

RESEARCH

Open Access



The response to influenza vaccination is associated with DNA methylation-driven regulation of T cell innate antiviral pathways

Hongxiang Fu^{1,5}, Harry Pickering², Liudmilla Rubbi¹, Ted M. Ross^{3,4}, Wanding Zhou⁵, Elaine F. Reed² and Matteo Pellegrini^{1*}

Abstract

Background The effect of vaccination on the epigenome remains poorly characterized. In previous research, we identified an association between seroprotection against influenza and DNA methylation at sites associated with the RIG-1 signaling pathway, which recognizes viral double-stranded RNA and leads to a type I interferon response. However, these studies did not fully account for confounding factors including age, gender, and BMI, along with changes in cell-type composition.

Results Here, we studied the influenza vaccine response in a longitudinal cohort vaccinated over two consecutive years (2019–2020 and 2020–2021), using peripheral blood mononuclear cells and a targeted DNA methylation approach. To address the effects of multiple factors on the epigenome, we designed a multivariate multiple regression model that included seroprotection levels as quantified by the hemagglutination-inhibition (HAI) assay test.

Conclusions Our findings indicate that 179 methylation sites can be combined as potential signatures to predict seroprotection. These sites were not only enriched for genes involved in the regulation of the RIG-I signaling pathway, as found previously, but also enriched for other genes associated with innate immunity to viruses and the transcription factor binding sites of BRD4, which is known to impact T cell memory. We propose a model to suggest that the RIG-I pathway and BRD4 could potentially be modulated to improve immunization strategies.

Keywords DNA methylation, Influenza virus, Influenza vaccine, Targeted bisulfite sequencing, RNA sequencing, Cell-type deconvolution

Background

Influenza vaccination has been recognized as the most effective method to prevent influenza virus infections and their complications, thereby reducing morbidity and mortality [1]. The development and deployment of influenza vaccines are guided by ongoing surveillance of circulating influenza strains, as the virus exhibits significant antigenic variability. This variability necessitates the annual reformulation of the influenza vaccine to match the predicted predominant circulating strains, highlighting the challenges in vaccine effectiveness and the need for continuous evaluation. Influenza vaccines work by triggering the production of antibodies against

*Correspondence:

Matteo Pellegrini
matteop@g.ucla.edu

¹ Department of Molecular Cell and Developmental Biology, University of California Los Angeles, Los Angeles, CA, USA

² Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

³ Department of Infectious Diseases, University of Georgia, Athens, GA, USA

⁴ Center for Vaccines and Immunology, University of Georgia, Athens, GA, USA

⁵ Center for Computational and Genomic Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA, USA



© The Author(s) 2024. **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.

the hemagglutinin surface glycoprotein [2]. These vaccines are designed to elicit hemagglutinin-specific antibody levels in the body, which can be quantified through the hemagglutination inhibition (HAI) assay. This assay measures the presence of functional antibodies capable of preventing the agglutination of Turkey red blood cells by the virus.

Numerous studies have sought to link the response to the influenza vaccine with various factors, including age, body mass index (BMI), and gene expression [3]. Despite these efforts, research into the epigenetic mechanisms influencing vaccine responses remains incomplete. Associating epigenetic signatures directly with vaccine responsiveness is challenging due to the influence of confounding factors, such as age and BMI, which also impact the epigenome. In this study, we focus on DNA methylation (DNAm), which is a key epigenetic mechanism that plays a crucial role in gene regulation, impacting a wide array of biological functions and disease processes [4–6]. It involves the addition of a methyl group to the fifth carbon position on cytosine residues. In the human genome, DNAm primarily occurs within the context of CpG dinucleotides. The distribution of methylation across the genome is uneven, with CpG islands typically exhibiting low levels of methylation [7]. Advances in next-generation sequencing (NGS) technologies have enabled the detailed profiling of DNAm patterns by analyzing bisulfite-treated DNA. Techniques such as whole-genome bisulfite sequencing (WGBS) and targeted bisulfite sequencing (TBS) offer precise measurements of DNAm at CpG sites [8].

In this study, we chose to profile our samples using TBS because it offers compelling advantages for the detailed analysis of DNAm patterns, especially in specific regions of interest within the genome. By focusing on predefined genomic regions, TBS enables researchers to achieve high-resolution, accurate measurements of methylation levels at CpG sites within specific target regions [9].

Previously, we identified longitudinal changes of DNAm and found that these changes were significant in genes associated with viral response pathways, such as the regulation of RIG-1 signaling, which leads to the production of interferons [10]. Complementary to our findings, another investigation utilized mice with site-specific genetic modifications and tools for targeted demethylation to establish a direct link between methylation patterns and the production of type I interferons [11]. Furthermore, an additional study pinpointed methylation sites that were associated with the humoral response to influenza vaccination. This research revealed that methylation sites predictive of humoral immune responses were notably enriched for the binding sites of polycomb-group proteins and the

FOXP2 transcription factor, underscoring the intricate relationship between epigenetic modifications and immune system function [12].

However, these studies did not address important confounders that may affect DNAm and influenza vaccination response. Notably, DNAm changes are associated with aging [13], and its patterns are highly specific to different cell types [14]. To address these potential confounding factors, including age, BMI, and cell-type proportions, our study adopted a multifactorial approach. We included the same cohort consecutively vaccinated over two years and conducted RNA sequencing (RNA-seq) for year 2020. This comprehensive approach led to the identification of key methylation sites associated not only with the previously identified RIG-I signaling pathway but also with other genes involved in innate immunity to viruses [10]. Our findings illuminate the complex interaction between DNAm and the influenza vaccine response, accounting for various potential confounders. The genes we identified may offer novel insights for future vaccine development.

Results

Study cohort and serological response to influenza vaccination

For our study, we focused on the cohorts comprising influenza vaccine recipients, enrolled at the University of Georgia between 2016 and 2020 [15]. From these cohorts, we selected a subset of 42 individuals who participated in the 2020 vaccination program (UGA5) and had also received their influenza vaccinations during the preceding 2019 program (UGA4). The administered vaccine was Fluzone (Fig. 1A), an inactivated, quadrivalent construct consisting of one strain from each of the four major subtypes (H1N1, H3N2, Yamagata, and Victoria) as described in detail in the Materials and Methods section.

Hemagglutination inhibition assay was used to quantify the amount of antibody to influenza viruses. For both cohorts, peripheral blood mononuclear cells were collected. Both targeted bisulfite sequencing and RNA sequencing were performed at baseline prior to vaccination and 28 days following vaccination (Fig. 1B). We sequenced approximately 5500 targeted regions within the genome that were selected based on different criteria as described in detail in the Methods section (Additional file 2: Supplementary Table 1). These sites were selected to be potentially involved in aging, viral responses, and cellular metabolisms. In total, we quantified levels of DNAm at approximately 100,000 CpG sites. The final methylation matrix contained 22,740 CpG sites that were covered by at least 20 reads in all the samples as described in detail in the Methods section. The multivariate model

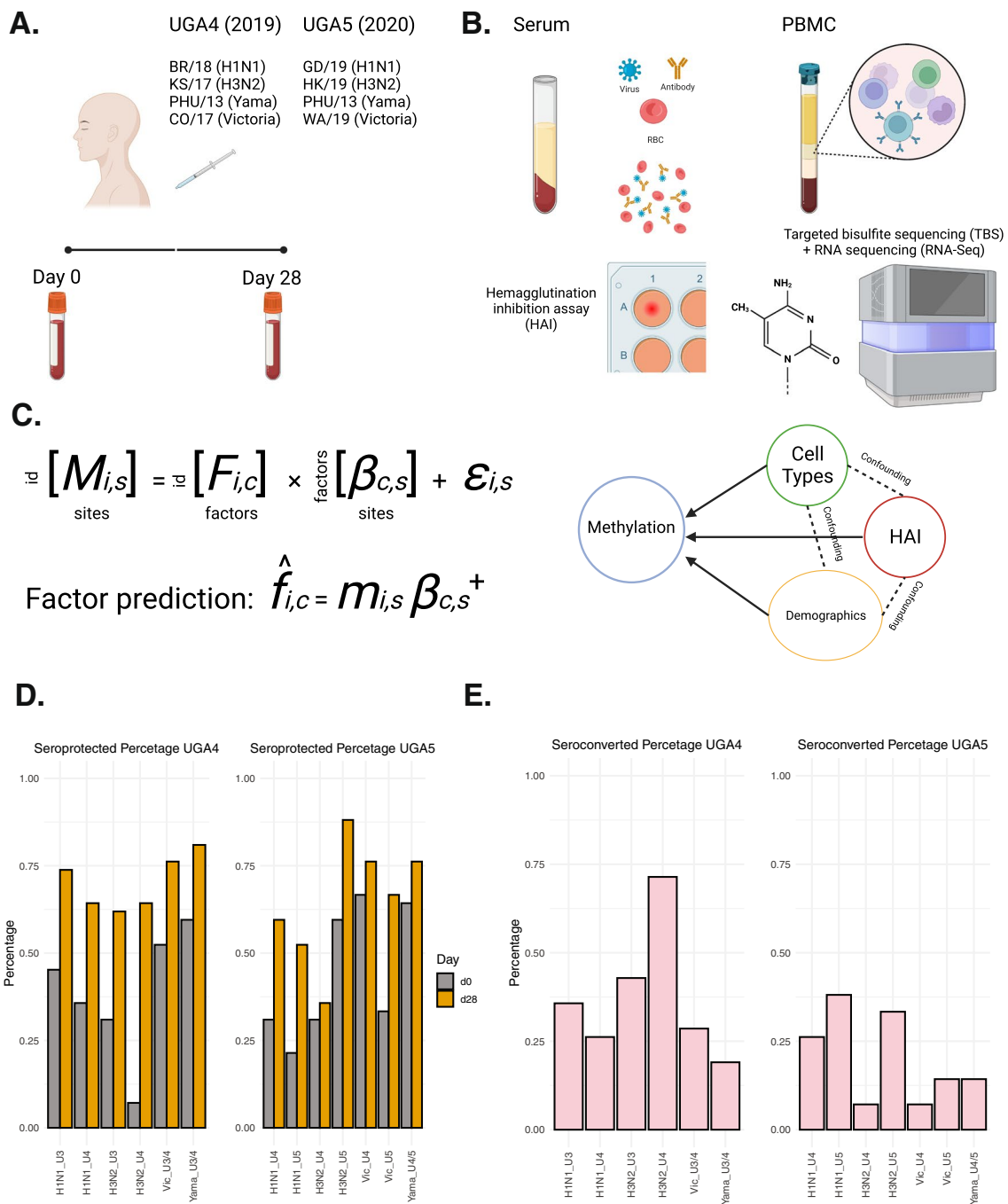


Fig. 1 Study schematic and design. **A** Vaccination summary of UGA4 (2019) and UGA5 (2020), the composition for quadrivalent Fluzone influenza vaccine in each cohort. **B** Hemagglutination inhibition assay, targeted bisulfite sequencing, and RNA sequencing schematic. **C** Model illustration, multivariate multiple regression, and pseudoinverse were used to address the effects of multiple factors on the epigenome. **D** The proportion of seroprotected subjects at day 0 and 28 days after receiving the vaccination. U3/4 indicates the vaccine strain is the same in UGA3 and UGA4. U4/5 indicates the vaccine strain is the same in UGA4 and UGA5. (E) The proportion of seroconverted subjects 28 days after receiving the vaccination

was then utilized to analyze the methylation levels with immune response under various cofounding variables (Fig. 1C).

Seroprotection is defined as HAI antibody titers $\geq 1:40$ or more post-vaccination, and seroconversion is defined as a fourfold increment in titer

compared with the baseline, leading to a titer of 1:40 or more. Each strain within the cohorts had a seroprotected percentage exceeding 50% at 28 days post-vaccination for both UGA4 and UGA5. In particular, more than 80% of the subjects were seroprotected against the H3N2 strain in UGA5 (Fig. 1D). H3N2 in UGA4 was the strain with the highest seroconverted rate, nearly 75% of the subjects seroconverted 28 days after vaccination. Overall, UGA4 participants had a higher percentage of seroconversion compared to participants in UGA5 (Fig. 1E).

Cell-type deconvolution across four time points was correlated with different serology quantifications

Methylation is known to be cell-type specific. To address the confounding factor of relative immune cell types in each PBMC sample, we conducted cell-type deconvolution analysis using CellFi [16] from the published methylation atlas of various human cell types [17]. We focused on cell types present in PBMCs and identified 37 differentially methylated regions (DMRs) for downstream cell-type abundance estimation (Fig. 2A). The relative abundance of 6 immune cell types was estimated across two cohorts at both time points (Fig. 2B). Our results align with the known composition of PBMC cells [18],

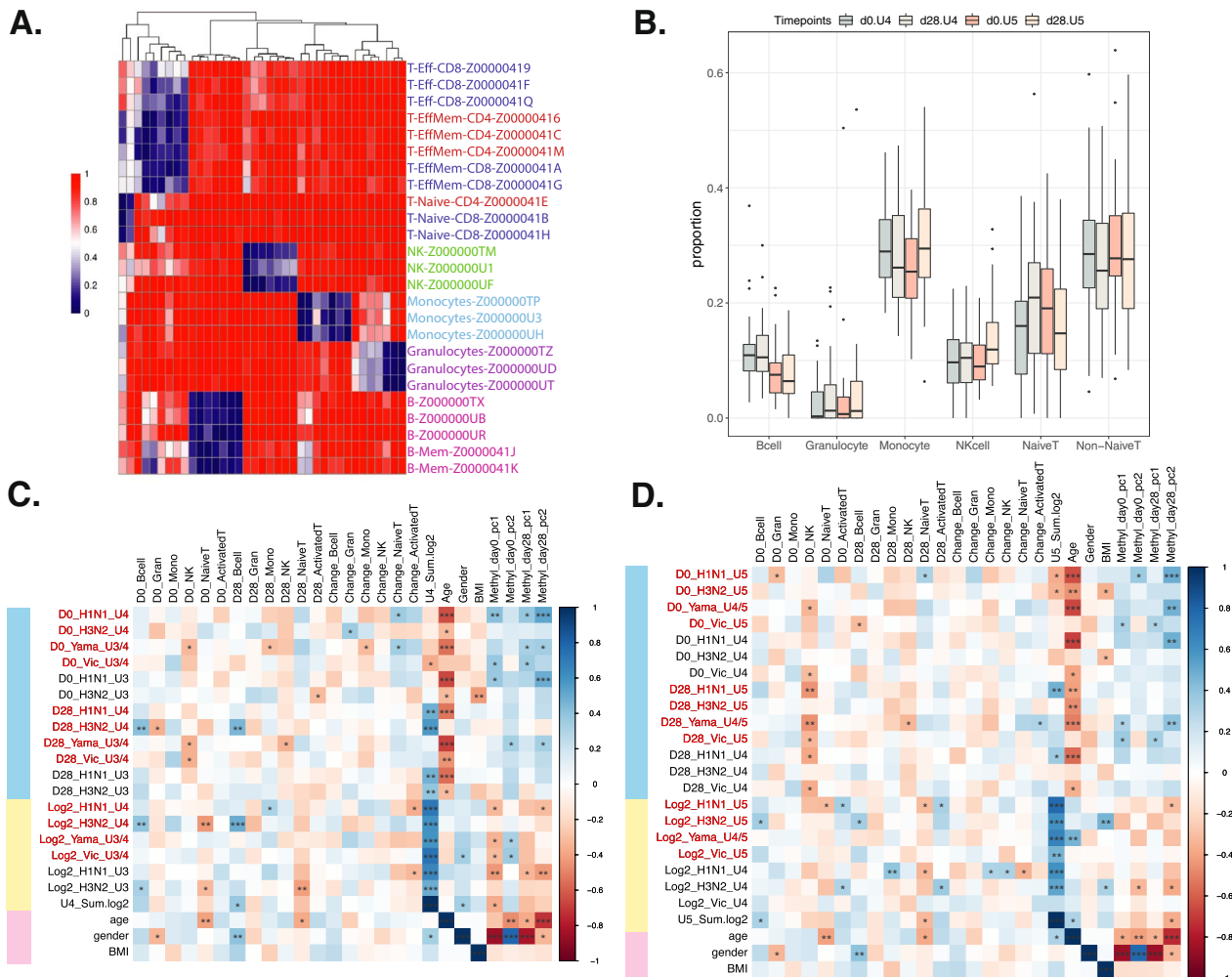


Fig. 2 Analysis of cell-type proportion estimation in two cohorts and their association with phenotypic traits. **A** 37 differentially methylated regions from blood tissues of the recently published methylation atlas, and blue indicates hypomethylation and red indicates hypermethylation. **B** Results of cell-type deconvolution for cohorts UGA4 and UGA5, assessed at two different time points. 6 cell types were selected which occur in PBMC. **C** Correlation of UGA4 cell-type estimates with various phenotypic traits. Blue rows are HAI measurements (U3/4 indicates the vaccine strain is the same in UGA3 and UGA4), yellow rows are log2 fold change of HAI from day 0 to day 28 (sum log2 is the sum of log2 fold change of the vaccine strain used in U4, responders have sum log2 ≥ 8), and pink rows are other factors. **D** Correlation of UGA5 cell-type estimates with various phenotypic traits (*Indicates $p < 0.05$, **Indicates $p < 0.01$, ***Indicates $p < 0.001$)

with ~70% being lymphocytes (including B cells, T cells, and Natural Killer (NK) cells), 25% monocytes, and a small percentage of granulocytes.

We analyzed the association between cell types and principal components of methylation with seroprotection levels (Fig. 2C, D). We observed various significant associations between cell-type estimates and serology trait measurements. These correlations matched previous findings. For example, day 0 NK cells showed negative associations with HAI titers for multiple vaccine strains, which is consistent with their role in inhibiting adaptive immune responses and impacting immune memory post-vaccination [19, 20]. Additionally, age was negatively associated for both years with HAI titers (Additional file 1: Fig. S1A–C), emphasizing its known impact on vaccine response. We also found correlations between the principal components of DNA methylation-level matrices and immune/biological traits. For example, the first three PCs showed strong gender association, and the fourth PC showed strong age association. They also demonstrated several significant positive correlations with immune responses to various strains, such as H1N1. This further reveals the importance of using a multivariate model to fully uncover the potential DNA methylation role in vaccination immune response.

Cross-validation of the multivariate multiple regression model showed nearly accurate predictions of all the variables

We utilized a multivariate multiple regression model framework to model the relationship between DNA methylation levels as the response and vaccine serological immune levels and various confounding variables as features. We calculated the Pearson correlation between these variables and found that the absolute values of the correlations between pairs of variables were all below 0.5 (Additional file 1: Fig. S2A). This indicates a lack of multicollinearity among them. The details of the model can be found in the methods section. To boost statistical power, we combined the two cohorts together and incorporated a year variable to mitigate the potential batch effect. We then applied leave-one-out cross-validation (LOOCV) to assess the model performance and prediction accuracy (Fig. 3A). This process revealed that our model yields accurate predictions of age, corroborating existing

research that methylation can be utilized to construct aging clocks [21]. Moreover, all the other variables also exhibited high Pearson correlation coefficients between predicted and true values, and all of them are statistically significant below the threshold of 0.05. To interpret the model, we examined the coefficient of the regression for each CpG site and each factor. Four Manhattan plots each show significant sites identified by t tests for the four vaccine strains administered to participants in 2019, corresponding to H1N1, H3N2, Yamagata, and Victoria strains (Fig. 3B–E). The full coefficient *p* value heatmap is shown in Fig. S2B. Sites with significant gender-specific methylations were all located on the X chromosome.

Significant HAI sites were enriched in virus response pathway and transcription factors binding sites

We performed enrichment analyses on the combined significant CpG sites identified as associated with HAI titer against each vaccine strain. To differentiate between the effects of hypermethylation and hypomethylation, we analyzed the model's coefficients separately. Positive coefficients suggest that an increase in methylation levels is associated with higher HAI. While hypermethylation of gene promoter regions generally correlates with repression of gene expression, methylation in gene bodies can be associated with active gene expression [22]. Conversely, hypomethylation often correlates with increased gene expression. 147 positive coefficients of HAI significant CpG sites were first mapped to the proximal genes, and gene set enrichment results showed viral response-related pathways against the full probe background to avoid selection bias. Notably, genes like C1QBP and RNF125, identified in this process, have been previously implicated in the RIG-I signaling pathway, as observed in the longitudinal study of the UGA4 cohort [10]. Type 1 interferon production is induced downstream of the RIG-I signaling pathway, and negative regulation of type 1 interferon production was significantly enriched in hypermethylated CpG sites (Table 1).

We extended our analysis to identify transcription factor binding sites enriched at our significant CpGs using the Cistrome database (Fig. 4B), and we filtered the sources of the ChIP data to be blood or immune related. Among the enriched transcription factors, Brd4 and NFkB1 are known to regulate immune responses [23,

(See figure on next page.)

Fig. 3 Multivariate multiple regression model LOOCV assessment and HAI significant coefficients. **A** LOOCV of the prediction from the multivariate multiple regression and pseudoinverse. The Pearson correlation between true and predicted of all the traits is all significant. **B** Manhattan plot of the significant coefficients of all the CpG sites for the 2019 H1N1 vaccine strain. **C** 2019 H3N2 vaccine strain. **D** 2019 Yamagata vaccine strain. **E** 2019 Victoria vaccine strain. Each dot represents the $-\log_{10}$ of FDR adjusted *p* values using Benjamini–Hochberg procedure for the coefficients, and the blue line indicates the significant threshold of adjusted *p* values < 0.05

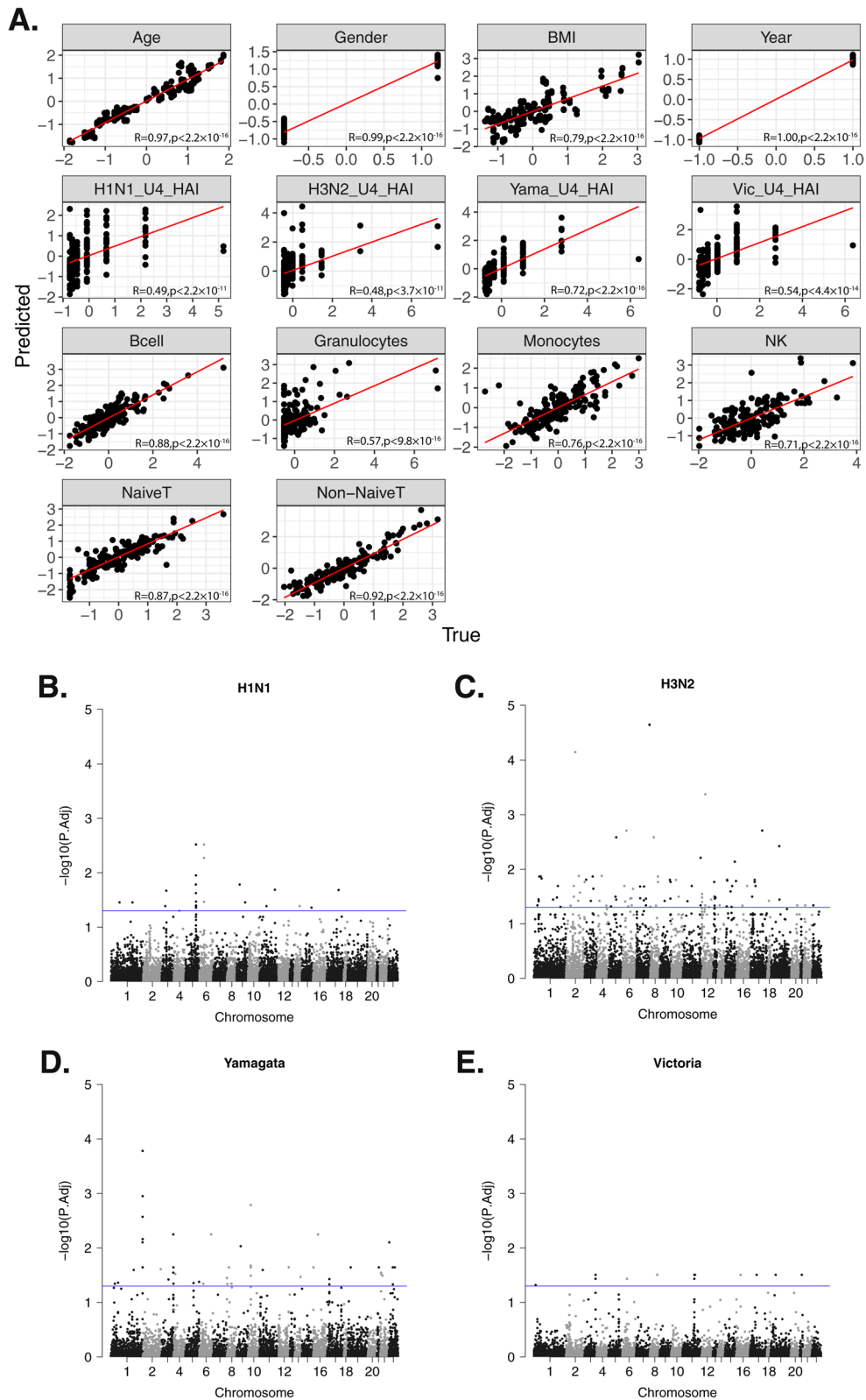


Fig. 3 (See legend on previous page.)

Table 1 Top 9 enriched pathways of significant positive HAI coefficients of the multimodal model

Pathway	Overlap	p value	Overlap genes
Regulation of defense response to virus (GO:0050688)	10/16	8.83382819583528E-09	ILRUN;PPM1B;RNF216;DHX9;ELMOD2;C1QBP;PCBP2;TNFAIP3;IFNLR1;ATG5
Negative regulation of response to biotic stimulus (GO:0002832)	7/15	2.37558081599061E-05	ILRUN;PPM1B;C1QBP;PCBP2;TNFAIP3;ATG12;ATG5
Negative regulation of type I interferon production (GO:0032480)	7/17	6.3909478276655E-05	ILRUN;PYCARD;PPM1B;RNF125;RNF216;ATG12;ATG5
Regulation of type I interferon production (GO:0032479)	7/22	4.19269386834906E-04	ILRUN;RNF125;DHX33;RNF216;ATG12;PQBP1;ATG5
Negative regulation of defense response to virus (GO:0050687)	5/11	4.5122299575726E-04	ILRUN;PPM1B;C1QBP;PCBP2;ATG5
Negative regulation of immune response (GO:0050777)	5/12	7.31430042128143E-04	PCBP2;TNFAIP3;FURIN;ATG12;ATG5
Regulation of defense response to virus by host (GO:0050691)	7/24	7.57836072328408E-04	RNF216;DHX9;SIN3A;TNFAIP3;IFNLR1;EIF2AK4;PQBP1
Negative regulation of protein localization to nucleus (GO:1900181)	3/4	0.0011758126952240700	ILRUN;SIN3A;SUFU
Negative regulation of signal transduction by P53 class mediator (GO:1901797)	3/4	0.0011758126952240700	PRKN;KDM1A;SNAI1

Significant sites were mapped to their proximal genes, and gene ontology enrichment was conducted against the full background set

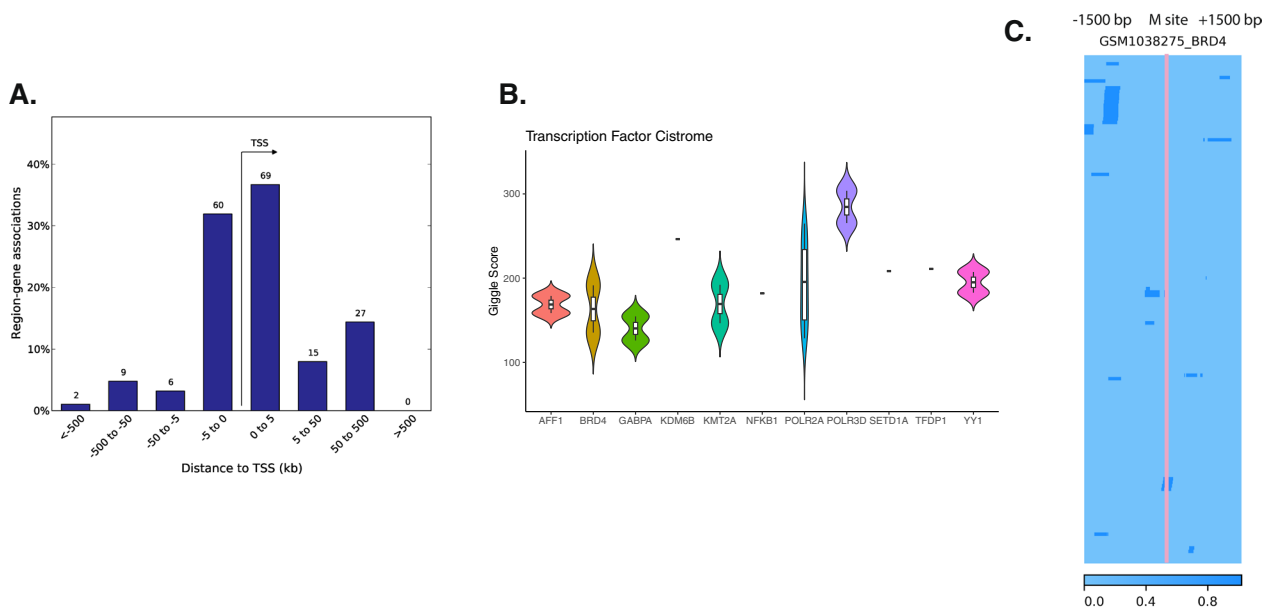


Fig. 4 Transcription factors binding sites enrichment of significant positive HAI coefficients. **A** Distance from significant positive HAI coefficient sites to transcription start sites percentage. **B** Significant transcription factor binding sites from the Cistrome database, specifically focusing on either blood or immune related sources. **C** Example of the overlap between methylation site and publicly available BRD4 ChIP-Seq data. The pink line indicates one of the significant methylation sites, and dark blue color marks the BRD4 transcription factor binding sites within 1500 base pairs of the methylation site

24]. We also cross-referenced our findings with publicly available Brd4 ChIP-Seq data [7], revealing the binding sites that are proximal to the methylation site regions (Fig. 4C). Brd4 was previously found to influence T cell memory and mediate RIG-I upregulation. It is also required by interferon regulatory factors in respiratory syncytial virus infection [25].

The significant HAI-positive coefficient sites also showed a predominant enrichment in H3K27ac, an active enhancer mark known for its association with active transcription. This finding was further supported by chromatin accessibility data, suggesting that most positive sites are located in regions of open chromatin, indicative of active transcriptional activity, and such a pattern

was less apparent in HAI-negative sites (Additional file 1: Fig. S3A–D). Our analysis also extended to examining the effects of age and Body Mass Index (BMI) on DNAm. We identified two transcription factors, Rest and Tal1, that significantly correlated with these traits, respectively (Additional file 1: Fig. S4A–D). Significant HAI-negative sites notably overlapped with LMO2 (Additional file 1: Fig. S5B), a transcription factor previously identified as a key factor in the influenza virus response [26].

Gene expression analysis of UGA5 complemented the methylation analysis result

For the UGA5 cohort, RNA sequencing (RNA-seq) data were analyzed both before and after vaccination for the same patients. Initially, we performed a differential gene expression analysis using the raw transcript count data. This approach identified 574 genes that were differentially expressed, meeting the genome-wide significance threshold (Fig. 5A). Notably, two genes, JAK3 and TYK2, exhibited a significant decrease in transcription 28 days post-vaccination (Fig. 5B). These genes are known to act as signal transducers in the adaptive immune STAT pathways, which are downstream of interferon lambda signals originating from the innate immune RIG-I pathway [27]. Interestingly, the genes corresponding to HAI-positive sites we found above did not show differential expression post-vaccination. However, we did observe a significant difference in the average transcripts per million (TPM) for the enriched Negative Regulation of Immune Response genes from HAI-positive sites on day 0 between vaccine responders and non-responders (Fig. 5C), based on change in HAI titers from pre- to 28 days post-vaccination, which suggests a potential link between baseline gene expression levels under methylation modulation

and vaccine response. We conducted additional correlation analyses between DNA methylation and gene expression linked to DNA methylation. Most correlations are negative in the promoter regions and positive in the gene body regions (Additional file 1: Fig. S6A). Specifically, we examined the gene TRIM47 and found that responders exhibit lower expression levels and higher methylation percentages on both day 0 and day 28 (Additional file 1: Fig. S6B).

In addition, we utilized RNA sequencing data from the UGA5 cohort to estimate cell-type abundances, further validating our methylation-based cell-type deconvolution results. The cell-type proportions in UGA5 demonstrated consistency between methylation and transcriptomic analyses (Additional file 1: Fig. S7A). A similar coefficient p value heatmap for selected methylation sites associated with differentially expressed genes is shown in Fig. S6B.

Discussion

We developed a multimodal approach to explore the joint influence of multiple factors on DNA methylation and the humoral immune response to influenza vaccination. In this study, we focused on 42 subjects vaccinated with quadrivalent Fluzone across two consecutive seasons (UGA4 in 2019, UGA5 in 2020), noting that the vaccine composition differed between the two years except for the Yamagata strain. Notably, the UGA5 cohort exhibited a lower seroconversion rate compared to UGA4, potentially influenced by previous vaccination histories and the concurrent COVID-19 pandemic [28, 29]. For our model, we focused on the four vaccine strains from the UGA4 cohort, for which HAI titers were also quantified in UGA5.

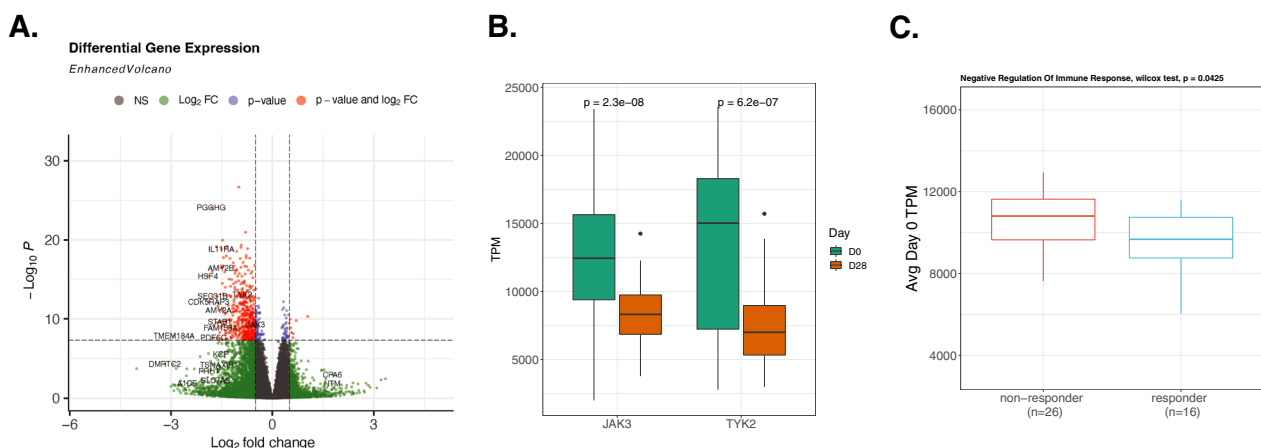


Fig. 5 RNA gene expression analysis of UGA5 cohort. **A** Differential gene expression analysis between day 0 and 28 days after vaccine injection, **B** TPM of JAK3 and TYK2 between day 0 and 28 days after vaccine injections. **C** Average day 0 TPM of Negative Regulation of Immune Response genes (PCBP2, TNFAIP3, FURIN, ATG12, ATG5) from significant HAI-positive coefficients enrichment between responder and non-responder

Methylation is a cell-type-specific epigenetic mechanism [17]. Our cell-type deconvolution provides an estimation of the proportion of PBMC cell types that can be included in the model. Interestingly, the proportion of NK cells was negatively correlated with vaccine immune response. By contrast, B cells and T cells were positively correlated with HAI seroresponse. NK cells play an important role in antiviral responses by expressing activation receptors in the innate immune system, whereas B cells and T cells function in the adaptive immune system producing antibodies and killing infected cells [30, 31].

When identifying methylation sites associated with serological immune responses to influenza, we also considered other factors, such as age and BMI, known to be associated with methylation changes. Our final multivariate multiple regression model demonstrated significant accuracy in predicting age, sex, cell types, and HAI levels against all four strains, indicating the importance of these factors in mediating the methylation at the measured sites. One hundred and forty seven hypermethylated CpG sites were associated with HAI levels against four vaccine strains, suggesting gene expression inhibition of proximal genes. These sites were mostly mapped to Negative Regulation to Defense Response to Virus and Type 1 Interferons including genes involved in the negative regulation of the RIG-I signaling pathway. Hypermethylation of those sites leads to the inhibited expression of these negative regulators and therefore likely upregulation of type 1 interferon production.

Furthermore, we investigated the overlap of these significant methylation sites with transcription factor binding sites, which identified BRD4, a well-studied member of the bromodomain and extra-terminal protein family in immune diseases and cancer [32]. The discovery that methylation sites which can predict influenza vaccine response overlap with public ChIP-Seq data from blood or immune cells suggests a potential role for BRD4 in regulating the vaccine response. It provides a molecular link between epigenetic modifications and immune efficacy. By implicating a specific transcriptional regulator, this discovery suggests that BRD4 could be a potential target for modulating vaccine responses through epigenetic mechanisms. For aging and BMI, we identified Rest and Tal1. Rest plays a crucial role in preventing senescence phenotypes, while Tal1 is implicated in high risk of obesity [33, 34]. These findings highlight Rest and Tal1 as potential targets for future studies focusing on immunosenescence and obesity in influenza vaccine response. Additionally, differential gene expression analysis of the UGA5 cohort highlighted JAK3 and TYK2 as significantly differentially expressed following vaccination, underscoring the potential involvement of the JAK signaling pathway in this network.

By combining the methylation multimodal model and differential gene expression results, we generated a final vaccine response model (Fig. 6). The methylation component (in red) mainly negatively regulates the RIG-I signaling pathway. After detecting the viral component, RIG-I signaling pathway is activated to drive the transcription of interferon production. The RIG-I signaling pathway was previously shown to be involved in antiviral responses and differential methylation analysis after influenza vaccination [10, 35, 36]. Two genes whose methylation is significantly associated with HAI act as negative regulators of the RIG-I signaling pathway. C1QBP and its receptors are targeted to the mitochondrial outer membrane, and the interaction with MAVS led to the disruption of RIG-I signaling pathway [37]. RNF125 interacts with RIG-I through proteasomal degradation after its conjugation to ubiquitin [38]. The transcription factor BRD4, coupled with RelA, mediates RIG-I upregulation, which enhances interferon response factor IRF1/7 expression by transcriptional elongation [23, 25]. The downstream output of interferons is regulated by ILRUN. ILRUN has UBA-like and NBR1-like domains which are essential for the inhibition of interferons [39].

The interferons produced by this signaling pathway can interact with the interferon lambda receptor 1 leading to signaling transduction. Interferon lambdas are innate immune cytokines that induce antiviral cellular responses [40]. The differentially expressed genes we identified are associated with the JAK/TYK2 signaling pathway which activates downstream STAT proteins. A previous study has shown that The JAK family, including Tyk2, consists of tyrosine kinases linked to receptors that serve as signal transducers [27]. The activation process of the JAK pathway begins when a cytokine, which can be interferon lambda, binds to IFNLR1. This binding induces a structural change in the receptor, which in turn activates and leads to the binding of JAK and Tyk2. These molecules form JAK dimers and then phosphorylate the receptor, facilitating the binding, phosphorylation, and subsequent pairing of STAT proteins (STAT1, STAT2, STAT3, STAT4, STAT5a, STAT5b, and STAT6) which are important for both innate and adaptive immune response [27].

We postulate that the association between DNAm patterns and the response to influenza vaccination is primarily mediated through the activation of antiviral pathways in T cells. This hypothesis was corroborated by analyzing publicly available gene expression data for human immune cells from the Human Protein Atlas [41]. Our analysis revealed that genes associated with HAI significant sites, namely RNF125, C1QBP, ILRUN, and BRD4, exhibit high levels of expression in T cells (Additional file 1: Fig. S8). Moreover, the expression of the interferon lambda receptor, predominantly observed in

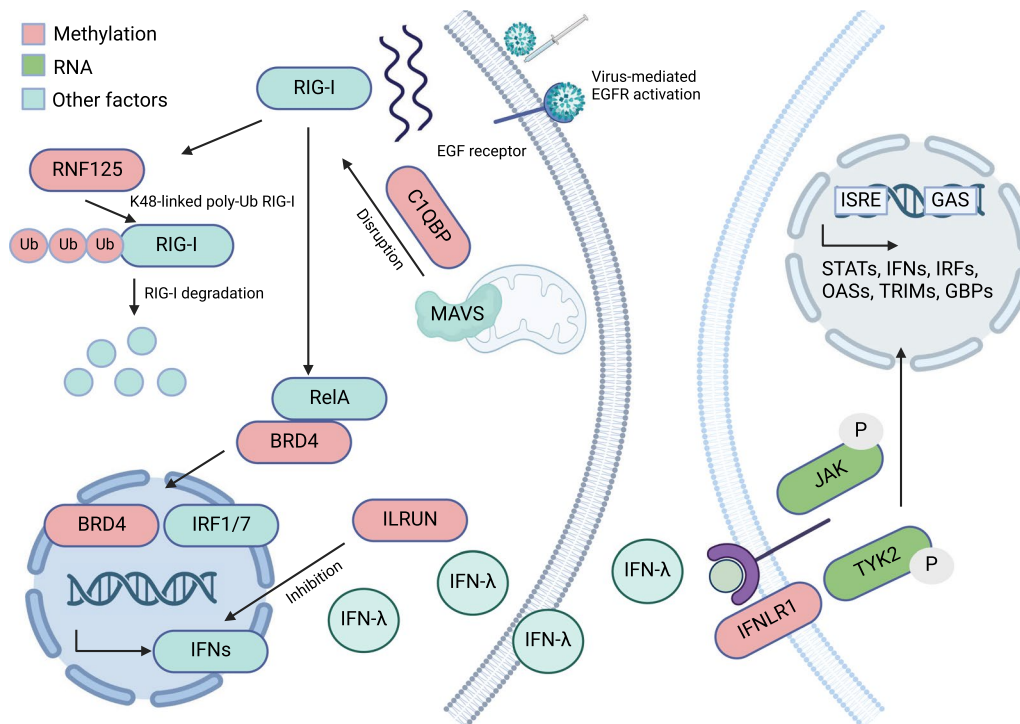


Fig. 6 Generated using Biorender (Additional file 3), the full proposed model of multimodal DNA methylation analysis and differential gene expression analysis of immune responses to influenza vaccine (T cell/B cell communication)

naive B cells, suggests a critical role for T cell to B cell communication in eliciting an immune response to the vaccine. These findings underscore the importance of T cell-mediated pathways in the context of DNAm and its impact on vaccine-induced immunity.

Overall, our findings suggest potential targets for enhancing influenza vaccine efficacy through epigenetic and transcriptomic regulation of RIG-I and related immune pathways. Future research directions include single-cell methylation analysis to dissect cell-type-specific methylation patterns and their implications for vaccine response. For example, some of the genes and transcription factors might only appear in specific cell types, which would provide a clearer picture of the cell-type-specific methylation association with vaccine response. Additionally, expanding the study to include more subjects could enable strain-specific analyses. These studies will provide more insights into the influenza vaccine response through the scope of epigenetics.

Conclusion

In conclusion, our study generated a multivariate multiple regression model to control for potential confounding variables when analyzing DNAm levels and their association with the influenza vaccine serological response. We identified significant methylation sites that are associated

with the negative regulation of interferon production, along with the transcription factor BRD4, which plays a role in this regulatory process. Additionally, our differential RNA analysis shed light on the involvement of the JAK family genes in the signal transduction pathways of interferons. These insights not only enhance our understanding of the epigenetic and transcriptomic dynamics influencing vaccine efficacy but also highlight potential targets for future influenza vaccine design. By identifying specific genes and pathways that could be modulated to improve vaccine responses, this model framework sets the stage for more targeted and effective influenza vaccination strategies.

Materials and methods

Vaccination and cohort subjects

As part of a longitudinal study led by the University of Georgia, Athens (UGA), a cohort of 690 individuals was recruited for five successive seasons from 2016 to 2020 (UGA 1–5) between the ages of 18 to 65 years old. Participants were administered the split-inactivated influenza vaccine, Fluzone, manufactured by Sanofi Pasteur. During the 2019–2020 season (UGA4), the vaccine strains incorporated were A/Brisbane/2018 (H1N1), A/Kansas/2017 (H3N2), B/Phuket/2013 (Yamagata lineage), and B/Colorado/2017 (Victoria lineage). In the following

2020–2021 season (UGA5), the strains included were A/Guangdong/2019 (H1N1), A/Hong Kong/2019 (H3N2), B/Phuket/2013 (Yamagata lineage), and B/Washington/2019 (Victoria lineage). The Institutional Review Board of the University of Georgia granted approval for the study protocols, informed consent procedures, and data gathering methodologies (IRB #3773). For this study, a subset comprising 42 participants, all of whom were enrolled in UGA 5 and had previously been part of the UGA4 season, were selected.

Hemagglutination-inhibition (HAI) assay

Blood samples were drawn from the participants on the day of vaccination (day 0), preceding vaccine injection, and then again on the seventh and twenty-eighth days post-vaccination (day 7 and day 28). Hemagglutination inhibition (HAI) assays were conducted on the day 0 and day 28 blood samples against each of the vaccine strains as well as additional strains, as elaborated in a previous study [15]. Briefly, hemagglutination inhibition (HAI) titer was ascertained by considering the reciprocal dilution of the final well consisting of non-agglutinated RBCs. For each plate, both positive and negative serum controls were incorporated. Using the World Health Organization (WHO) and the European Committee for Medicinal Products' guidelines for evaluating influenza vaccines, seroprotected subjects were defined as an HAI titer equal to or greater than 1:40, and seroconverted subjects were defined as a fourfold increment in titer compared with the baseline, leading to a titer of 1:40 or more [42].

Targeted bisulfite sequencing (TBS-seq)

DNA was isolated from each participant's blood samples using the standard phenol–chloroform extraction approach previously described [43]. For the construction of the TBS-seq library, 500 ng of the extracted DNA was used. The NEBNext Ultra II Library prep kit along with custom pre-methylated adapters (IDT) was used to perform adapter ligation and dA-tailing [9]. The probes for targeted bisulfite sequencing (TBS-seq) were chosen based on specific parameters. A first set of probes was selected from previously published DNA methylation-based aging clocks [13]. A second set was chosen to be within promoter regions of genes that respond to viral infections like SARS and influenza [44–47]. A third set included immune cell-type-specific loci as established by the Blueprint Epigenome Project [48]. The last set was chosen based on their correlation with metabolic phenotypes [49, 50]. Their coordinates, based on GRCH38, can be found in Supplementary Table 1.

The purified libraries were subsequently hybridized with the biotinylated probes under the given conditions: 2 min at 98 °C; 14 cycles of (98 °C for 20 s; 60 °C

for 30 s; 72 °C for 30 s); 72 °C for 5 min; maintained at 4 °C. Following the bisulfite treatment on captured DNA, PCR amplification was carried out using KAPA HiFi Ura-cil+. An examination of library quality was carried out using the high-sensitivity D1000 assay on a 2200 Agilent TapeStation. The libraries were then sequenced as 150 paired-end bases using two NovaSeq6000 (Sp lane).

TBS-seq data process and methylation calling

Cutadapt (v2.10) was used to remove adapter sequences from the demultiplexed FastQ files [51]. Subsequently, BSBolt (v1.3.0) was employed to align the reads to the GRCh38 reference genome [52]. The markdup function in Samtools was used to eliminate PCR duplicates [53]. The final step in this process involved calling methylation using the BSBolt CallMethylation function. To generate the methylation matrix, each CpG site was required to be covered by a minimum of 20 reads in all samples using the AggregateMatrix function in BSBolt.

RNA sequencing and analysis (RNA-seq)

RNA was isolated from peripheral blood mononuclear cells (PBMC) samples. Libraries were prepared for samples that passed quality control using the KAPA stranded-mRNA kit. Prepared libraries were sequenced on the Illumina HiSeq3000 platform. After quality control by FastQC, reads were aligned to the GRCh38 human reference genome using STAR [54] and count tables were generated using R package Rsubread [55]. For this analysis, transcripts were included with a minimum count of ≥ 5 reads in $\geq 10\%$ of samples. To normalize sequencing depth and gene length, transcripts per million (TPM) was used to quantify gene counts. Differential gene expression analysis between day 0 and day 28 as the condition design was performed using DESeq2. We used the raw count data, which DESeq2 internally performed its normalization [56]. Cell-type abundances in PBMC cells were estimated from RNA-seq data using CIBERSORT with default parameters [57].

Cell-type deconvolution for PBMC

B cells, Granulocytes, Monocytes, Natural Killer cells, Naive T cells, and Non-Naive T cells were selected as reference cell types within PBMCs. Their whole-genome bisulfite sequences were obtained from a recently published DNA methylation atlas [17]. Cell-type abundance estimation was performed using CellFi, which employs nonnegative least square methods for deconvolution [16]. We first identified 37 differentially methylated regions that were specifically hypomethylated in each of the six cell types (Additional file 4: Supplementary Table 2). We used nonnegative least square regression to estimate cell-type abundance using these 37 regions.

Multivariate multiple regression model

A multivariate multiple regression model was used to account for different factors that affect DNA methylation and influenza vaccination response:

$$[M_{is}] = [F_{ic}] \times [\beta_{cs}] + \epsilon_{is}$$

where M is the methylation matrix, i is the i th sample, and s is the s th methylation site. F is the factor matrix, and c is the c th factor. For the factor matrix, we included both the demographic variables (age, gender and BMI) and the cell-type compositions from the deconvolution method described above. To control for the year of the two cohorts, we also added a year variable where 0 is for UGA4 and 1 is for UGA5. The coefficients β were estimated through least squares estimation, and ϵ is the random error term.

$$\hat{f}_{ic} = m_{is} \times \beta_{cs}^+$$

For the prediction model, the pseudoinverse of the beta estimation was used (depicted as the plus sign) to predict the phenotype values based on the methylation levels. We utilized leave-one-out cross-validation to provide a more accurate prediction model assessment.

Enrichment analysis of significant coefficients sites

We selected the methylation sites that have significant coefficients for each trait after false discovery rate correction for multiple testing for both positive and negative coefficients separately. Gene ontology biological process enrichment was performed by mapping the significant sites to their proximal genes using biomaRt and then measuring enrichment of the gene sets using Enrichr and GREAT [58, 59]. We used the entire set of target regions as the background set for these analyses. Transcription factor binding, histone marks and variants, and chromatin accessibility overlap analyses were performed using the Cistrome Data Browser [60]. Genomic region sets that overlap with the query set were also computed using LOLA [61]. Published BRD4 ChIP-Seq data were used to check for BRD4 enrichment analysis with one example of the significant methylation site from our model [62].

Abbreviations

BMI	Body mass index
DNAm	DNA methylation
DMR	Differential methylation region
GEO	Gene Expression Omnibus
HAI	Hemagglutination inhibition assay
LOOCV	Leave-one-out cross-validation
NGS	Next-generation sequencing
PMBC	Peripheral blood mononuclear cells
RBC	Red blood cells
TPM	Transcripts per million
TBS	Targeted bisulfite sequencing
UGA	University of Georgia cohorts

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-024-01730-x>.

Additional file 1: Figure S1. **A** Age distribution of this cohort in 2019. **B** Association of age with UGA3 and UGA4 vaccine strain responsiveness in 2019. **C** Association of age with UGA4 and UGA5 vaccine strain responsiveness in 2020. Figure S2. **A** Correlation between model features of the multivariate multiple regression model. **B** Full coefficient p value heatmap of the model. **C** Overlay of Manhattan plots for four vaccine strains. Red dashed lines indicate the adjusted p value < 0.05 threshold. Figure S3. Histone mark and variants & chromatin accessibility enrichment from Cistrome. **A** Significant positive HAI coefficient for histone mark and variants. **B** Significant positive HAI coefficient for chromatin accessibility. **C** Significant negative HAI coefficient for histone mark and variants. **D** Significant negative HAI coefficient for chromatin accessibility. Figure S4. Transcription factors binding sites enrichment of age and BMI. **A** Significant positive coefficient for age. **B** Significant negative coefficient for age. **C** Significant positive coefficient for BMI. **D** Significant negative coefficient for BMI. Figure S5. Transcription factors binding sites enrichment of significant negative HAI coefficients. **A** Distance from significant negative HAI coefficient sites to transcription start sites percentage. **B** Significant transcription factor binding sites from the Cistrome database, specifically focusing on either blood or immune related sources. **C** Overlap between methylation site and publicly available ChIP-Seq data from LOLA. Figure S6. **A** Correlation between DNA methylation and gene expression between responders and non-responders at day 0 and day 28 of UGA5. Promoter regions are +/- 1000 base pairs from TSS. **B** An example gene TRIM47's expression and its corresponding methylation site methylation level at day 0 and day 28. Figure S7. **A** Estimates of cell-type proportions from RNA expression. **B** Heatmap of coefficients for selected DNA methylation sites mapped to differentially expressed genes. Figure S8. Human protein atlas immune cell gene expression. **A-E** RNF125, C1QBP, ILRUN, BRD4, IFNLR1.

Additional file 2: The coordinates (GRCH38) of the probes (5500 targeted regions) that cover the selected CpG sites.

Additional file 3: Biorender publication license for Fig. 6.

Additional file 4: Differential methylated regions for CellFi methylation deconvolution.

Additional file 5: Describing participants' demographics, age, BMI, and antigen titer for each strain of the HAI assay for UGA4 in 2019.

Additional file 6: Describing participants' demographics, age, BMI, and antigen titer for each strain of the HAI assay for UGA5 in 2020.

Additional file 7: Raw count matrix for UGA5 RNA-seq data.

Acknowledgements

The authors would like to thank all UGA4 and UGA5 influenza vaccine program participants. The authors thank Hannah Hanley for program coordination, as well as the members of the vaccine research unit collection, processing, and analysis teams, including Brittany Baker, Jasmine Burris, Michael Carlock, Jasper Gattiker, Omar Hamwy, Lauren Howland, Katie Mailloux, and Emma Whitesell. We also thank Jonathan Murrow, Brad Phillips, Kim Schmitz, and the entire members of the UGA Clinical Trials Evaluation Unit.

Author contributions

MP and EFR were the primary instigators of this study, overseeing its design and securing the necessary funding. The study's conceptual underpinnings were established by HP. HF led the data analysis and wrote the manuscript. The provision of the UGA cohorts was facilitated by TMR, and valuable contributions to the sequencing data were made by LR. Interpretation of results and significant revisions to the manuscript were the collaborative effort of HF, HP, WZ, EFR, and MP. Every author has made critical intellectual contributions to this project and engaged in significant discussions. After a critical review, all authors have given their endorsement to the final manuscript.

Funding

Financial support for the influenza vaccination program was accorded by the National Institute of Allergy and Infectious Diseases (NIAID), an affiliate of the US National Institutes of Health (NIH), as per the Department of Health and Human Services Contract 75N93019C00052. Additional support was granted by the University of Georgia (US) via the UGA-001 grant. Moreover, T.M.R. is supported by the Georgia Research Alliance as an Eminent Scholar. The UGA Clinical Trials Evaluation Unit benefited from the backing of the National Center for Advancing Translational Sciences, a division of the National Institutes of Health, under the auspices of Award Number UL1TR002378. It must be noted that the aforementioned funders maintained no involvement in the conceptualization, execution, interpretation, or manuscript preparation of this study, nor did they influence the decision to publish. The content within this publication does not necessarily reflect the opinions or policies of the Department of Health and Human Services. Moreover, any mention of specific trade names, commercial products, or organizations does not constitute an endorsement by the US Government.

Availability of data and materials

All the raw sequence FastQ data and processed methylation CGmap data are uploaded to Gene Expression Omnibus (GEO) and are accessible through GEO series accession number GSE263782. Metadata is uploaded as Additional files 5 and 6 for UGA4 and UGA5. The raw RNA count matrix is uploaded as Additional file 7.

Declarations

Ethics approval and consent to participate

The University of Georgia Institutional Review Board has granted approval for this study. Informed consent for participation in the study was secured in written form from all participants, along with their parents or legal guardians.

Consent for publication

The manuscript is approved by all authors for publication. Additionally, written authorization was received for publication from all parties involved, including the participants themselves, their parents, or their respective legal guardians.

Competing interests

The authors declare that they have no competing interests.

Received: 25 April 2024 Accepted: 12 August 2024

Published online: 21 August 2024

References

- Osterholm MT, Kelley NS, Sommer A, Belongia EA. Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis. *Lancet Infect Dis*. 2012;12(1):36–44.
- Webby R. Understanding immune responses to the influenza vaccine. *Nat Med*. 2016;22(12):1387–8.
- Castrucci MR. Factors affecting immune responses to the influenza vaccine. *Hum Vaccin Immunother*. 2017;14(3):637–46.
- Palou-Márquez G, Subirana I, Nonell L, Fernández-Sanlés A, Elosua R. DNA methylation and gene expression integration in cardiovascular disease. *Clin Epigenetics*. 2021;13(1):75.
- Jin Z, Liu Y. DNA methylation in human diseases. *Genes Dis*. 2018;5(1):1–8.
- Mazzone R, Zwergel C, Artico M, Taurone S, Ralli M, Greco A, et al. The emerging role of epigenetics in human autoimmune disorders. *Clin Epigenet*. 2019;11(1):34.
- Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology*. 2013;38(1):23–38.
- Li Y, Tollefsbol TO. DNA methylation detection: bisulfite genomic sequencing analysis. *Methods Mol Biol*. 2011;791:11–21.
- Morselli M, Farrell C, Rubbi L, Fehling HL, Henkhaus R, Pellegrini M. Targeted bisulfite sequencing for biomarker discovery. *Methods*. 2021;187:13–27.
- Fu H, Pickering H, Rubbi L, Ross TM, Reed EF, Pellegrini M. Longitudinal analysis of influenza vaccination implicates regulation of RIG-I signaling by DNA methylation. *Sci Rep*. 2024;14(1):1455.
- Gao Z, Li W, Mao X, Huang T, Wang H, Li Y, et al. Single-nucleotide methylation specifically represses type I interferon in antiviral innate immunity. *J Exp Med*. 2021;218(3):e20201798.
- Zimmermann MT, Oberg AL, Grill DE, Ovsyannikova IG, Haralambieva IH, Kennedy RB, et al. System-wide associations between DNA-methylation, gene expression, and humoral immune response to influenza vaccination. *PLoS ONE*. 2016;11(3):e0152034.
- Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet*. 2018;19(6):371–84.
- Scott CA, Duryea JD, MacKay H, Baker MS, Laritsky E, Gunasekara CJ, et al. Identification of cell type-specific methylation signals in bulk whole genome bisulfite sequencing data. *Genome Biol*. 2020;21(1):156.
- Abreu RB, Kirchenbaum GA, Clutter EF, Sautto K, Ross TM. Preexisting subtype immunodominance shapes memory B cell recall response to influenza vaccination. *JCI Insight [Internet]*. 2020 Jan 16 [cited 2023 Jun 2];5(1). Available from: <https://insight.jci.org/articles/view/132155#SEC4>.
- Morselli M, Farrell C, Montoya D, Gören T, Sabırlı R, Türkçüer İ, et al. DNA methylation profiles in pneumonia patients reflect changes in cell types and pneumonia severity. *Epigenetics*. 2022;17(12):1646–60.
- Loyfer N, Magenheimer J, Peretz A, Cann G, Bredno J, Klochender A, et al. A DNA methylation atlas of normal human cell types. *Nature*. 2023;613(7943):355–64.
- Kleiveland CR. Peripheral Blood Mononuclear Cells. In: Verhoeckx K, Cotter P, López-Expósito I, Kleiveland C, Lea T, Mackie A, et al., editors. *The impact of food bioactives on health: in vitro and ex vivo models* [Internet]. Cham (CH): Springer; 2015 [cited 2024 Jan 21]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK500157/>.
- Rydzynski C, Daniels KA, Karme EP, Brooks TR, Mahl SE, Moran MT, et al. Generation of cellular immune memory and B-cell immunity is impaired by natural killer cells. *Nat Commun*. 2015;6(1):6375.
- Rydzynski CE, Cranert SA, Zhou JQ, Xu H, Kleinstein SH, Singh H, et al. Affinity maturation is impaired by natural killer cell suppression of germinal centers. *Cell Rep*. 2018;24(13):3367–3373.e4.
- Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14(10):R115.
- Wang Q, Xiong F, Wu G, Liu W, Chen J, Wang B, et al. Gene body methylation in cancer: molecular mechanisms and clinical applications. *Clin Epigenetics*. 2022;14(1):154.
- Wang N, Wu R, Tang D, Kang R. The BET family in immunity and disease. *Sig Transduct Target Ther*. 2021;6(1):1–22.
- Li Q, Verma IM. NF- κ B regulation in the immune system. *Nat Rev Immunol*. 2002;2(10):725–34.
- Tian B, Yang J, Zhao Y, Ivanciuc T, Sun H, Garofalo RP, et al. BRD4 couples NF- κ B/RelA with airway inflammation and the iRF-RIG-I amplification loop in respiratory syncytial virus infection. *J Virol*. 2017;91(6):e00007–17.
- Chen G, Li H, Hao M, Li X, Dong Y, Zhang Y, et al. Identification of critical genes and pathways for influenza A virus infections via bioinformatics analysis. *Viruses*. 2022;14(8):1625.
- Rusiñol L, Puig L. Tyk2 targeting in immune-mediated inflammatory diseases. *Int J Mol Sci*. 2023;24(4):3391.
- Kim HK, Min KD, Cho S. Analysis of the effectiveness of non-pharmaceutical interventions on influenza during the Coronavirus disease 2019 pandemic by time-series forecasting. *BMC Infect Dis*. 2023;23(1):717.
- Qiu Z, Cao Z, Zou M, Tang K, Zhang C, Tang J, et al. The effectiveness of governmental nonpharmaceutical interventions against COVID-19 at controlling seasonal influenza transmission: an ecological study. *BMC Infect Dis*. 2022;22(1):331.
- Yokoyama WM. Natural killer cell immune responses. *Immunol Res*. 2005;32(1–3):317–25.
- Vitetta ES, Berton MT, Burger C, Kepron M, Lee WT, Yin XM. Memory B and T cells. *Annu Rev Immunol*. 1991;9:193–217.
- Donati B, Lorenzini E, Ciarrocchi A. BRD4 and cancer: going beyond transcriptional regulation. *Mol Cancer*. 2018;17(1):164.
- Rocchi A, Carminati E, De Fusco A, Kowalska JA, Floss T, Benfenati F. REST/NRSF deficiency impairs autophagy and leads to cellular senescence in neurons. *Aging Cell*. 2021;20(10):e13471.

34. Samaan Z, Lee YK, Gerstein HC, Engert JC, Bosch J, Mohan V, et al. Obesity genes and risk of major depressive disorder in a multiethnic population: a cross-sectional study. *J Clin Psychiatry*. 2015;76(12):e1611-1618.
35. Cao W, Liepkalns JS, Kamal RP, Reber AJ, Kim JH, Hofstetter AR, et al. RIG-I ligand enhances the immunogenicity of recombinant H7 HA protein. *Cell Immunol*. 2016;304–305:55–8.
36. Hao Q, Jiao S, Shi Z, Li C, Meng X, Zhang Z, et al. A non-canonical role of the p97 complex in RIG-I antiviral signaling. *EMBO J*. 2015;34(23):2903–20.
37. Xu L, Xiao N, Liu F, Ren H, Gu J. Inhibition of RIG-I and MDA5-dependent antiviral response by gC1qR at mitochondria. *Proc Natl Acad Sci U S A*. 2009;106(5):1530–5.
38. Arimoto K, Takahashi H, Hishiki T, Konishi H, Fujita T, Shimotohno K. Negative regulation of the RIG-I signaling by the ubiquitin ligase RNF125. *Proc Natl Acad Sci*. 2007;104(18):7500–5.
39. Ambrose RL, Brice AM, Caputo AT, Alexander MR, Tribolet L, Liu YC, et al. Molecular characterisation of ILRUN, a novel inhibitor of proinflammatory and antimicrobial cytokines. *Heliyon*. 2020;6(6):e04115.
40. Evans JG, Novotny LA, Meissner EG. Influence of canonical and non-canonical IFNLR1 isoform expression on interferon lambda signaling. *Viruses*. 2023;15(3):632.
41. Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, et al. Impact of genetic polymorphisms on human immune cell gene expression. *Cell*. 2018;175(6):1701–15.
42. Hannoun C, Megas F, Piercy J. Immunogenicity and protective efficacy of influenza vaccination. *Virus Res*. 2004;103(1–2):133–8.
43. Guha P, Das A, Dutta S, Chaudhuri TK. A rapid and efficient DNA extraction protocol from fresh and frozen human blood samples. *J Clin Lab Anal*. 2018;32(1):e22181.
44. Rehwinkel J, Gack MU. RIG-I-like receptors: their regulation and roles in RNA sensing. *Nat Rev Immunol*. 2020;20(9):537–51.
45. Sun L, Liu S, Chen ZJ. SnapShot: pathways of antiviral innate immunity. *Cell*. 2010;140(3):436–436.e2.
46. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. 2020;181(2):271–280.e8.
47. Lukassen S, Chua RL, Trefzer T, Kahn NC, Schneider MA, Muley T, et al. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *EMBO J*. 2020;39(10):e105114.
48. Martens JHA, Stunnenberg HG. BLUEPRINT: mapping human blood cell epigenomes. *Haematologica*. 2013;98(10):1487–9.
49. Gomez-Alonso MDC, Kretschmer A, Wilson R, Pfeiffer L, Karhunen V, Seppälä I, et al. DNA methylation and lipid metabolism: an EWAS of 226 metabolic measures. *Clin Epigenet*. 2021;13(1):7.
50. Reed ZE, Suderman MJ, Relton CL, Davis OSP, Hemani G. The association of DNA methylation with body mass index: distinguishing between predictors and biomarkers. *Clin Epigenetics*. 2020;12(1):50.
51. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011;17(1):10–2.
52. Farrell C, Thompson M, Tosevska A, Oyetunde A, Pellegrini M. Bisulfite Bolt: a bisulfite sequencing analysis platform. *GigaScience*. 2021;10(5):giab033.
53. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
54. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
55. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res*. 2019;47(8):e47.
56. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
57. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–7.
58. McLean CY, Bristol D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*. 2010;28(5):495–501.
59. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*. 2009;4(8):1184–91.
60. Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, et al. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res*. 2019;47(D1):D729–35.
61. Sheffield NC, Bock C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and bioconductor. *Bioinformatics*. 2016;32(4):587–9.
62. Lovén J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell*. 2013;153(2):320–34.

Publisher's Note

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