

RESEARCH

Open Access



DNA methylation patterns re-establish during clonal expansion restoring cell-to-cell epialleles heterogeneity

Davide Costabile¹, Maddalena Russo², Rosa Della Monica¹, Michela Buonaiuto¹, Giulia De Riso^{1,3}, Marta Sabbarese¹, Giuseppe Catapano⁴, Sergio Coccozza¹, Roberta Visconti⁵, Lorenzo Chiariotti^{1,2*} and Mariella Cuomo^{1,2*}

Abstract

It is still debated whether the combination of methylated and unmethylated CpGs, defined as epialleles, is stably propagated during cell division or whether epiallele diversity can dynamically emerge over time. By analyzing epiallele profiles in putative single-cell-derived clones, we sought to track the behaviour of DNA methylation profiles through cell divisions to determine whether epialleles are faithfully copied from mother to daughter cells, or instead a heterogeneous epiallele population is generated, giving rise to cell-to-cell methylation variability. Our findings revealed that: i) clonal expansion did not result in the maintenance of a single dominant epiallele configuration; ii) expanding clones progressively re-established a broad repertoire of epialleles similar to that observed in the mixed population; iii) although the major epiallele structures were partially preserved, each clone displayed a distinct epiallele distribution characterized by differences in epiallele frequencies relative to the bulk population. Notably, these clone-specific epiallele distributions remained relatively stable as cell numbers increased during proliferation. Overall, our findings suggest that epiallele diversity reemerges during clonal expansion, consistent with the dynamic behavior of DNA methylation at individual CpG sites and cell-to-cell variability in DNA methylation patterns

Keywords DNA methylation, Epiallele, Epigenetic heterogeneity

Introduction

Despite being considered a stable epigenetic modification [1–4], increasing evidence suggests that DNA methylation establishment is more dynamic and heterogeneous than previously anticipated [5–9]. Assessing DNA methylation heterogeneity in bulk cells requires high-coverage amplicon bisulfite sequencing analyses [10–13]. These analyses reveal the combinatorial patterns of DNA methylation across individual DNA molecules, referred to as epialleles [11, 12, 14, 15]. Previous studies have demonstrated that epiallele analysis can capture the oscillatory behavior of DNA methylation since epiallele patterns reflect the degree and the relative tendency of specific CpG sites to be methylated within a locus [10–12, 14, 16,

*Correspondence:

Lorenzo Chiariotti
lorenzo.chiariotti@unina.it
Mariella Cuomo
mariella.cuomo@unina.it

¹CEINGE Advanced Biotechnologies “Franco Salvatore”, via G. Salvatore 486, 80145 Naples, Italy

²Department of Molecular Medicine and Medical Biotechnologies, University of Naples “Federico II”, Via S. Pansini 5, 80131 Naples, Italy

³AOU Federico II, via S. Pansini 5, 80131 Naples, Italy

⁴Neurosurgery Unit, “Ospedale del Mare” Hospital, Napoli, Italy

⁵Department of Pharmacy, University of Naples “Federico II”, Via D. Montesano 49, 80131 Naples, Italy



17]. Notably, we previously reported a high degree of epipolymorphism in brain development, suggesting that the epiallele composition may serve as an epigenetic “identity card” that distinguishes specific brain regions and cell types [11, 14, 17]. However, it remains unclear whether each epiallele configuration is clonally transmitted through cell divisions or whether they are dynamically re-established over successive generations. Specifically, it is unknown whether a single cell preserves its epiallele composition in its progeny or whether daughter cells progressively acquire a broader spectrum of epialleles as a consequence of DNA methylation plasticity. To address these questions, we isolated putative single cells and expanded them into individual clones, ensuring that the observed epiallele composition originated from a putative single progenitor cell. By tracking epiallele trajectories at previously characterized loci [18, 19] and across different genomic regions, we observed that cell expansion does not lead to the fixation of a single (or few) epiallele configuration, which likely reflects epiallele clonal transmission. Instead, it is associated with the reemergence of diverse epiallele profiles with dominant epialleles preserved across cell divisions. Moreover, in some cases, the relative frequencies of epialleles subtly shift, leading each clone to establish a new stable methylation equilibrium. Notably, the tendency of the single CpG sites to be methylated or not methylated remained conserved, supporting a model in which DNA methylation dynamics generate structured epiallele diversity while preserving locus-specific CpG methylation tendencies. In this framework, epigenetic heterogeneity emerges as a controlled system, shaped by the average methylation level of the specific genomic region and independent of the specific genomic context.

Results

Experimental setup

To unravel the trajectories that epiallele profiles follow from a putative single cell, we designed an experimental system consisting of different conditions (Fig. 1). We diluted a confluent U251 “mother” plate (MP, ~4 million cells) to generate putative single, separated cells that were grown until they formed single clones, here defined as diluted mother plate (DMP). Several well-separated clones were then collected from the DMP. Several well-isolated clones were collected from the DMP and expanded independently until reaching ~20,000, ~40,000, ~200,000 cells, and full confluence (~4 million cells), referred to as 20k, 40k, 200k, and confluent clones (CC), respectively. In parallel, the remaining unpicked clones from the DMP (approximately 50–100 clones) were pooled and collected as a reference population (here defined as Pool). DNA methylation was analyzed across these experimental conditions at four levels of progressively increasing resolution: (i) average methylation levels across all analyzed CpG sites; (ii) methylation level at each single-CpG site; (iii) epiallele classes distributions, grouping molecules by the number of methylated CpGs irrespective of position; and (iv) full epiallele profiles, accounting for the specific positional arrangement of methylated and unmethylated CpGs along individual DNA molecules. Assuming that DNA methylation is not bimodally distributed, regions displaying intermediate methylation levels are expected to be the most informative for assessing epiallele diversity and DNA methylation heterogeneity, compared with loci showing either very low or very high methylation. On this basis, we initially focused on the *DAO* and *DDO* genes as model loci, as they exhibit intermediate DNA methylation levels [11, 14, 17, 19] together with highly reproducible epiallele distributions [11, 14, 17, 19]. We subsequently extended the same analytical framework to additional genomic

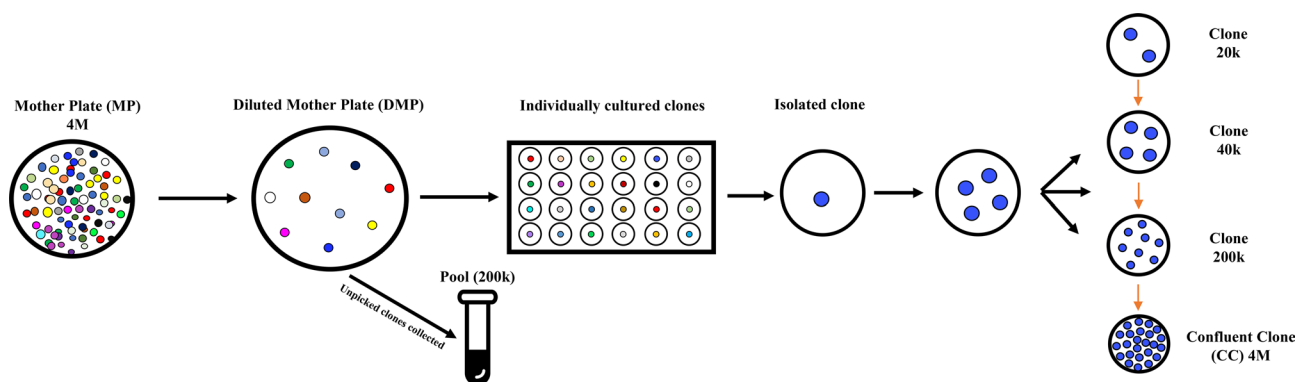


Fig. 1 Schematic representation of the experimental setup. In the mother plate (MP), each colored sphere represents a cell. In the diluted mother plate (DMP), each colored sphere represents a group of cells derived from the same mother cell (clone). Clones from the DMP were picked and plated individually into single wells to grow separately, while unpicked clones were pooled and collected. When a consistent number of cells was reached, each clone was split into four separate plates. The cell number thresholds before harvesting for analysis were about 20,000, 40,000, 200,000 and 4 million cells

regions and contexts, including the *ALB* and *LGALS1* promoters, *SRR1* and *SRR2* enhancers, *ALB* intron, *LINE1* and *Sat2* repetitive elements, and *ADAMTS2* late-replicating regions in the context of solo-WCGW sites.

DNA methylation and epiallele distribution at *DAO* and *DDO* promoters in U251 cells

As first, we analyzed DNA methylation at *DAO* and *DDO* promoters in U251 MP cells (Supplementary Figure 1a). Both loci displayed intermediate levels of DNA methylation (49.3% at *DAO* and 27.7% at *DDO*) (Fig. 2A) with DNA methylation levels not equally distributed across the analyzed CpG sites, indicating substantial cell-to-cell heterogeneity within the cell population (Fig. 2B). Epiallele-based analyses further revealed a complex and heterogeneous epiallele composition at both loci, characterized by the coexistence of multiple epiallele classes (Fig. 2C) and a wide diversity of epiallele profiles (Fig. 2D). At *DAO* promoter, the distribution of epiallele classes resembled a gaussian curve, suggesting a high degree of heterogeneity in epiallele composition (Fig. 2C). For *DDO* gene, the 1-Meth and the 2-Meth classes were found as the most abundant epiallele classes (Fig. 2C). Finally, the epiallele distribution analysis detected most different epiallele profiles, indicating an heterogeneous epialleles composition of the U251 MP for both *DAO* and *DDO* genes (Fig. 2D). Collectively, our data indicate that *DAO* and *DDO* represent suitable model loci for

capturing DNA methylation heterogeneity and for investigating epiallele dynamics during cell propagation.

Analysis of epiallele distributions at the *DAO* and *DDO* promoters in putative single-cell derived clones

We next investigated how DNA methylation and epiallele organization at *DAO* and *DDO* loci evolve during single-cell clonal expansion. DNA methylation at the *DAO* and *DDO* promoters was initially evaluated by analyzing both average methylation levels (Fig. 3A, G) and methylation at individual CpG sites (Fig. 3B, H). Both measures were highly consistent among MP, Pool, and all clonal populations, indicating the overall stability of DNA methylation levels during cell propagation (Fig. 3A, B, G, H). We then delved into epiallele classes and distribution analyses (Fig. 3C, I). All the analyzed clones exhibited multiple epiallele classes at both *DAO* and *DDO* loci, suggesting that epiallele classes diversity was reestablished (Fig. 3C, I). No significant differences (ANOVA followed by Tukey's post hoc test; Fig. 3C, I) in each epiallele class were detected among the MP, Pool and clones (Fig. 3C, I). Consistently, Jensen-Shannon distance analysis revealed no significant changes in overall epiallele classes distributions among clones compared to MP for either *DAO* or *DDO* (Fig. 3D, J). We next tracked epiallele profiles by performing Principal Component Analysis (PCA) (Fig. 3E, K). Principal components 1 and 2 explained only a small fraction of the total variance, suggesting that, the overall

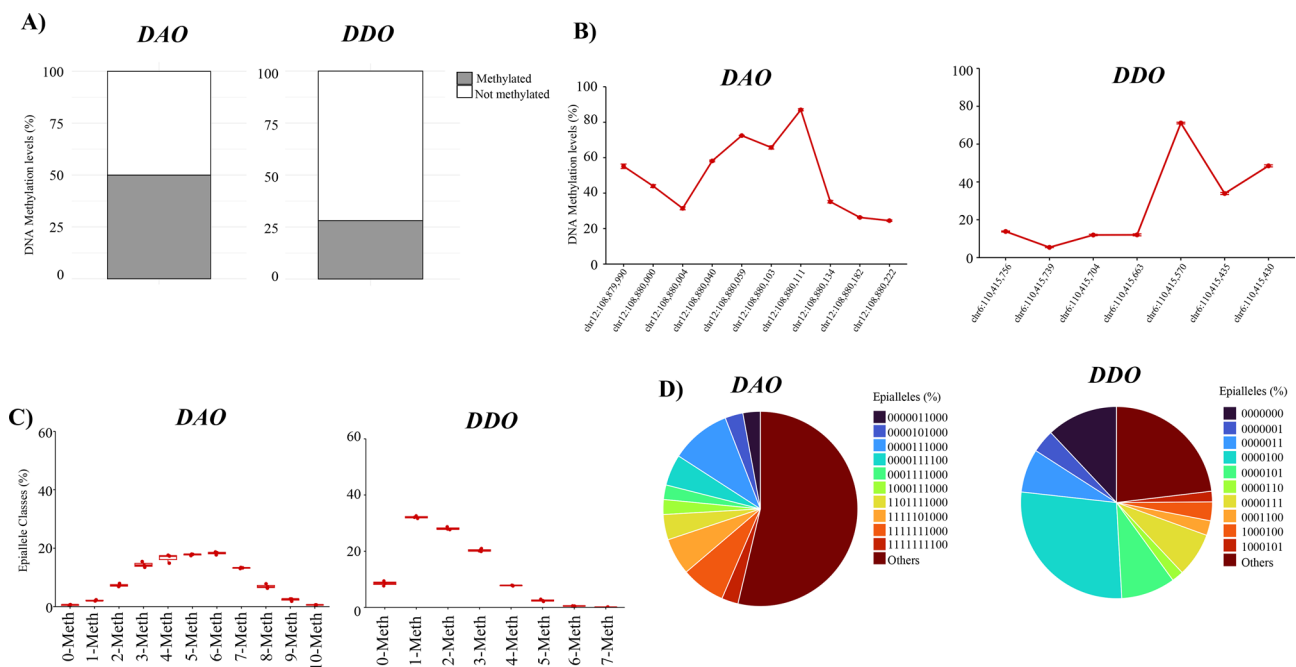


Fig. 2 DNA methylation profiles at *DAO* and *DDO* promoters in the U251 cell lines. **A**) Bar plot representing the average methylation levels at the *DAO* (left) and *DDO* (right) promoters in the U251 MP ($n=3$). The average of methylated (in gray) and unmethylated (in white) molecules are shown. **B**) line graphs representing the DNA methylation at single CpG level of the U251 MP ($n=3$) at the *DAO* (left) and *DDO* (right) promoters. **C**) box plots showing the epiallele classes distribution at the *DAO* (left) and *DDO* (right) promoters in the U251 MP ($n=3$). **D**) pie chart showing, in terms of percentage, the most represented epiallele of the U251 MP at the *DAO* (left) and *DDO* (right) promoters

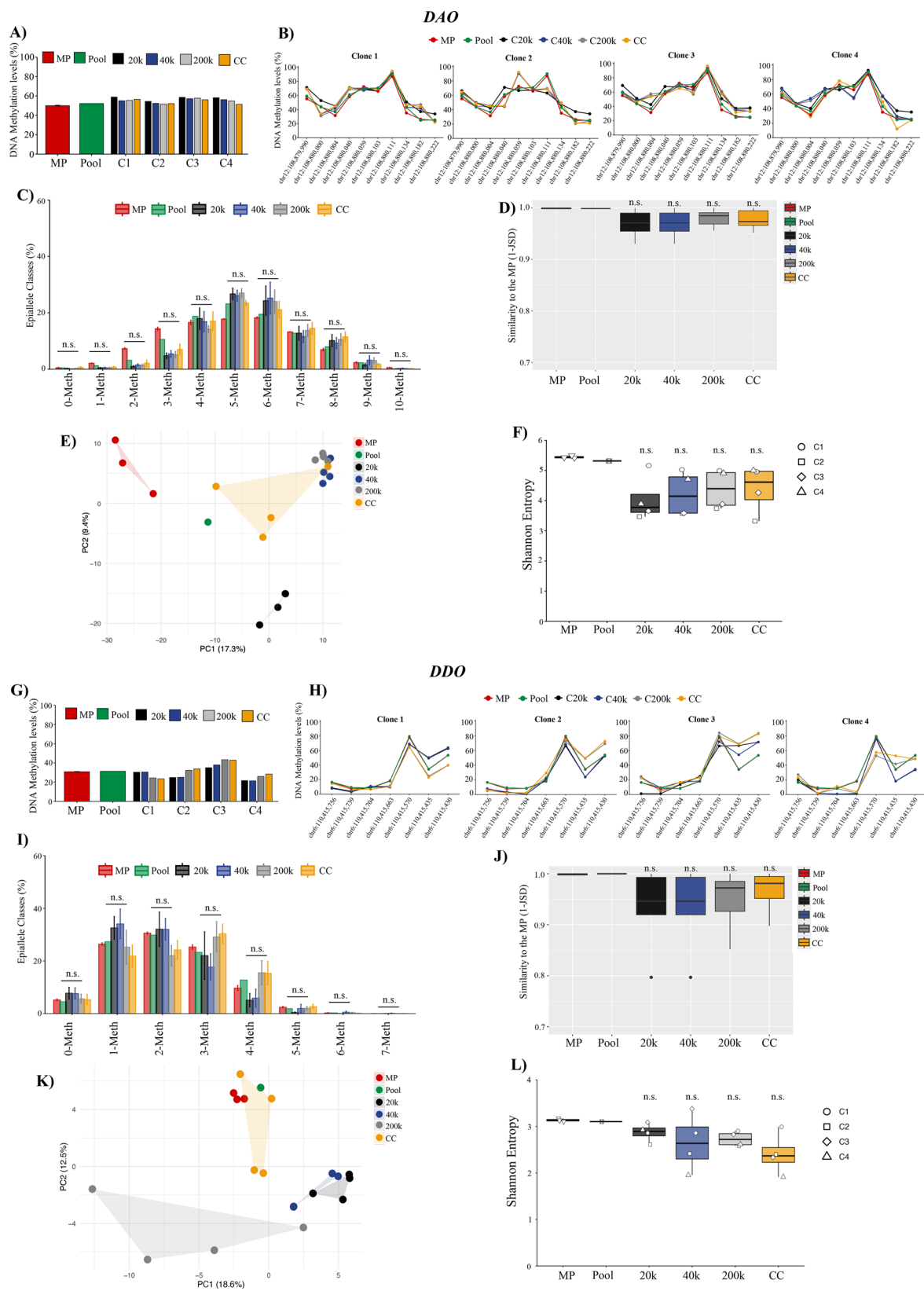


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 *DAO* and *DDO* DNA methylation analysis at putative single-cell derived clones. **A, G**) DNA average methylation levels at **A**) *DAO* and **G**) *DDO* promoters in all the experimental conditions is shown (MP = red; Pool = green; C20k = black; C40k = blue; C200k = gray; CC = orange). **B, H**) single-CpG DNA methylation levels at **B**) *DAO* and **H**) *DDO* promoters are shown for each clone. **C, I**) epiallele classes (%) considering all the epialleles with the same number of methylated CpG sites for **C**) *DAO* and **I**) *DDO* promoters are shown for each experimental condition. Statistical analyses were conducted by using one-way ANOVA followed by Tukey's HSD post-hoc test for multiple pairwise comparisons among experimental conditions (n.s. = non-significant). **D, J**) comparison of the epiallele classes profiles using pairwise Jensen–Shannon divergence, expressed as $1 - \text{JSD}$ for **D**) *DAO* and **J**) *DDO* promoters. Similarity distributions within each experimental condition (MP, Pool, C20k, C40k, C200k, CC) are shown. Pairwise comparisons against MP were performed using Wilcoxon tests with BH-adjusted p -values (n.s. = non-significant). **E, K**) Principal component analysis (PCA) plot of **E**) *DAO* and **K**) *DDO* epiallele profiles in each experimental sample. Each point represents individual samples colored by experimental group (MP = red; Pool = green; C20k = black; C40k = blue; C200k = gray; CC = orange). Shaded polygons indicate the convex hull of each group. Percentages on the axes represent the variance explained by PC1 and PC2. **F, L**) boxplots representing the distribution of Shannon entropy considering the epiallele abundance profiles at **F**) *DAO* and **L**) *DDO* promoters for each sample within experimental groups, with colors indicating group identity (MP = red; Pool = green; C20k = black; C40k = blue; C200k = gray; CC = orange). Each point represents an individual sample, with the symbol shape identifying the clone (C1 = circle, C2 = square, C3 = diamond, C4 = triangle) (n.s. = non-significant). All experimental points were compared to the MP group using Wilcoxon rank-sum tests with benjamini–Hochberg correction. All statistical analyses were conducted by comparing MP ($n = 3$) and clones ($n = 4$ at each experimental time points)

epiallele profiles remained largely conserved among the clones and the MP at both the *DAO* and *DDO* promoters. However, for the *DAO* gene, we observed a trend for C20k to form a cluster that was separated from both the MP and the other clones. We also observed a trend for C40k, C200k, and CC to form a cluster that partially diverged from the MP and the Pool (Fig. 3E). A similar pattern was observed for *DDO*: C20k and C40k tended to cluster together; C200k formed a distinct cluster; and CC clustered very close to MP and the pooled sample (Fig. 3K). Finally, we quantified epiallele complexity by calculating Shannon entropy (Fig. 3F, L). Consistent with the PCA results, the analysis of epiallele Shannon entropy did not reveal statistically significant differences among samples. Nevertheless, MP and Pool showed the highest entropy values compared to the clones for both the *DAO* and *DDO* genes (Fig. 3F, L), suggesting a slight reduction in epigenetic complexity following clonal expansion. To further quantify the similarity between the epiallele composition of MP and the clones at the *DAO* and *DDO* promoters, we performed Pearson correlation tests based on epiallele frequencies. All clones showed statistically significant correlations with the MP (Supplementary Fig. S1), indicating that the observed similarity is highly unlikely to occur by chance. We also performed an additional bisulfite PCR sequencing experiments incorporating unique molecular identifiers (UMIs) to test eventual PCR bias for the epiallele analyses at *DAO* and *DDO* promoters. Analysis of UMI-tagged reads showed a very low duplication rate (<1%) indicating that PCR duplicates have a negligible impact on our data and supporting the validity of interpreting sequencing reads as representative of independent DNA molecules (Supplementary Figure S2A). Importantly, we observed highly concordant epiallele distributions between previous experiment without UMIs and the experiment after UMI-based deduplication (Supplementary Figure S2B–E). Moreover, to test whether the observed behavior was cell line-specific, we repeated the same analyses with an additional non-tumoral cell line, ARPE-19, and observed a comparable pattern. In

this context as well, all analyzed clones displayed restoration of the full spectrum of epiallele classes at both the *DAO* and *DDO* loci, consistent with the re-establishment of heterogeneous epiallele profiles during clonal expansion (Supplementary Figure 3). Collectively, these results suggest that heterogeneous epiallele repertoires are progressively re-established within the expanding population. Despite originating from a putative single cell, all clones progressively maintain the tendency of each CpG site to be methylated or unmethylated, similar to what was observed in mixed populations.

To investigate whether clone-specific differences in epiallele frequencies might be influenced by culture conditions during single-cell clonal expansion (particularly cell density), we performed a control experiment in which cells were maintained at a comparable density throughout the expansion. A population of ~40,000 MP cells was serially expanded through seven consecutive mitotic doublings, corresponding to the divisions needed for 40,000 cells to reach confluence, while maintaining a constant population size of ~40,000 cells. At each doubling, half of the cells were collected for methylation analyses, and the remaining half was further expanded (see Supplementary Fig. 3A for the experimental design). Under these conditions, no differences in epiallele classes distribution or in epiallele distribution entropy were observed at both *DAO* or *DDO* promoters (Supplementary Fig. 3B, C), suggesting that epiallele diversity, variability, and complexity remain stable during mitotic divisions. Even though this control experiment does not fully separate founder effects and/or stochastic methylation changes, our results suggest that the subtle differences observed between MP and putative single-cell-derived clones may reflect the single-cell origins and the isolated culture conditions, which can shift the relative frequencies of epialleles even though the full epiallele repertoire is restored and the hierarchical DNA methylation distribution at single CpG sites is maintained.

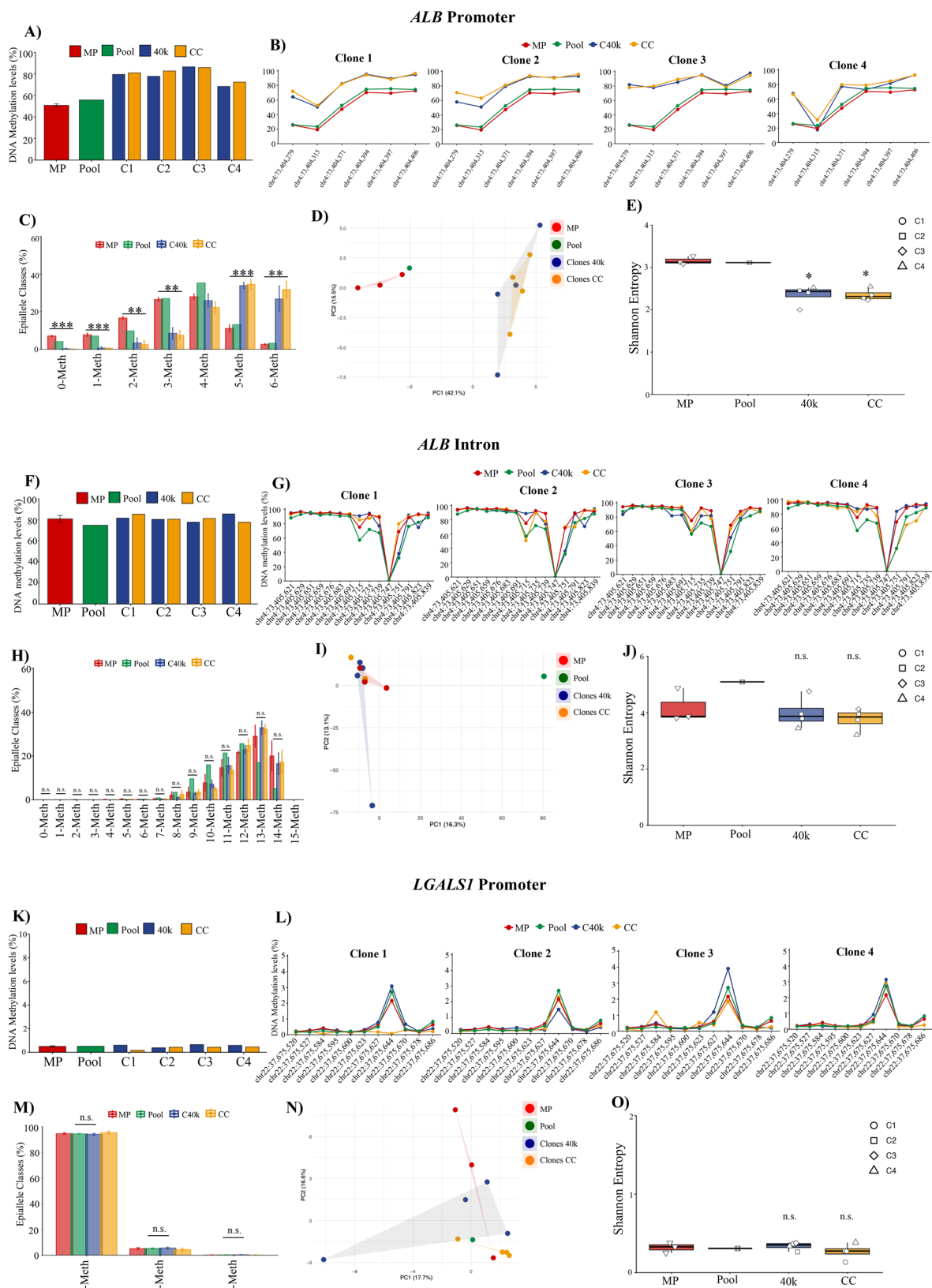


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 DNA methylation profiles at *ALB* promoter, *ALB* intron and *LGALS1* promoter in putative single-cell derived clones. **A, F, K**) DNA average methylation levels at **A**) *ALB* promoter, **F**) *ALB* intron and **K**) *LGALS1* promoter across the experimental conditions is shown (MP = red; Pool = green; C40k = blue; CC = orange). **B, G, L**) single-CpG DNA methylation profiles at **B**) *ALB* promoter, **G**) *ALB* intron and **L**) *LGALS1* promoter for each clone. **C, H, M**) distribution of epiallele classes (%) considering all the epialleles with the same number of methylated CpG sites for **C**) *ALB* promoter, **H**) *ALB* intron and **M**) *LGALS1* promoter are shown for each experimental condition. Statistical analyses were conducted by using one-way ANOVA followed by Tukey's HSD post-hoc test for multiple pairwise comparisons among experimental conditions (n.S. = non-significant; **p*-value < 0.05; ***p*-value 0.01; ****p*-value < 0.001). **D, I, N**) Principal component analysis (PCA) plot based on epiallele profiles distribution at **D**) *ALB* promoter, **I**) *ALB* intron and **N**) *LGALS1* in all experimental samples. Each sample is shown as colored point (MP = red; Pool = green; C40k = blue; CC = orange). Shaded polygons indicate the convex hull of each group. Percentages on the axes represent the variance explained by PC1 and PC2. **E, J, O**) boxplots showing the Shannon entropy values calculated on epiallele abundance profiles in **E**) *ALB* promoter, **J**) *ALB* intron and **O**) *LGALS1* within each experimental group. Different colors indicate groups (MP = red; Pool = green; C40k = blue; CC = orange), while point shapes denote clones (C1 = circle; C2 = square; C3 = diamond; C4 = triangle). Statistical comparisons versus MP were performed using Wilcoxon rank-sum tests with BH correction (n.S. = non-significant; **p*-value < 0.05). All statistical analyses were conducted by comparing MP (*n* = 3) and clones (*n* = 4 at each experimental time points)

Context-dependent epiallele trajectories across selected functional genomic regions

Next, we extended the DNA methylation analyses to additional selected functional genomic regions including promoters, intronic, repetitive, and regulatory regions, using the same analytical framework. Based on the observed concordance between the 20k and 40k clones, as well as between the 200k clones and the CC samples, at the DAO and DDO loci, subsequent analyses were performed using the 40k clones and the CC samples. First, we analyzed *ALB* and *LGALS1* promoter regions. At the *ALB* promoter, all clones showed increased average DNA methylation compared to MP and Pool samples, at both 40k cell numbers and confluence (Fig. 4A). This increase was consistently reflected at the single CpG methylation level, even though the relative methylation hierarchy among CpGs was largely preserved. (Fig. 4B). Accordingly, the distribution of epiallele classes in clones significantly differed from MP and Pool samples, except for the 4-Meth class (ANOVA with Tukey's post hoc test; Fig. 4C). Epiallele-based PCA showed a clear separation between MP/Pool samples and all clones, at both 40k cell number and confluence, indicating a coordinated shift in epiallele architecture associated with clonal expansion (Fig. 4D). Consistently, Shannon entropy analysis revealed a significant reduction in epiallele complexity in clonal populations compared to MP and Pool samples (Fig. 4E). This may reflect a contraction of epiallele heterogeneity driven by the enrichment of highly methylated epialleles. Despite the global increase in DNA methylation observed in all clones, Pearson correlation coefficient revealed significant correlations between the epiallele structures of three out of four clones and the MP (Supplementary Figure 1). This may indicate that even in the presence of a coordinated increase in DNA methylation, epiallele reorganization does not occur randomly, but rather follows constrained trajectories that preserve similarity to the original epiallele architecture. Consistent with observations at DAO and DDO, the epiallele spectrum was re-established following clonal growth, suggesting that even in the presence of global methylation changes, epiallele diversity is restored and reorganized.

To assess whether the differences observed at the *ALB* promoter could be influenced by clonal cell density, we performed the control experiment keeping cells at a comparable cell density throughout the expansion. We found no differences in epiallele classes or epiallele distribution entropy at the *ALB* promoter compared to the MP (Supplementary Fig. 3D, E), suggesting that the observed changes in DNA methylation in putative single-cell-derived clones may be related to the cell density effect. To determine whether the average methylation differences were confined to the *ALB* promoter or extended across the gene body, we analyzed an intronic region of the *ALB* gene. This region displayed high and highly conserved DNA methylation across all samples, at both the average level (Fig. 4F) and single CpG resolution (Fig. 4G). Epiallele class composition was stable across conditions, with no significant differences between MP and clones (Fig. 4H). Accordingly, PCA showed no sample separation, and Shannon entropy analysis revealed no significant changes in epiallele complexity (Fig. 4I, J). Additionally, Pearson correlation analysis revealed a statistically significant similarity between the epiallele profiles of the *ALB* intronic region of all clones and MP (Supplementary Fig. 1). These data suggest that the intronic region of *ALB* gene recapitulates the epiallele dynamics observed at DAO and DDO, indicating that the cell density-associated effect is likely restricted to the *ALB* promoter.

At the *LGALS1* promoter, we found low DNA methylation levels across all analyzed samples (Fig. 4K) and single CpG methylation (Fig. 4L) trend were nearly similar among samples. Consistent with this low methylation state, epiallele class analysis revealed the presence of only a limited subset of methylation classes, with only three (out of eleven) epiallele classes detected across all samples, that were conserved in all analyzed samples (Fig. 4M). Accordingly, PCA did not reveal any separation between MP and clones (Fig. 4N), and Shannon entropy analysis showed no significant differences in epiallele complexity among samples (Fig. 4O). Moreover, a significant Pearson correlation between clones and MP epiallele profiles was found (Supplementary Fig. 1). Thus, the *LGALS1* promoter recapitulate the epiallele dynamics

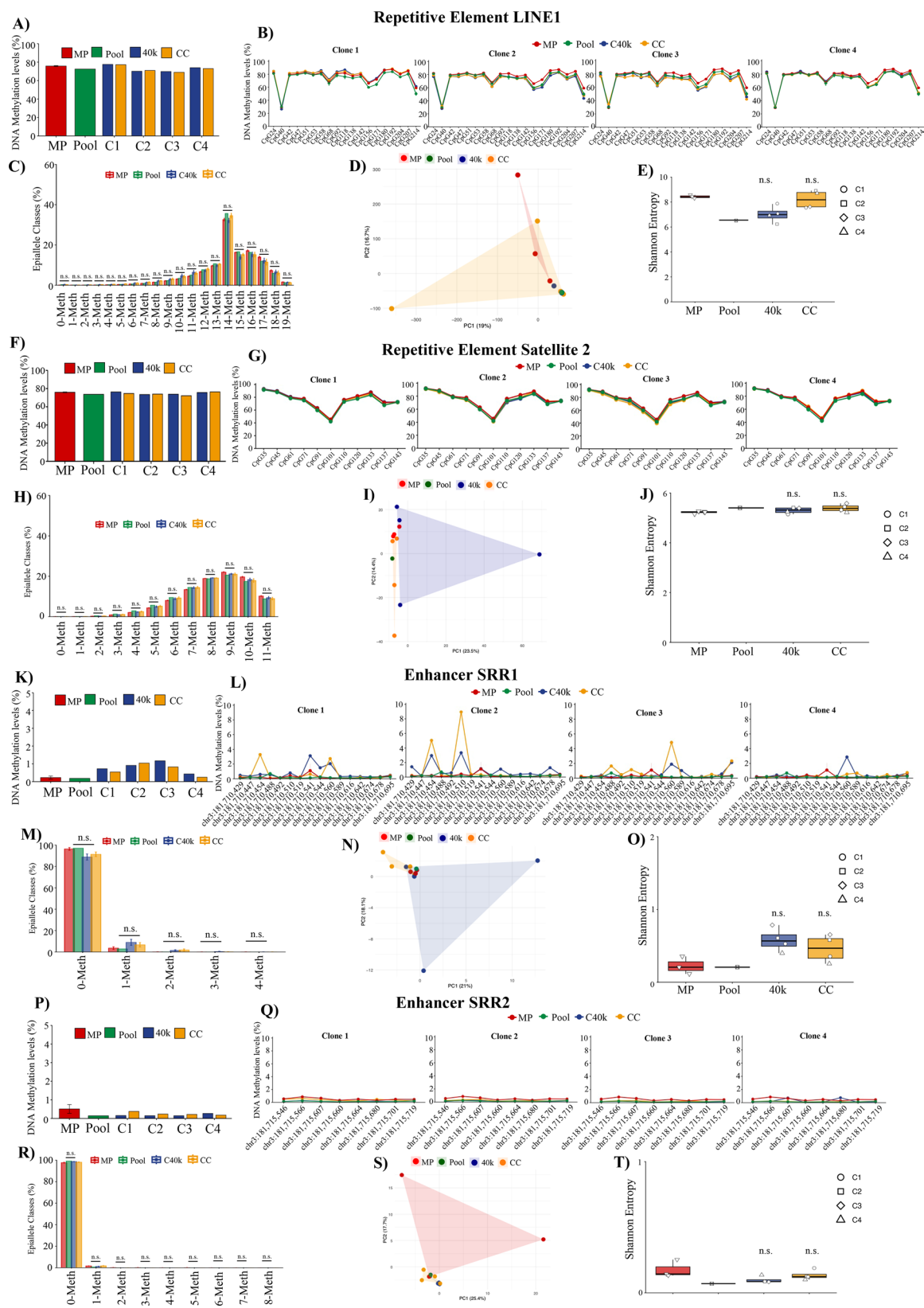


Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 DNA methylation profiles at repetitive elements, enhancers and late-replicating regions in putative single-cell derived clones. **A, F, K, P**) DNA average methylation levels at **A**) *LINE1* and **F**) *Sat2* repetitive elements and **K**) *SRR1* and **P**) *SRR2* enhancers and across the experimental conditions is shown (MP = red; Pool = green; C40k = blue; CC = orange). **B, G, L, Q**) single-CpG DNA methylation profiles at *LINE1* (**B**) and *Sat2* (**G**) repetitive element and at *SRR1* (**L**) and *SRR2* (**Q**) enhancers in MP, Pool and in each clone at 40,000 cells (40k) and confluence (CC). **C, H, M, R**) distribution of epiallele classes (%) considering all the epialleles with the same number of methylated CpG sites for *LINE1* (**C**) and *Sat2* (**H**) repetitive elements and *SRR1* (**M**) and *SRR2* (**R**) enhancers in each experimental condition. Statistical analyses were conducted by using one-way ANOVA followed by Tukey's HSD post-hoc test for multiple pairwise comparisons among experimental conditions (n.s. = non-significant). **D, I, N, S**) Principal component analysis (PCA) plot based on epialleles abundance of **D**) *LINE1* and **I**) *Sat2* repetitive elements and **N**) *SRR1* and **S**) *SRR2* enhancers in all experimental samples. Each sample is shown as colored point (MP = red; Pool = green; C40k = blue; CC = orange). Shaded polygons indicate the convex hull of each group. Percentages on the axes represent the variance explained by PC1 and PC2. **E, J, O, T**) boxplots showing the Shannon entropy values of epiallele abundance profiles in **E**) *LINE1* and **J**) *Sat2* repetitive elements in **O**) *SRR1* and **T**) *SRR2* enhancers within each experimental group. Different colors indicate groups (MP = red; Pool = green; C20k = black; C40k = blue; C200k = gray; CC = orange), while point shapes denote clones (C1 circle; C2 square; C3 diamond; C4 triangle). Statistical comparisons versus MP were performed using Wilcoxon rank-sum tests with benjamini–Hochberg correction (n.s. = non-significant)

with a strong conservation of epiallele organization during clonal expansion.

Next, we analyzed repetitive elements and regulatory regions, including *LINE1*, *Sat2*, and two *SOX2* enhancers, named *SRR1* and *SRR2*. Due to the highly repetitive nature of *LINE1* and *Sat2*, the analyses reported here are not attributed to individual genomic loci, but rather represent the combined methylation signal of multiple repeat copies amplified by the primer sets. The *LINE1* and *Sat2* loci displayed uniformly high and conserved DNA methylation across the MP, Pool, and all clones (Fig. 5A, F), with similar single CpG methylation levels (Fig. 5B, G), stable distributions of predominantly highly methylated epiallele classes (Fig. 5C, H), and no significant changes in epiallele distribution (Fig. 5D, I) or complexity (Fig. 5E, L). Thus, the overall epiallele spectrum was fully re-established after clonal expansion at these repetitive genomic loci. Consistently, Pearson correlation analyses showed significant similarity between clones and MP (Supplementary Fig. 1). We next analyzed the *SOX2* regulatory enhancers *SRR1* and *SRR2*. Both regions exhibited low DNA methylation levels, which were homogeneous across MP and clones, both at the average level (Fig. 5K, P) and at single CpG resolution (Fig. 5L, Q). Epiallele class composition was similarly conserved, with the unmethylated (0-Meth) class as the most abundant across all samples (Fig. 5M, R). Consistently, neither epiallele distribution (Fig. 5N, S) nor Shannon entropy analysis (Fig. 5O, T) revealed significant differences between MP and clones. Also for the enhancers, epialleles of the clones significantly correlated to the epialleles of the MP (Supplementary Fig. 1). Together, these results suggest that DNA methylation patterns do not segregate clonally, even at distal regulatory elements such as *SOX2* enhancers, supporting the presence of a regulated, dynamically maintained epiallele architecture.

Epiallele dynamics in late-replicating regions containing solo WCGW sites

Given the dynamic yet stochastic distribution of epialleles observed across promoters, intragenic regions, repetitive elements, and enhancers, we extended our

analysis to late-replicating genomic regions enriched for WCGW CpG contexts (Fig. 6). These regions are known to exhibit reduced fidelity of DNA methylation maintenance during replication, likely due to limited activity of the DNMT1–UHRF1 machinery in late S phase and have been associated with increased epigenetic variability [20–22]. By this way, we aimed to assess whether the well-orchestrated methylation patterns observed in other genomic contexts are preserved under conditions of intrinsically reduced methylation maintenance fidelity. To this end, we analyzed a region within the *ADAMTS2* gene containing three CpG sites, all embedded in solo-WCGW motifs. In this region, two of the three CpG sites consistently displayed low methylation levels across MP and Pool samples (Fig. 6A, B), whereas the remaining site showed higher methylation and was also the site exhibiting the largest variability across clones (Fig. 6B). Notably, this CpG largely drove the differences observed in both global and single-site average methylation levels between MP and clones (Fig. 6B). No significant differences in epiallele classes distributions were found, likely due to the limited number of CpG sites (three), which constrains the number of possible epiallele combinations (only nine) (Fig. 6C). Nevertheless, changes in epiallele composition between MP and individual clones were suggested by PCA (Fig. 6D), and although not statistically significant, Shannon entropy analysis indicated a trend toward reduced epiallele heterogeneity, particularly in Clone 4, which is characterized by a marked loss of methylation at the first CpG site (Fig. 6E). Consistently, Pearson correlation comparing epiallele distance matrices revealed a significant correlation between MP and individual clones (Supplementary Fig. 1), indicating that despite clone-specific shifts in epiallele composition, indicating that the epiallele frequency profiles were not independent of those observed in the parental population. We therefore extended the analysis to a second downstream region of *ADAMTS2* containing one solo-WCGW CpG site. No differences were observed in average methylation levels across samples (Fig. 6F). At the single-site level, the WCGW CpG site remained remarkably stable, whereas the non-WCGW CpG sites exhibited significant

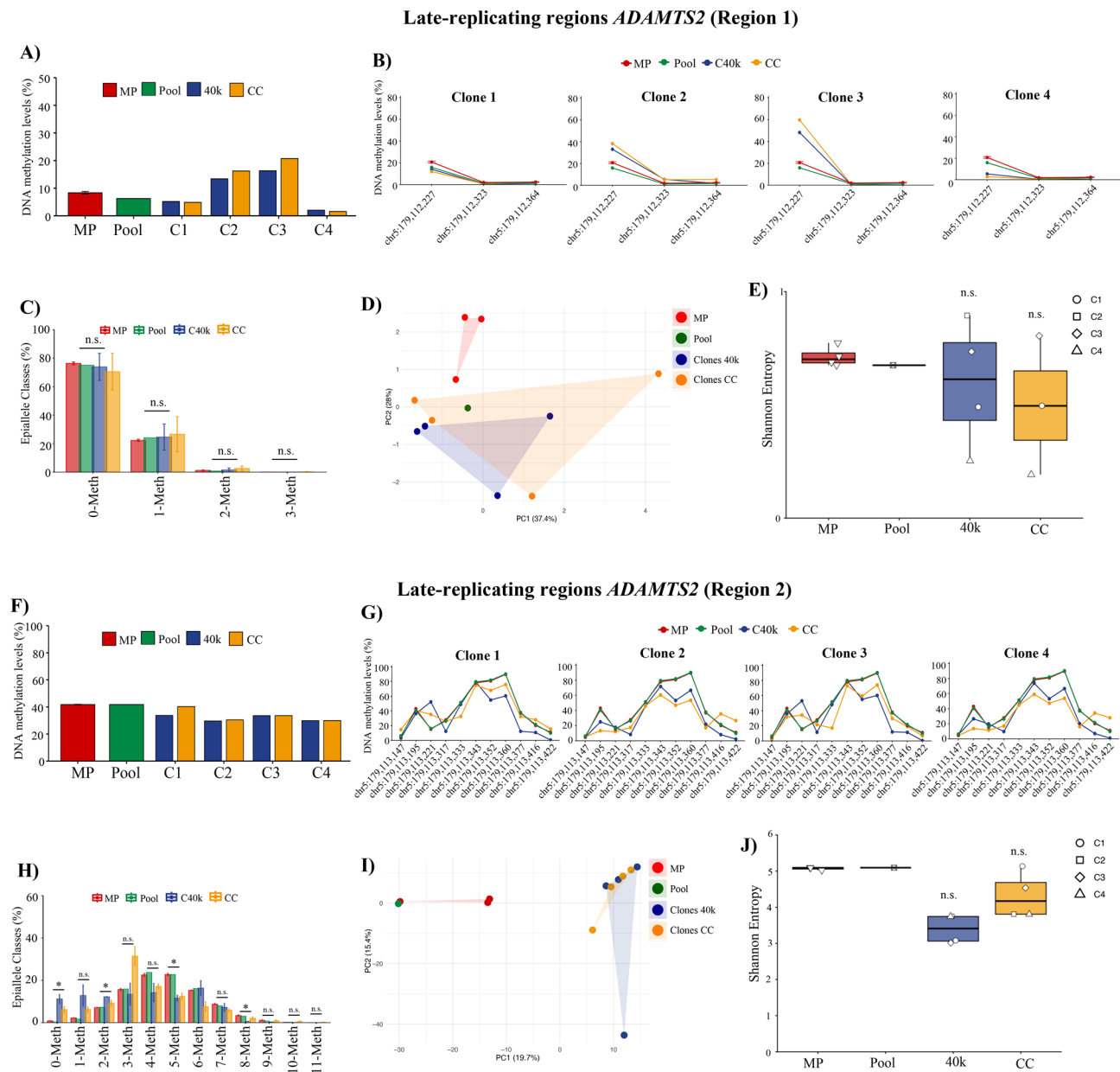


Fig. 6 DNA methylation profiles at late replicating regions in putative single-cell derived clones at “solo-WCGW” context. **A, F**) DNA average methylation values in *ADAMTS2* “WCGW solo” **A**) region 1 and **F**) region 2 among all the experimental conditions (MP=red; Pool=green; C40k=blue; CC=orange) and all the analyzed clones. **B, G**) line graphs representing single-CpG DNA methylation profiles at *ADAMTS2* “WCGW solo” **B**) region 1 and **G**) region 2 under the different experimental conditions, shown separately for each clone. **C, H**) epiallele classes (percentage) for *ADAMTS2* “WCGW solo” **C**) region 1 and **H**) region 2 in each experimental condition. Statistical analyses were conducted by using one-way ANOVA followed by Tukey’s HSD post-hoc test for multiple pairwise comparisons among experimental conditions (n.s. = non-significant; **p*-value < 0.05). **D, I**) Principal component analysis (PCA) plot for *ADAMTS2* “WCGW solo” **D**) region 1 and **I**) region 2. Each sample is shown as colored point (MP=red; Pool=green; C40k=blue; CC=orange). Shaded polygons indicate the convex hull of each group. Percentages on the axes represent the variance explained by PC1 and PC2. **E, J**) Shannon entropy values based on epiallele frequencies at *ADAMTS2* “WCGW solo” **E**) region 1 and **J**) region 2 for the different experimental conditions. Different colors indicate groups (MP=red; Pool=green; C20k=black; C40k=blue; C200k=gray; CC=orange), while point shapes denote clones (C1 circle; C2 square; C3 diamond; C4 triangle). Statistical comparisons versus MP were performed using Wilcoxon rank-sum tests with benjamini–Hochberg correction (n.s. = non-significant).

methylation changes across clones, despite a largely conserved hierarchy of CpG methylation levels across sites (Fig. 6G). In contrast to the first region, epiallele classes analysis revealed significant changes, particularly affecting the 0-Meth, 2-Meth and 5-Meth classes (Fig. 6H).

Consistently, PCA showed a clear separation of MP and Pool samples from the clones, indicating a global shift in epiallele distribution (Fig. 6I). This was further supported by significant differences in Shannon entropy, pointing to altered epiallelic heterogeneity within the

clonal populations (Fig. 6). Nevertheless, Pearson correlation test comparing epiallele frequencies between MP and individual clones revealed significant correlations, indicating that despite changes in epiallele composition and heterogeneity, the epiallele frequency profiles were not independent of those observed in the parental population. (Supplementary Fig. 1). Collectively, these results indicate that late-replicating WCGW-enriched regions are characterized by elevated epiallelic dynamics, with frequent site-specific and combinatorial changes. Although clonal expansion is associated with substantial redistribution of epialleles and altered heterogeneity, the persistence of significant Pearson correlations suggests that epiallele variation arises within a conserved structural framework rather than through a pure stochastic emergence of a new epialleles profiles. These findings support a model in which late-replicating WCGW contexts represent epigenetically permissive regions in which epiallele diversity is continuously regenerated across cell divisions rather than clonally fixed.

Discussion

The hypothesis that the methylation state at each CpG site in different cells is subject to periodic fluctuation emerged from our and other previous studies, showing that DNA methylation is highly polymorphic and may follow a deterministic turnover [7, 23–25]. Pioneering works showed that CpG methylation at certain promoters may fluctuate on a time scale of minutes as part of transcriptional regulatory cycles, suggesting that DNA methylation can participate in highly dynamic regulatory processes rather than representing a static epigenetic mark [26, 27]. Over time, an increasing number of novel tools have been developed to assess the overall epiallele dynamics [28–32]. However, it remains unclear whether, in a single cell at a given locus, the specific epiallele configuration is stably inherited by the daughter cells, resulting in a cell population with the same epiallele, or epiallele diversity may be progressively re-established during cell proliferation, giving rise to daughter cells with all possible combinations of epialleles. In the present work, we sought to investigate this question by using a putative single-cell-derived clonal system and a comprehensive epiallele-based analytical framework applied across multiple functional genomic contexts. Instead of observing the propagation of a single dominant epiallele during clonal expansion, we found that a broad repertoire of epialleles was progressively re-established in the expanding cell populations originating from isolated cells. These observations suggest that epiallele diversity may emerge during clonal expansion rather than being strictly maintained as a single clonal configuration. Importantly, the epiallele dynamics were strongly influenced by the initial methylation state of each locus.

Regions characterized by high or uniformly low average methylation exhibited limited epiallele heterogeneity and strong conservation of epiallele organization, whereas regions with intermediate methylation levels displayed greater epiallele diversity and plasticity. These findings suggest that the average methylation state of a locus represents an important determinant of the range of epiallele configurations that can be generated during clonal expansion. Despite shifts in epiallele frequencies in certain genomic contexts, particularly at promoter regions, a key feature across all regions was the preservation of site-specific methylation tendencies. Individual CpG sites retained a consistent propensity to be methylated or unmethylated, even when their absolute methylation levels or the relative abundance of epialleles changed. Notably, epiallele dynamics was also observed in late-replicating genomic regions enriched for solo-WCGW CpG contexts, which have previously been associated with reduced fidelity of DNA methylation maintenance during replication. Previous studies have suggested that accelerated methylation loss in these regions may result from inefficient maintenance methylation by DNMT1 during late S phase, while transcriptional activity and chromatin-associated features such as H3K36me3 may partially protect CpG sites from progressive methylation erosion [20]. These mechanisms were not directly tested in the present study and we may not exclude that the wide epiallele scenario observed in a mixed cell population may be the result of a stochastic DNA methylation process, potentially arising from an unfaithful copy of the hemimethylated strand of DNA by the DNMT1 [6, 33, 34]. However, these observations provide a potential mechanistic framework for interpreting our results. The persistence of site-specific methylation tendencies across clones may reflect underlying local genomic and chromatin constraints that shape maintenance fidelity and influence epiallele organization. While our experimental design does not directly track methylation inheritance through successive divisions of individual cells, the consistent re-emergence of complex epiallele repertoires across independently expanded clones is compatible with a model in which DNA methylation patterns operate within a dynamic equilibrium. In this scenario, epiallele composition would not be purely random but shaped by flexible yet constrained processes governing methylation maintenance and local genomic context.

Limitations of the study

Although our study provides insights into the organization of epiallele diversity during clonal expansion, several limitations should be considered when interpreting our findings. First, gene expression was not directly assessed in this study, and therefore we cannot establish a functional relationship between epiallele composition and

transcriptional activity at the analyzed loci. Future studies integrating epiallele profiling with transcriptional measurements will be required to determine whether specific epiallele configurations influence gene expression states. Second, our analyses were restricted to a limited number of genomic regions, selected to represent different functional contexts including promoters, regulatory elements, intragenic regions and repetitive sequences. While this targeted approach allowed us to achieve very high sequencing depth and robust characterization of epiallele architectures, the results cannot be assumed to represent genome-wide methylation dynamics. Consequently, the conclusions of this study should be interpreted within the scope of locus-specific analyses, and future genome-wide investigations will be necessary to determine whether the patterns observed here extend to broader genomic contexts.

Methods

Cell culture and establishment of clones

U251MG glioblastoma cell lines (ATCC, Manassas, VA, USA) were cultured at 37 °C in the presence of 5% CO₂ in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, #32430–027) media containing 10% fetal bovine serum (FBS) (Gibco, #26140079), 5% Penicillin/Streptomycin (10,000 U/mL) (Gibco™ #15140122), 1X MEM Nonessential Amino Acids (100X) (Gibco™ #11140050), 1 mM Sodium Pyruvate (100 mM) (Gibco™ #11360070), and 1X GlutaMAX™ Supplement (100X) (Gibco™ #35050061). Cells were detached using Trypsin-EDTA 1X (EuroClone, #ECB3052D). To generate clones, 8000 cells were counted and plated into a 150 mm tissue culture dish (Corning, #CLS430599). Cells were thoroughly dissociated by repeated pipetting prior to plating to ensure a putative single-cell suspension. Individual cells were visually identified under the microscope and their growth was monitored over time until colony formation. Colonies selected for further expansion originated from spatially isolated single cells that were tracked during the initial stages of proliferation. After 96 hours, individual groups of cells were picked under microscopic guidance and placed in a 24-well dish (Corning, #CLS3527). Each clone was grown individually, avoiding any contact with other clones. After 2 cell duplications, clones were detached and plated in three 60 mm tissue culture dishes (Corning, #430,166). The cells were grown to 40,000 cells (clone 40,000), to 200,000 cells (clone 200,000) and until the confluence state (confluent clone). After 96 hours, cells designated to grow until the confluence were plated in a 100 mm culture dish to prevent over confluence. To replicate the number of mitoses occurring between clone 40k and clone 200k and between clone 200k and confluent clone, 40,000 cells and 200,000 cells were counted and plated into 60 mm tissue culture dish and 100 mm tissue

culture dish respectively. After each mitotic division, cells were split to maintain the same number of cells per plate. The experiments were conducted with three biological replicates ($n=3$) of the MP, while four independently derived clones were analyzed at each time point ($n=4$). For the statistical analyses, the clone data were pooled to represent the clonal population.

DNA extraction

DNA was extracted from cell pellets using DNeasy Blood & Tissue Kit (QIAGEN, #69506) following the manufacturer's protocol. The quality of each DNA sample was evaluated at NanoDrop One (Thermo Scientific) for absorbance 260/280 ratio and absorbance 260/230 ratio. Quantification of the DNA was performed using Qubit 2.0 Fluorometer with the dsDNA broad range assay kit (Invitrogen, Q32850).

Bisulfite treatment and amplicon library preparation

Bisulfite treatment was performed using EZ DNA Methylation Kit (Zymo Research, D5002) following the manufacturer's protocol. Briefly, genomic DNA from cell pellets (1000 ng) was converted with "C/T conversion reagent" and eluted in 25 µL of H₂O. Two amplification steps were performed to obtain an amplicon library that was sequenced with Illumina MiSeq sequencer. All the primers and the relative thermal protocols used for the amplification of all the genomic regions are listed in Supplementary Table 1. First amplification reactions were performed by using FastStart High Fidelity PCR System (Roche, #03553400001). The reactions were performed in 30 µL total volume: 3 µL of buffer 10x, 0.9 µL of 5 mM forward and reverse primers, 3.6 µL of 25 mM MgCl₂, 0.6 µL of 10 mM dNTPs mix, 0.3 µL of Taq, 5 µL of bisulfite-treated template DNA, and H₂O up to 30 µL. To remove primers dimers AMPure XP purification magnetic beads (Beckman-Coulter, #A63881) were used following the manufacturer's protocol with a ratio between AMPure Beads volume and PCR products volume of 0.8. Purified PCR products were run on 1.5% agarose gel (Sigma-Aldrich, #A9539) with a 100 bp ladder (New England BioLabs, #N3231L). Second PCR steps were performed using KAPA Taq HotStart (Sigma-Aldrich, #KK1508) in 30 µL total volume: 15 µL of 2x master mix, 1.2 µL forward and reverse primers from Nextera XT Index Kit v2 (Illumina, #15052163), 5 µL first PCR product, and H₂O up to 30 µL. Thermocycler settings were: 95 °C for 2 min followed by 8 cycles at 95 °C 20s, 55 °C for 30 s, 72 °C for 40 s, followed by a final extension step at 72 °C for 5 min. All amplicons were quantified using Qubit 2.0 Fluorometer (Invitrogen, #Q32866). Amplicons were pooled at equimolar ratio and then diluted to final concentration of 8 picomolar. Phix Control libraries (Illumina, #15017666) were combined with normalized libraries [15% (v/v)]

to increase diversity of base calling during sequencing. Amplicons library was subjected to sequencing using v2 reagents kits on Illumina MiSeq system (Illumina, # SY-410–1003). Paired-end sequencing was performed in 251 cycles per read (251×2). An average of 50,000 reads/sample were used for bioinformatic analysis.

Sequence analysis

As first, sequencing quality was assessed at the sample level using FastQC (Supplementary Table 2). Across all the library, an average of 91.6% of bases showed a Phred quality score ≥ 30 . For UMI-based amplicon sequencing, the data were processed using UMI-tools to extract 8-bp UMIs from both ends of paired-end reads. Primer sequences were removed with Cutadapt. UMI sequences were subsequently trimmed, and resulting amplicon sequences were analyzed. Sequence heterogeneity was assessed by quantifying unique sequences and their frequencies using Biostrings in RStudio. Duplication rates were calculated based on total reads and unique UMI counts (Supplementary Figure S4A). Paired-end reads obtained from Illumina Miseq platform were assembled using PEAR tool [35] with a minimum of 40 overlapping residues as threshold. After, FASTQ assembled reads were converted in a FASTA format using PRINSEQ tool [36]. The obtained sequences were analyzed using ampliMethProfiler pipeline software (<https://sourceforge.net/projects/amplimethprofiler>). Bisulfite conversion efficiency was evaluated by analyzing the non-CpG cytosines within the sequenced region, which are expected to be fully converted after bisulfite treatment. Across all samples, the conversion rate exceeded 99%, indicating efficient bisulfite conversion (Supplementary Table 2). Quality filtered FASTA files have been produced by AmpliMethProfiler for each sample. AmpliMethProfiler also generates a tabular format file (BIOM format) containing the amount of epialleles in all samples. The BIOM table was normalized to the same number of sequences per sample using a rarefaction procedure implemented in QIIME [37].

Statistical analyses

The similarity between Epiallele Classes profiles between two conditions was quantified as 1- JSD, where JSD denotes the Jensen Shannon distance. This metric measures the degree of dissimilarity between two discrete distributions P_1 and P_2 , and is defined as:

$$\sqrt{\frac{D(P_1, \bar{P}) + D(P_2, \bar{P})}{2}}$$

where $\bar{P}$ represents the average distribution of two methylation classes profiles, and $D(P_1, \bar{P})$ and $D(P_2, \bar{P})$

denote the Kullback-Leibler (KL) divergences between each profile and the average distribution, respectively, as previously described [38]. For epiallele analysis, we computed the Shannon index from the BIOM table using QIIME, in order to quantify DNA methylation heterogeneity. To evaluate the relatedness between epialleles whose frequency increases in clones compared to MP, we selected those with a frequency difference above the 95th percentile for DAO and above the 90th percentile for DDO. We then computed the average Hamming distance between these epialleles using the *stringdistmatrix* function from the *stringdist* R package. To assess whether the observed distances were lower than expected by chance, we compared them to the distribution of distances obtained by randomly sampling the same number of epialleles from the MP 10,000 times. Epiallele entropy was quantified for each sample using the Shannon entropy index implemented in the *vegan* R package [39]. Differences in Shannon entropy between the MP group and the other groups were assessed using two-sided Wilcoxon rank-sum tests. Resulting *p-values* were corrected for multiple comparisons using the Benjamini–Hochberg false discovery rate (FDR) method. Pairwise Euclidean distance matrices were calculated from epiallelic frequency profiles using the *vegdist* function from the *vegan* package. A reference distance matrix was generated for MP samples, and separate distance matrices were computed for each clone group (C1–C4). Similarity between the MP distance matrix and each clone-specific distance matrix was assessed using Pearson correlation test. For each comparison, the Pearson correlation coefficient (*r*) and associated *p-value* were reported. Data visualization was performed using *ggplot2* (H. [40]). All statistical analyses were performed in R version 4.5.2.

Supplementary information

The online version contains supplementary material available at <https://doi.org/10.1186/s13062-026-00844-x>.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements

The authors thank the facilities of CEINGE-Advanced Biotechnologies “Franco Salvatore” for the support. The authors thank Dr. Mariachiara De Luca for technical and administrative support.

Author contributions

Davide Costabile: Conceptualization, Methodology, Formal analysis, Validation, Investigation, Writing - Original Draft. Maddalena Russo: Investigation, Formal analysis. Rosa Della Monica: Writing - Review & Editing, Validation. Michela Buonaiuto: Writing - Review & Editing, Validation. Giulia De Riso: Formal analysis, Writing - Review & Editing. Marta Sabbarese: Writing - Review & Editing. Giuseppe Catapano: Writing - Review & Editing. Roberta Visconti: Writing - Review & Editing. Sergio Cocozza: Writing - Review & Editing. Lorenzo Chiariotti: Conceptualization, Writing - Original Draft, Funding acquisition.

Mariella Cuomo: Conceptualization, Supervision, Methodology, Formal analysis, Writing - Original Draft.

Data availability

DNA methylation raw data generated in this study are available on European Nucleotide Archive (ENA) database with the accession number PRJEB108340

Declarations

Ethical approval

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 16 March 2026 / Accepted: 15 May 2026

Published online: 02 June 2026

References

1. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* 2002;16(1):6–21. <https://doi.org/10.1101/gad.947102>.
2. Lande-Diner L, Zhang J, Ben-Porath I, Amariglio N, Keshet I, Hecht M, et al. Role of DNA methylation in stable gene repression. *J Biol Chem.* 2007;282(16):12194–200. <https://doi.org/10.1074/jbc.M607838200>.
3. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet.* 2013;14(3):204–20. <https://doi.org/10.1038/nrg3354>.
4. Weber M, Schübeler D. Genomic patterns of DNA methylation: targets and function of an epigenetic mark. *Curr Opin Cell Biol.* 2007;19(3):273–80. <https://doi.org/10.1016/j.ccb.2007.04.011>.
5. Brown SE, Fraga MF, Weaver ICG, Berdasco M, Szyf M. Variations in DNA methylation patterns during the cell cycle of HeLa cells. *Epigenetics.* 2007;2(1):54–65. <https://doi.org/10.1041/epi.2.1.3880>.
6. Landan G, Cohen NM, Mukamel Z, Bar A, Molchadsky A, Brosh R, et al. Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. *Nat Genet.* 2012;44(11):1207–14. <https://doi.org/10.1038/ng.2442>.
7. Rulands S, Lee HJ, Clark SJ, Angermueller C, Smallwood SA, Krueger F, et al. Genome-scale oscillations in DNA methylation during exit from pluripotency. *Cell Syst.* 2018;7(1):63–76.e12. <https://doi.org/10.1016/j.cels.2018.06.012>.
8. Saito D, Suyama M. Linkage disequilibrium analysis of allelic heterogeneity in DNA methylation. *Epigenetics.* 2015;10(12):1093–98. <https://doi.org/10.1080/15592294.2015.1115176>.
9. Xie H, Wang M, de Andrade A, Bonaldo Mde F, Galat V, Arndt K, et al. Genome-wide quantitative assessment of variation in DNA methylation patterns. *Nucleic Acids Res.* 2011;39(10):4099–108. <https://doi.org/10.1093/nar/gkr017>.
10. Affinito O, Palumbo D, Fierro A, Cuomo M, De Riso G, Monticelli A, et al. Nucleotide distance influences co-methylation between nearby CpG sites. *Genomics.* 2020;112(1):144–50. <https://doi.org/10.1016/j.ygeno.2019.05.007>.
11. Cuomo M, Florio E, Della Monica R, Costabile D, Buonaiuto M, Di Risi T, et al. Epigenetic remodelling of Fxyd1 promoters in developing heart and brain tissues. *Sci Rep.* 2022;12(1):6471. <https://doi.org/10.1038/s41598-022-10365-y>.
12. De Riso G, Fiorillo DFG, Fierro A, Cuomo M, Chiariotti L, Miele G, et al. Modeling DNA methylation profiles through a Dynamic equilibrium between methylation and demethylation. *Biomolecules.* 2020;10(9):1271. <https://doi.org/10.3390/biom10091271>.
13. Qu W, Tsukahara T, Nakamura Y, Yurino H, Hashimoto S, Tsuji S, et al. Assessing cell-to-cell DNA methylation variability on individual long reads. *Sci Rep.* 2016;6(1):21317. <https://doi.org/10.1038/srep21317>.
14. Florio E, Keller S, Coretti L, Affinito O, Scala G, Errico F, et al. Tracking the evolution of epialleles during neural differentiation and brain development: D-Aspartate oxidase as a model gene. *Epigenetics.* 2017;12(1):41–54. <https://doi.org/10.1080/15592294.2016.1260211>.
15. Russo G, Tramontano A, Iodice I, Chiariotti L, Pezone A. Epigenome chaos: stochastic and deterministic DNA methylation events drive cancer evolution. *Cancers (Basel).* 2021;13(8):1800. <https://doi.org/10.3390/cancers13081800>.
16. Cuomo M, Borrelli L, Della Monica R, Coretti L, De Riso G, di Durazzo DL, et al. DNA methylation profiles of Tph1A and BDNF in Gut and brain of L. Rhamnosus-treated Zebrafish. *Rhamnosus-Treat Zebrafish. Biomolecules.* 2021;11(2):142. <https://doi.org/10.3390/biom11020142>.
17. Cuomo M, Keller S, Punzo D, Nuzzo T, Affinito O, Coretti L, et al. Selective demethylation of two CpG sites causes postnatal activation of the dao gene and consequent removal of D-serine within the mouse cerebellum. *Clin Epigenet.* 2019;11(1):149. <https://doi.org/10.1186/s13148-019-0732-z>.
18. Affinito O, Scala G, Palumbo D, Florio E, Monticelli A, Miele G, et al. Modeling DNA methylation by analyzing the individual configurations of single molecules. *Epigenetics.* 2016;11(12):881–88. <https://doi.org/10.1080/15592294.2016.1246108>.
19. Keller S, Punzo D, Cuomo M, Affinito O, Coretti L, Sacchi S, et al. DNA methylation landscape of the genes regulating D-serine and D-aspartate metabolism in post-mortem brain from controls and subjects with schizophrenia. *Sci Rep.* 2018;8(1):10163. <https://doi.org/10.1038/s41598-018-28332-x>.
20. Endicott JL, Nolte PA, Shen H, Laird PW. Cell division drives DNA methylation loss in late-replicating domains in primary human cells. *Nat Commun.* 2022;13(1):6659. <https://doi.org/10.1038/s41467-022-34268-8>.
21. Shipony Z, Mukamel Z, Cohen NM, Landan G, Chomsky E, Zeliger SR, et al. Dynamic and static maintenance of epigenetic memory in pluripotent and somatic cells. *Nature.* 2014;513(7516):115–19. <https://doi.org/10.1038/nature13458>.
22. Zhou W, Dinh HQ, Ramjan Z, Weisenberger DJ, Nicolet CM, Shen H, et al. DNA methylation loss in late-replicating domains is linked to mitotic cell division. *Nat Genet.* 2018;50(4):591–602. <https://doi.org/10.1038/s41588-018-0073-4>.
23. Farlik M, Sheffield NC, Nuzzo A, Datlinger P, Schönegger A, Klughammer J, et al. Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. *Cell Rep.* 2015;10(8):1386–97. <https://doi.org/10.1016/j.celrep.2015.02.001>.
24. Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, et al. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci USA.* 2010;107(19):8689–94. <https://doi.org/10.1073/pnas.1002720107>.
25. Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, Peat J, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat Methods.* 2014;11(8):817–20. <https://doi.org/10.1038/nmeth.3035>.
26. Kangaspekka S, Stride B, Métivier R, Polycarpou-Schwarz M, Ibberson D, Carmouche RP, et al. Transient cyclical methylation of promoter DNA. *Nature.* 2008;452(7183):112–15. <https://doi.org/10.1038/nature06640>.
27. Métivier R, Gallais R, Tiffocche C, Le Péron C, Jurkowska RZ, Carmouche RP, et al. Cyclical DNA methylation of a transcriptionally active promoter. *Nature.* 2008;452(7183):45–50. <https://doi.org/10.1038/nature06544>.
28. Li S, Garrett-Bakelman F, Perl AE, Luger SM, Zhang C, To BL, et al. Dynamic evolution of clonal epialleles revealed by methclone. *Genome Biol.* 2014;15(9):472. <https://doi.org/10.1186/s13059-014-0472-5>.
29. Pezone A, Tramontano A, Scala G, Cuomo M, Riccio P, De Nicola S, et al. Tracing and tracking epiallele families in complex DNA populations. *NAR Genomics Bioinf.* 2020;2(4):lqaa096. <https://doi.org/10.1093/nargab/lqaa096>.
30. Sarnataro A, De Riso G, Coccoza S, Pezone A, Majello B, Ament S, et al. A novel workflow for the qualitative analysis of DNA methylation data. *Comput Struct Biotechnol J.* 2022;20:5925–34. <https://doi.org/10.1016/j.csbj.2022.10.027>.
31. Scala G, Affinito O, Palumbo D, Florio E, Monticelli A, Miele G, et al. ampMethProfiler: a pipeline for the analysis of CpG methylation profiles of targeted deep bisulfite sequenced amplicons. *BMC Bioinf.* 2016;17(1):484. <https://doi.org/10.1186/s12859-016-1380-3>.
32. Wong NC, Pope BJ, Candiloro IL, Korbie D, Trau M, Wong SQ, et al. MethPat: a tool for the analysis and visualisation of complex methylation patterns obtained by massively parallel sequencing. *BMC Bioinf.* 2016;17(1):98. <https://doi.org/10.1186/s12859-016-0950-8>.
33. Hay AD, Kessler NJ, Gebert D, Takahashi N, Tavares H, Teixeira FK, et al. Epigenetic inheritance is unfaithful at intermediately methylated CpG sites. *Nat Commun.* 2023;14(1):5336. <https://doi.org/10.1038/s41467-023-40845-2>.
34. von Meyenn F, Iurlaro M, Habibi E, Liu NQ, Salehzadeh-Yazdi A, Santos F, et al. Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. *Mol Cell.* 2016;62(6):983. <https://doi.org/10.1016/j.molcel.2016.06.005>.
35. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics.* 2014;30(5):614–20. <https://doi.org/10.1093/bioinformatics/btt593>.
36. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics.* 2011;27(6):863–64. <https://doi.org/10.1093/bioinformatics/btr026>.

37. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7(5):335–36. <https://doi.org/10.1038/nmeth.f.303>.
38. De Riso G, Sarnataro A, Scala G, Cuomo M, Della Monica R, Amente S, et al. MC Profiling: a novel approach to analyze DNA methylation heterogeneity in genome-wide bisulfite sequencing data. *NAR Genomics Bioinf*. 2022;4(4):lqac096. <https://doi.org/10.1093/nargab/lqac096>.
39. Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, et al. *vegan*: Community Ecology Package; 2025.
40. Wickham H. *Elegant graphics for data analysis*. Springer-Verlag New; 2016. <https://doi.org/ggplot2>.

Publisher's Note

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