log-rank test. Green and red indicate direct and inverse relationships, respectively, between methylation and survival.

**Additional file 6.** Supplementary tables

### Abbreviations

ICB: Immune checkpoint blockade; NSCLC: Non-small-cell lung cancer; DMR: Differentially methylated regions; TMB: Tumor mutational burden; PFS: Progression-free survival; DMTis: DNA methyltransferase inhibitors

### Acknowledgements

Not applicable

### Authors' contributions

J.Y.K. performed all data analyses. J.Y.K., J.K.C., and H.J. wrote the manuscript. J.K.C. and H.J. conceived and supervised the study. The authors read and approved the final manuscript.

### Funding

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation of Korea funded by the Ministry of Science and ICT (NRF-2017M3A9A7050612 and NRF-2019M3A9B6064688).

### Availability of data and materials

Our cohort dataset (methylation chip and RNA-seq data) is available at Gene Expression Omnibus under accession number GSE119144 and GSE135222. The raw data for the exome sequencing of our cohort is available at European Genome-phenome Archive under accession number EGAS00001003731.

### Ethics approval and consent to participate

We complied with all relevant ethical regulations for work with human participants. Informed consent was obtained. This study was approved by



the institutional review board at Samsung Medical Center (2018-03-130 and 2013-10-112).

#### Consent for publication

All authors give consent for the publication of the manuscript.

#### Competing interests

The authors declare no competing financial interests.

#### Author details

<sup>1</sup>Department of Bio and Brain Engineering, KAIST, Daejeon 34141, Republic of Korea. <sup>2</sup>Penta Medix Co., Ltd., Seongnam-siGyeonggi-do, 13449, Republic of Korea. <sup>3</sup>Cancer Ageing and Somatic Mutation Programme, Wellcome Sanger Institute, Cambridge, UK.

Received: 28 February 2020 Accepted: 20 July 2020

Published online: 06 August 2020

#### References

- Doroshow DB, Sanmamed MF, Hastings K, Politi K, Rimm DL, Chen L, et al. Immunotherapy in non-small cell lung cancer: facts and hopes. *Clin Cancer Res*. 2019;25:4592–602.
- Sui H, Ma N, Wang Y, Li H, Liu X, Su Y, et al. Anti-PD-1/PD-L1 therapy for non-small-cell lung cancer: toward personalized medicine and combination strategies. *J Immunol Res Hindawi*. 2018;2018.
- Herbst RS, Baas P, Kim DW, Felip E, Pérez-Gracia JL, Han JY, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet*. 2016;387:1540–50.
- Brahmer J, Reckamp KL, Baas P, Crinò L, Eberhardt WEE, Poddubskaya E, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N Engl J Med*. 2015;373:123–35.
- Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol*. 2019;30:44–56.
- Davis AA, Patel VG. The role of PD-L1 expression as a predictive biomarker: an analysis of all US food and drug administration (FDA) approvals of immune checkpoint inhibitors. *J Immunother Cancer Journal for Immunotherapy of Cancer*. 2019;7:1–8.
- Ørskov AD, Treppendahl MB, Skovbo A, Holm MS, Friis LS, Hokland M, et al. Hypomethylation and up-regulation of PD-1 in T cells by azacytidine in MDS/AML patients: a rationale for combined targeting of PD-1 and DNA methylation. *Oncotarget*. 2015;6:9612–26.
- Loo Yau H, Ettayebi I, De Carvalho DD. The cancer epigenome: exploiting its vulnerabilities for immunotherapy. *Trends Cell Biol [Internet]. Elsevier Ltd*; 2019;29:31–43. Available from: <https://doi.org/10.1016/j.tcb.2018.07.006>.
- Balgkouranidou I, Liloglou T, Lianidou ES. Lung cancer epigenetics: emerging biomarkers. *Biomark Med England*. 2013;7:49–58.
- Lu F, Zhang H-T. DNA methylation and non-small cell lung cancer. *Anat Rec (Hoboken)*. United States; 2011;294:1787–1795.
- Duruiseaux M, Esteller M. Lung cancer epigenetics: from knowledge to applications. *Semin Cancer Biol [Internet]. Elsevier*; 2018;51:116–28. Available from: <https://doi.org/10.1016/j.semcancer.2017.09.005>.
- Szejniuk WM, Robles AI, McCulloch T, Falkmer UGI, Røe OD. Epigenetic predictive biomarkers for response or outcome to platinum-based chemotherapy in non-small cell lung cancer, current state-of-art. *Pharmacogenomics J United States*. 2019;19:5–14.
- Goltz D, Gevensleben H, Vogt TJ, Dietrich J, Golletz C, Bootz F, et al. CTLA4 methylation predicts response to anti-PD-1 and anti-CTLA-4 immunotherapy in melanoma patients. *JCI insight*. 2018;3:0–9.
- Goltz D, Gevensleben H, Dietrich J, Dietrich D. PD-L1 (CD274) promoter methylation predicts survival in colorectal cancer patients. *Oncoimmunology*. 2017;6:e1257454.
- Fröhlich A, Loick S, Bawden EG, Fietz S, Dietrich J, Diekmann E, et al. Comprehensive analysis of tumor necrosis factor receptor TNFRSF9 (4-1BB) DNA methylation with regard to molecular and clinicopathological features, immune infiltrates, and response prediction to immunotherapy in melanoma. *EBioMedicine*. 2020;52:102647.
- Ye Z, Dong H, Li Y, Ma T, Huang H, Leong HS, et al. Prevalent homozygous deletions of type I interferon and defensin genes in human cancers associate with immunotherapy resistance. *Clin cancer Res an Off J Am Assoc Cancer Res*. 2018;24:3299–308.
- Shin DS, Zaretsky JM, Escuin-Ordinas H, Garcia-Diaz A, Hu-Lieskovan S, Kalbasi A, et al. Primary resistance to PD-1 blockade mediated by JAK1/2 mutations. *Cancer Discov [Internet]*. 2017;7:188 LP – 201. Available from: <http://cancerdiscovery.aacrjournals.org/content/7/2/188.abstract>.
- Zaretsky JM, Garcia-Diaz A, Shin DS, Escuin-Ordinas H, Hugo W, Hu-Lieskovan S, et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N Engl J Med [Internet]. Massachusetts Medical Society*; 2016;375:819–29. Available from: <https://doi.org/10.1056/NEJMoa1604958>.
- Lee J, Lee AJ, Lee J-K, Park J, Kwon Y, Park S, et al. Mutalisk: a web-based somatic MUTation AnalySiS toolkit for genomic, transcriptional and epigenomic signatures. *Nucleic Acids Res [Internet]*. 2018;46:W102–8. Available from: <https://doi.org/10.1093/nar/gky406>.
- Wang S, Jia M, He Z, Liu X-S. APOBEC3B and APOBEC mutational signature as potential predictive markers for immunotherapy response in non-small cell lung cancer. *Oncogene [Internet]*. 2018;37:3924–36. Available from: <https://doi.org/10.1038/s41388-018-0245-9>.
- Delpu Y, Cordelier P, Cho WC, Torrisani J. DNA methylation and cancer diagnosis. *Int J Mol Sci*. 2013;14:15029–58.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci*. 2005;102:15545–50.
- Lau TS, Chung TKH, Cheung TH, Chan LKY, Cheung LWH, Yim SF, et al. Cancer cell-derived lymphotoxin mediates reciprocal tumour-stromal interactions in human ovarian cancer by inducing CXCL11 in fibroblasts. *J Pathol*. 2014.
- DeJarnette JB, Sommers CL, Huang K, Woodside KJ, Emmons R, Katz K, et al. Specific requirement for CD3epsilon in T cell development. *Proc Natl Acad Sci U S A [Internet]. National Academy of Sciences*; 1998;95:14909–14. Available from: <http://www.jstor.org/stable/46643>.
- Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. *Immunity*. 2018;48:812–830.e14.
- Campbell PJ. Cliques and schisms of cancer genes. *Cancer cell*. 2017.
- Zhang H, Deng Y, Zhang Y, Ping Y, Zhao H, Pang L, et al. Cooperative genomic alteration network reveals molecular classification across 12 major cancer types. *Nucleic Acids Res*. 2017;45:567–82.
- Ding W, Feng G, Hu Y, Chen G, Shi T. Co-occurrence and mutual exclusivity analysis of DNA methylation reveals distinct subtypes in multiple cancers. *Front cell Dev Biol*. 2020;8:20.
- Szczurek E, Beerenwinkel N. Modeling mutual exclusivity of cancer mutations. *PLOS Comput Biol [Internet]. Public Library of Science*; 2014;10: e1003503. Available from: <https://doi.org/10.1371/journal.pcbi.1003503>.
- Ma J, Yu X, Guo L, Lu SH. DUSP6, a tumor suppressor, is involved in differentiation and apoptosis in esophageal squamous cell carcinoma. *Oncol Lett*. 2013.
- Woo JS, Srikanth S, Kim K-D, Elsaesser H, Lu J, Pellegrini M, et al. CRACR2A-mediated TCR signaling promotes local effector Th1 and Th17 responses. *J Immunol [Internet]*. 2018;201:1174 LP – 1185. Available from: <http://www.jimmunol.org/content/201/4/1174.abstract>.
- Shi Z, Liu J, Yu X, Huang J, Shen S, Zhang Y, et al. Decreased expression of TMEM173 predicts poor prognosis in patients with hepatocellular carcinoma. *PLoS One*. 2016;.
- Sun Y, Cao FL, Qu LL, Wang ZM, Liu XY. MEG3 promotes liver cancer by activating PI3K/AKT pathway through regulating AP1G1. *Eur Rev Med Pharmacol Sci*. 2019;.
- Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. *Science (80- ) [Internet]*. 2015;347:1260419. Available from: <http://science.sciencemag.org/content/347/6220/1260419.abstract>.
- Koochekpour S, Zhang Y, Beroukhir M, Hsieh CL, Hofer MD, Zhau HE, et al. Amplification and overexpression of prosaposin in prostate cancer. *Genes Chromosom Cancer*. 2005;.
- Saghafinia S, Mina M, Riggi N, Hanahan D, Ciriello G. Pan-cancer landscape of aberrant DNA methylation across human tumors. *Cell Rep [Internet]. Elsevier Company*; 2018;25:1066-1080.e8. Available from: <https://doi.org/10.1016/j.celrep.2018.09.082>.
- Pfeifer GP. Defining driver DNA methylation changes in human cancer. *Int J Mol Sci*. 2018;19:1–13.

38. Dong L, Ren H. Blood-based DNA methylation biomarkers for early detection of colorectal cancer. *J Proteomics Bioinform* [Internet]. 2018/06/26. 2018;11:120–6. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/30034186>.
39. Wei JH, Haddad A, Wu KJ, Zhao HW, Kapur P, Zhang ZL, et al. A CpG-methylation-based assay to predict survival in clear cell renal cell carcinoma. *Nat Commun Nature Publishing Group*. 2015;6:1–11.
40. Topper MJ, Vaz M, Marrone KA, Brahmer JR, Baylin SB. The emerging role of epigenetic therapeutics in immuno-oncology. *Nat Rev Clin Oncol* [Internet]. Springer US; 2019; Available from: <https://doi.org/10.1038/s41571-019-0266-5>.
41. Chiappinelli KB, Zahnow CA, Ahuja N, Bylin SB. Combining epigenetic and immunotherapy to combat cancer. *Cancer Res*. 2016;76:1683–9.
42. Luo N, Nixon MJ, Gonzalez-Ericsson PI, Sanchez V, Opalenik SR, Li H, et al. DNA methyltransferase inhibition upregulates MHC-I to potentiate cytotoxic T lymphocyte responses in breast cancer. *Nat Commun* [Internet]. Springer US; 2018;9:1–11. Available from: <https://doi.org/10.1038/s41467-017-02630-w>.
43. Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, Akman B, et al. Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. *Cell*. 2015;162:974–86.
44. Schwartz LH, Litière S, De Vries E, Ford R, Gwyther S, Mandrekar S, et al. RECIST 1.1 - update and clarification: From the RECIST committee. *Eur J Cancer*. 2016;.
45. Jung H, Kim HS, Kim JY, Sun JM, Ahn JS, Ahn MJ, et al. DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat Commun* [Internet]. Springer US; 2019;10:1–12. Available from: <https://doi.org/10.1038/s41467-019-12159-9>.
46. Kim K, Kim HS, Kim JY, Jung H, Sun J-M, Ahn JS, et al. Predicting clinical benefit of immunotherapy by antigenic or functional mutations affecting tumour immunogenicity. *Nat Commun* [Internet]. 2020;11:951. Available from: <https://doi.org/10.1038/s41467-020-14562-z>.
47. Davalos V, Gil-Bazo I, Xiao Q, Perez L, Moran S, Ramos I, et al. Epigenetic prediction of response to anti-PD-1 treatment in non-small-cell lung cancer: a multicentre, retrospective analysis. *Lancet Respir Med*. 2018;6:771–81.
48. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics*. 2017;33:3982–4.
49. Jerome A, Hastie T, Simon N, Tibshirani R. Package ‘glmnet’. 2017;.
50. Warnes GR. gplots: various R programming tools for plotting data. *J Phycol*. 2012;.
51. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;.
52. Leiserson MDM, Reyna MA, Raphael BJ. A weighted exact test for mutually exclusive mutations in cancer. *Bioinformatics*. 2016;.
53. Therneau TM, Grambsch PM. Modeling survival data: extending the Cox model. *Technometrics*. 2000;.
54. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* [Internet]. 2015 [cited 2015 Apr 5];43:D447–52.
55. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* [Internet]. 2010 [cited 2014 May 26];38:W214–20.
56. Camp R, Dolled-Filhart M, Rimm D. X-tile a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res*. 2004;10:7252–9.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

