Targeted Genomic and Epigenomic Dual Analysis for the Same Region of a Single Gene

Yun Bao[1]*, Grace Zhao[1]*, Heng Wang[1]*, Ashraf Wahba[1], Yingmin Wang[1], Bilan Hsue[2], Ling Su[3], Josh Wang[4], Carole Durand[1], Shengrong Lin[1], Xiaolin Wu[3]

[1]Agilent Technologies, Santa Clara, CA, [2]Agilent Technologies, La Jolla, CA, [3]Frederick National Laboratory for Cancer Research, Frederick, MD, [4]Agilent Technologies, MD

*Equal contribution

AGBT 2025

POSTER #405

Agilent Trusted Answers

Introduction

Integrating genomic and epigenomic data is increasingly crucial for a more comprehensive understanding of biological processes, particularly in complex diseases like cancer. While many studies focus methylation analysis on promoter regions and genomic analysis on exons, research indicates the importance of examining omics information across entire genes, including UTRs and introns.

Current whole-genome technologies for simultaneous genomic and epigenomic analysis can be costly, requiring deep sequencing for low tumor content and often exhibiting limited coverage in high-GC regions. To address these limitations, we developed a novel strategy for simultaneous genomic and methylation analysis of the same region within a single gene from a single DNA input. This approach offers a more affordable solution with improved coverage uniformity, particularly in high-GC regions.

The standard target capture strategy for the Agilent Avida Duo workflow uses two sets of probes targeting different genomic regions for genomic alteration and methylation analysis. Each set captures both Watson and Crick strands, making it ideal for low-input samples like cfDNA. The modified target capture strategy we describe here, however, uses two sets of probes that target the same genomic region. Critically, each set captures only one strand (either Watson or Crick). This single-strand capture requires ~10 ng gDNA and is therefore not recommended for applications that use less input material.

Using ARID1A as a model, we implemented this strategy with the Agilent Avida Duo target enrichment system, enabling sequential capture and acquisition of both genomic and methylation information.

To optimize the quantity of required sequencing reads and coverage uniformity, we divided the 90 kb of ARID1A target region into high-uniqueness and low-uniqueness regions and tested various panel mixture ratios. We also compared two fragmentation methods (enzymatic shearing and Covaris sonication), demonstrating comparable sequencing results for both genomic alteration and methylation analysis.

Pilot studies using tumor gDNA successfully demonstrated the feasibility of our approach for simultaneous targeted genomic and epigenomic analysis of the same region within a single gene.

Materials and Methods

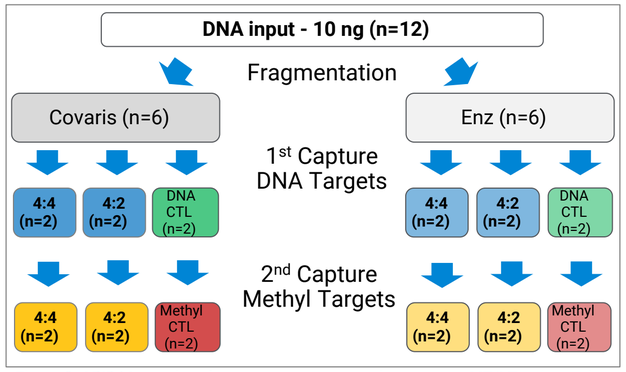


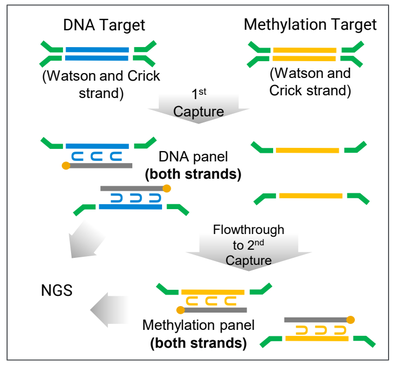
Figure 2: ARID1A panel performance was compared across two MT/LT ratios and two gDNA fragmentation methods. A modified target capture method, targeting the same ARID1A region, was evaluated with the standard method targeting different regions as a control.

In the Agilent Avida Duo workflow, the first capture is for DNA panel capture, and the second capture is for methylation panel. Two different sub-pool mix ratios of ARID1A panel design (Table 1) were used for both DNA and methylation capture. As controls, the Agilent Avida DNA Focused Cancer Panel (27 kb, Cat# 5280-0049) and a methylation panel (TMS 200, 60 kb) were used for the standard Avida Duo workflow (Agilent, Cat# G9440A). As starting material, 10 ng of gDNA (Cat# 5190-8849) was fragmented by either sonication or enzymatic shearing in a 50 µL volume, using two replicate reactions per condition.

Sequencing read budgets for the ARID1A panels were 10 M PE for DNA panel, and 5 M PE for Methyl panel. Sequencing read budgets for the control panels were 5 M PE for both DNA and Methyl.

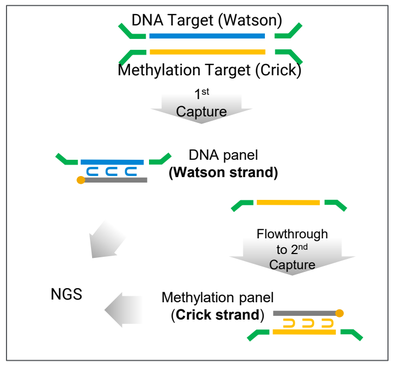
Technology

Standard method captures different regions



Left: The standard target capture method for the Avida Duo workflow captures different target regions for genomic alteration and methylation analysis. Two sets of bridge probes (blue and yellow) target distinct genomic regions- DNA target (blue) and methylation target (yellow), capturing both Watson and Crick strands for a given target. This strategy is optimized for cfDNA and other low input applications (a few ng DNA).

Modified method captures the same region



Right: The modified target capture method targets the same region for both DNA (blue) and methylation (yellow) analysis. Two sets of bridge probes (blue and yellow) target the same region, capturing either the Watson or the Crick strand. Note that this approach was not tested with less than 10 ng gDNA.

Figure 1. The Agilent Avida Duo workflow enables analysis of DNA and methylation, capturing targets from two different target regions or a single target region.

Agilent Avida target enrichment technology utilizes an interlocked, three-dimensional structure for synergistic indirect capture of target DNA. This scaffold forms only when two or more bridge probes ("C" shapes) hybridize to the same target and are stabilized by a biotinylated anchor probe (dark grey bar with orange circle). This synergistic hybridization ensures highly specific capture, enabling binding to streptavidin beads (not shown). Avida's high capture efficiency allows direct capture from unamplified DNA, facilitating a unique dual workflow for simultaneous genomic and methylation analysis from limited samples without splitting. This strategy maximizes recovery and minimizes turnaround time.

ARID1A and ARID panel designs

ARID1A

- ARID1A, a member of the SWI/SNF chromatin remodeling family, plays a crucial role in gene transcription.
- ARID1A acts as a tumor suppressor, regulating cell growth and division.
- Mutations in ARID1A are implicated in various cancers. Beyond well-established variants, intronic variations also contribute to cancer by affecting splicing, gene expression, and other mechanisms. Methylation signatures may aid cancer treatment stratification. Further research is needed to fully elucidate the roles of both genetic and epigenetic changes across the entire ARID1A gene in cancer.

ARID1A panel - genomic region 90 kb, two isoform cDNA 8585 bp, 7934 bp

- Panel composition: ~75% high-uniqueness (MT, "More Than" uniqueness cut off) and ~25% low-uniqueness (LT, "Less Than" uniqueness cut off, primarily intronic low-complexity regions). MT and LT probe groups were blended at different ratios (see Table 1) to optimize both target recovery and on-target rate.
- Panel synthesis: Four sub-pools (MT+, LT+, MT-, LT-), representing Watson (+) and Crick (-) strands.

Materials and Methods

MT/LT ratio	Strand targeted	Panel type	MT (ul)	LT (ul)
4:4	-	DNA	4	4
	+	Methyl	4	4
4:2	-	DNA	4	2
	+	Methyl	4	2

Table 1: For ARID1A panels, two MT/LT sub-pool ratios were used to formulate capture panels for DNA (-) and methylation (+) strands.

Results and Discussion

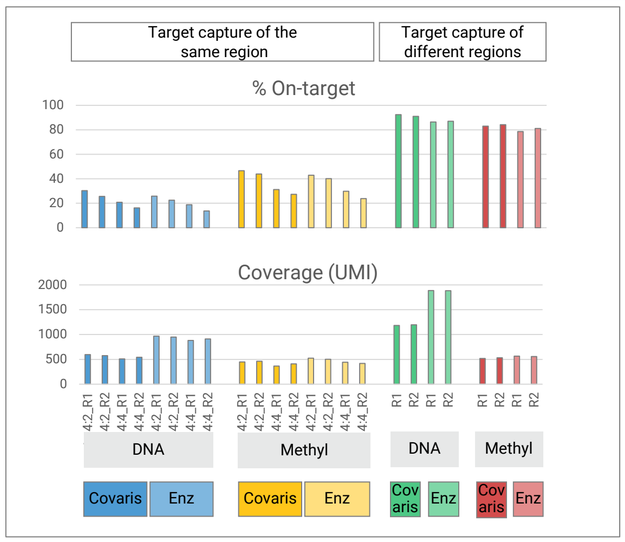


Figure 3. Sequencing performance: panel MT/LT ratio and fragmentation method comparison.

The ARID1A panels, using a modified design approach, target the same region: the "-" strand panel enriches DNA, and the "+" strand panel enriches methylation. The 4:2 panel showed slightly higher recovery and on-target rate than the 4:4 mix.

Control panels used in the standard target capture strategy are targeting both strands, in different regions, and mainly on exonic regions (all MT regions). The recovery and on-target are much higher than ARID1A panel.

For both ARID1A and control panels, enzymatic shearing (Enz) fragmentation yielded higher recovery, but slightly lower on-target rate compared to sonication (Covaris).

Two replicates (R1, R2) were examined for each condition.

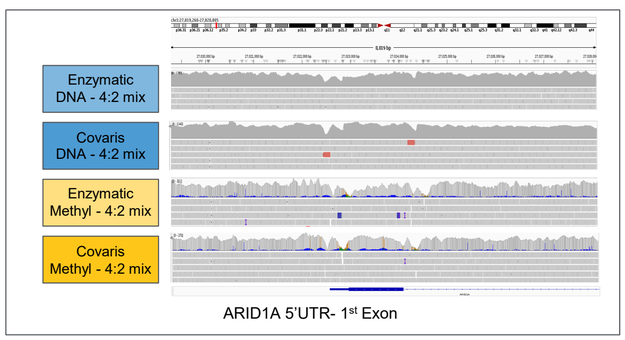


Figure 4. Coverage of the ARID1A 5' UTR and the first exon, regions that are challenging to sequence due to high GC content.

Uniform coverage of ARID1A exons and introns was achieved for both genomic and methylation panels, with comparable coverage between enzymatic fragmentation (Enzymatic) and sonication (Covaris) fragmentation methods.

Results and Discussion

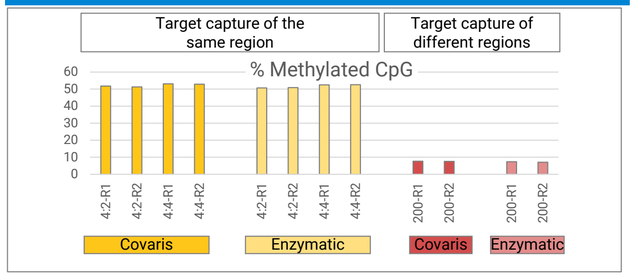


Figure 5. Bisulfite conversion performance: comparison of panel ratios and fragmentation methods

The reported values represent the percentage of cytosine (C) that remained unconverted after bisulfite treatment. The levels of methylated CpG bisulfite conversation rate were similar across different panel ratios (4:4, 4:2) and fragmentation methods (Covaris, dark yellow/ burgundy and enzymatic shearing (Enzymatic, lighter yellow/ lighter burgundy). Two replicates (R1, R2) were examined for each condition. The bisulfite conversion rate for all samples was between 98.9% and 99.3%.

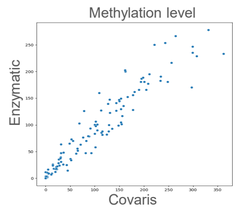


Figure 6. Methylation levels at each probe cluster region are comparable between enzymatic shearing (Enzymatic) and sonication (Covaris) fragmentation methods.

4:2 panel mix (n=2).

The Pearson correlation coefficient is 0.98.

Sample ID	Capture	Type	Raw (M)	Raw median cover	Coverage	% mapped	% on-target	% >0.2X mean
1	DNA	Breast cancer	8	2265	1483	99.5	13.6	91
2	DNA	Breast cancer	6	1631	1125	99.8	13.2	90.4
3	DNA	Breast cancer	9.9	4799	3455	99.8	22.1	89.9
4	DNA	Breast cancer	4.5	1393	854	99.8	14.2	91
5	DNA	Breast cancer	9.3	1959	1294	99.8	10	91.5
6	DNA	Breast cancer	4.9	1103	804	99.8	9.9	91.2
7	DNA	Normal	13.4	2634	1542	99.7	9.2	92.6
Ctl	DNA	Control	18.5	4957	2348	99.8	11.7	93.4
1	Methyl	Breast cancer	5.5	1882	1377	68.7	28.4	93.9
2	Methyl	Breast cancer	4.7	1718	828	64.6	32.4	93.9
3	Methyl	Breast cancer	5.3	2246	645	68	32.9	94.1
4	Methyl	Breast cancer	2.9	1596	409	75.2	35.5	93.6
5	Methyl	Breast cancer	4.8	1205	573	75.6	17.4	94.4
6	Methyl	Breast cancer	3.4	968	823	75.4	19.8	95.7
7	Methyl	Normal	5.6	2539	1281	68.6	37.4	92.8
Ctl	Methyl	Control	7.4	2005	1003	74.1	20.9	94.7

Table 2. Sequencing metrics for ARID1A genomic and methylation analysis of breast tissue gDNA

No genomic variants or methylation alterations were detected in the gDNA (20-50 ng) from tumor (n=7) and normal (n=1) breast tissue samples, as expected, because ARID1A mutations are reported in only ~5-7% of all breast cancers. The gDNA was fragmented using Covaris. Panel mix is 4:2. Control DNA input is Agilent female gDNA.

Conclusions

The Agilent Avida Duo workflow offers a unique solution for multi-omic analysis, enabling the investigation of both genomic variants and methylation alterations from a single sample input, eliminating the need for sample splitting.

While the standard Avida Duo protocol targets two distinct regions for genomic and methyl targets, this study demonstrates a new approach to analyze both genomic and methylation changes within the *same* region.

By implementing a strand-specific capture strategy, we investigated genomic and methylation changes concurrently without modifying the existing NGS sample preparation chemistry. This method captures only one strand of DNA at a time, allowing for simultaneous analysis of both modalities in a defined region.

Furthermore, we addressed the challenges associated with low complexity sequencing in specific panel regions. Leveraging the modularity of the Avida hybridization system, we explored various combinations of high-uniqueness and low-uniqueness sequence regions to optimize both target recovery and on-target rate. This adaptable approach allows for customization based on the specific research needs.

Finally, we successfully demonstrated the feasibility of integrating enzymatic fragmentation into the Avida Duo workflow, achieving acceptable recovery and on-target rates. The Agilent Avida Duo workflow also provides reasonable coverage of high GC regions, such as the 5' UTR, for both DNA and methylation analysis, expanding its applicability to diverse genomic regions of interest.

 PR7001-3891