

# 6-base genome reveals the reality of DNA methylation biomarkers of ageing

Mark Consgar<sup>1</sup>, Thao Huynh<sup>1</sup>, Fabio Puddu<sup>1</sup>, Annelie Johansson<sup>1</sup>, Emriia Lleshi<sup>1</sup>, Robert Crawford<sup>1</sup>, Tom Charlesworth<sup>1</sup>, Robert J Osborne<sup>1</sup>, Páidí Creel<sup>1</sup>

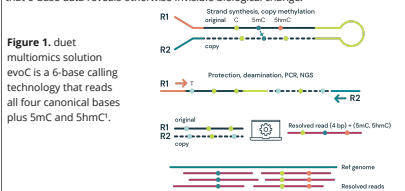
<sup>1</sup> biomodal Ltd, The Trinity Building, Chesterford Research Park, Cammelburg, UK

## 1. Introduction

DNA methylation is well established as a biomarker of ageing and has been used in 'clocks' to determine biological age vs chronological age. However ageing studies to date have evaluated total DNA methylation through a modified cytosine (modC) readout using array and bisulphite techniques. However, a modC readout is a conflation of 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) which have distinct, and opposing biological functions.

Here we use 6-base sequencing (Fig 1. a) using duet evoC, a multiomic solution that detects the four canonical DNA bases along with 5mC and 5hmC, to comprehensively profile the epigenome of multiple tissues across the ageing spectrum.

We use the modality XPLR software to explore the levels and distribution of the two DNA modifications across the genome. These explorations reveal differential distributions of the two methylation marks at genes and promoters and reinforce that 6-base data reveals otherwise invisible biological change.



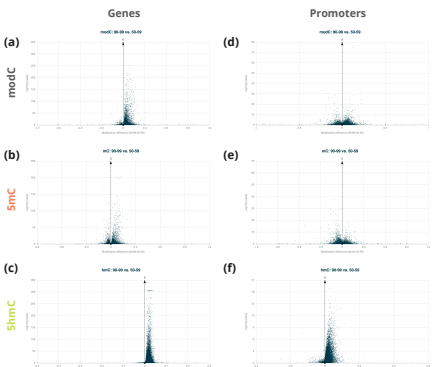
## 4. Differential methylation analysis reveals age-specific 5mC and 5hmC signatures

Figure 4. DMR plots were generated using the modality XPLR analysis software

To explore changes in methylation during ageing, modC (a and d), 5mC (b and e) and 5hmC (c and f) DMRs were calculated for gene (a - c) or promoter (d - f) regions comparing data from colon and breast tissues from two individuals ages 54 and 93. Plotted are extent of change on the x axis against the significance of those changes on the y axis, with the most significant changes top left and top right for loss and gain of methylation, respectively.

Comparing DMRs across promoters, the distribution of modC DMRs appear most similar to 5mC DMRs which is expected due to the higher level of 5mC constituting the majority of a conflated modC signal. In contrast the 5hmC DMRs for promoters have a different distribution where most differentially hydroxymethylated regions have a gain of 5hmC, indicating they are on an activating trajectory.

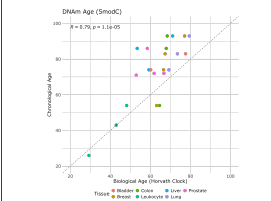
For gene DMRs a modC readout indicates there is more gain of methylation than loss, and this is reflected in both the 5mC and 5hmC DMRs. However gain of 5mC and gain of 5hmC have opposite functional impacts that is lost if just evaluating modC data. There are genes that are being deactivated (gain of 5mC) but also many genes that are being activated (gain of 5hmC). Overall we see both gain and loss of 5mC or 5hmC, supporting the hypothesis that epigenetic information is lost with age<sup>2</sup>.



## 2. 6-base whole genome data enables biological age estimation

Figure 2. Using 6-base data to mimic a modC readout demonstrates good alignment between chronological and biological age

With 6-base data it is possible to create an artificial modC readout by combining 5mC and 5hmC data. This enables comparison to tools developed using legacy methylation sequencing technologies. However with 6-base data it is also possible to use the functionally distinct 5mC and 5hmC information to derive a better understanding of biological processes of ageing.

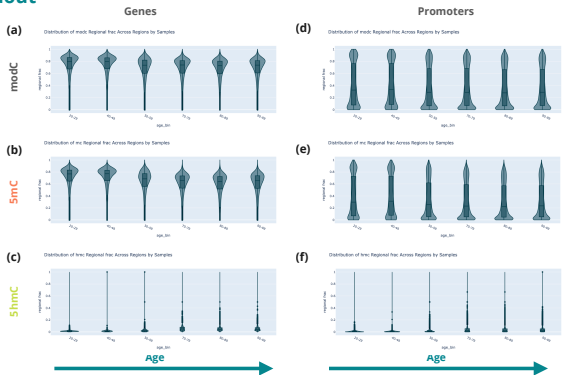


## 3. The 6-base genome reveals the reality of functionally distinct 5mC and 5hmC dynamics of ageing that are invisible with a modC readout

Figure 3. Violin plots were generated using the modality XPLR analysis software to assess modC, 5mC and 5hmC levels

modC (a and d), 5mC (b and e) and 5hmC (c and f) fractions across genes were used to explore methylation dynamics with age. With a modC readout there appears to be no change in methylation fraction over these regions, however considering 5mC and 5hmC separately reveals the otherwise invisible reality.

6-base data also enables us to develop mechanistic hypotheses where decreases in 5mC are associated with transcriptional activation, especially when coupled with increases in 5hmC. This data suggests that the loss of 5mC is coupled to a gain of 5hmC, which would indicate opposite mechanistic changes as we age than would be expected with modC data alone. We could then hypothesise that there is an increase in gene activation as we age, which might be driven by stochastic demethylation mediated by the TET enzymes. However, these data are derived from multiple tissues that are not balanced across the age categories represented here. Therefore further analysis is required to explore the robustness of these observations and evaluate whether we can derive reliable mechanistic insight.



## 5. Understanding functional implications of methylation changes associated with ageing

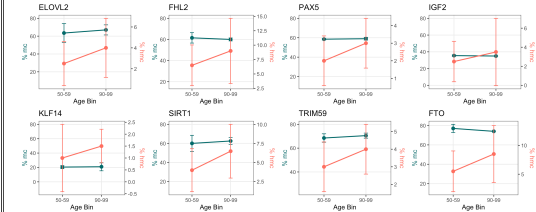


Figure 5. Plotting 5mC and 5hmC levels at ageing associated genes using modality XPLR text-based DMR BED files. Ageing related genes with significant 5hmC DMRs when comparing data from colon and breast tissues from two individuals ages 54 and 93 (Fig 4). For these regions the 5mC changes were not significant when performing a 5mC DMR analysis.

These data reinforce the importance of understanding 5hmC dynamics. For these genes in these samples, changes in 5mC did not meet the significance threshold and there was also no modC DMR. These changes would be invisible when using only modC or mC information, which would be expected to increase chromatin accessibility and transcriptional activity given the gain in 5hmC.

## 5. Conclusion

This study demonstrates the value of moving beyond legacy epigenetic sequencing technologies that provide a modC readout. Plotting 6-base data from multiple tissues we see global changes in 5mC and 5hmC that are obscured when using modC data. More specific comparisons 6-base data reveals region-specific changes in 5mC and 5hmC across both gene and promoter regions. We see both gain and loss of 5mC and almost entirely gain of 5hmC with age when comparing colon and breast tissue. This gain of 5hmC is once again obscured in a modC readout. Here, the 6-base data reveals a subset of ageing specific genes have significant changes in 5hmC but not 5mC or modC, which could indicate activation with age, in these samples.

While the possibility to make general claims based on this comparison is limited due to a single pair of individuals being explored, we feel it suggests a phenomenon which should be further explored in a designed experiment using a larger cohort.

## 6. References

- Füllgraber J, et al. Simultaneous sequencing of genetic and epigenetic bases in DNA. *Nat Biotechnol.* 2023 Oct;41(10):1457-1464.
- Yang JH et al. Loss of epigenetic information as a cause of mammalian ageing. *Cell.* 2023;186(2):305-326.e27. doi:10.1016/j.cell.2022.12.027



View online  
bioModal