

Determination of DNA methylation on promotor activity

Stephan Drotler, *University of Salzburg*

Abstract—DNA methylation occurs on Cytosines in a CpG context, which accumulate predominantly in regulatory regions of the genome. Epigenetic treatments of loci associated with genetic disorders, is current state of research while some drugs are already approved for certain diseases. In malignant transformations, regions coding for tumor suppressor genes are often silenced by hypermethylation. A known locus for non-small cell lung cancer (NSCLC) is the p16 promoter, on chromosome 9p21 which controls the INK4a/ARF tumor suppressor transcripts [1].

In this paper we use human lung cancer in-vitro model (H1299), to assess the effect of the unspecific hypomethylating agent decitabine (DAC) on the p16 promoter. After initial treatment of a H1299 cells with DAC, pyrosequencing is performed to determine the methylation changes on a CpG island in the INK4a/ARF locus. Furthermore, sequencing of LINE1, a repetitive element associated with genomic instability is used as a marker for genome-wide demethylation. To further demonstrate the effect of methylated CpG-rich promoters on gene transcription, a methylation specific luciferase assay in HEK293T cells is executed. The sequencing products yielded low-confidence results, yet showing hypomethylation of treated LINE1 and p16 areas. Expression analysis confirms higher transcription for unmethylated constructs, however data acquisition was also error-prone. As a consequence, similar data from the module 2017, serves as a control for the presented results.

Key Words— BT conversion, NSCLC, Pyrosequencing, DNA methylation, LINE1, Cancer, human cell culture, Luciferase assay, Decitabine, p16, INK4a/ARF

I. INTRODUCTION

CpG islands are genetic regions with an elevated occurrence of CpG sites, often found in regulatory regions of the genome. They serve as the only recognition motif for DNA methyl transferases (DNMTs), which catalyze the methylation of Cytosine to 5-methylcytosine (5mC). The human genome encodes three DNMTs enzymes. DNMT1 is responsible for the re-establishment of methylation patterns from replicated, hemi-methylated DNA. DNMT3a and DNMT3b, are de-novo DNMTs and establish new methylation patterns. Methylated CpGs in promotor regions correlate with low expression of the concerned genes, while unmodified CpGs are associated with higher transcription. This mechanism presents the most efficient control for the cellular proteome, thus contributing to an immense intra- and intercellular complexity. Methylation states can reflect environmental exposures, disease incidents, aging, cellular development and differentiation. Additionally, CpG methylation conduces genomic stability by silencing transposable elements. Quantification of methylation levels is therefore of high interest in many life science branches [2].

The experiments were performed at the Biomedical Institute, Billrothstraße 11, University of Salzburg from 02.12.19 to 13.12.19 under the supervision of Florian Wolff.

Tumor suppressor genes transcribe surveillance proteins which are often part of the cell-cycle control or DNA repair. Inactivating mutations or hypermethylation of those genes or their respective promotor, favor tumor growth and are therefore frequently observed in cancer cells. The INK4a/ARF (Inhibitors of CDK4/Alternative Reading Frame) can be transcribed into several distinct transcripts which have an integral role in cell-cycle control mechanism. Of interest for this work is the p16INK4a gene product, which inhibits phosphorylation of the retinoblastoma (RB) protein by D-Cyclins and cyclin dependent kinases (CDKs). Phosphorylation of RB by CDKs prevents RB from binding to and inhibiting E2F1 transcription factors which are needed for S-phase entry [2].

Reversion of the hypermethylation and subsequent expression of silenced tumor suppressors can make tumor cells more prone for senescence or even apoptosis. Additionally, unspecific hypomethylation can be accompanied by upregulation of endogenous retroviral (ERV) constructs. This process can lead to viral mimicry of the cell and might result in the induction of potent anti-tumor response. Within this research subject, hypomethylating agents like decitabine are of great interest. DAC is already approved for the treatment of elderly patients with acute myeloid leukemia (AML) or myelodysplastic syndrome (MDS). The effect of DAC for patients with NSCLC is subject of this work [2][3].

Pyrosequencing with prior BT conversion and expression analysis via luciferase is used to determine the promotor methylation state and activity. These techniques are explained in further detail in the Methods section.

II. METHODS AND MATERIALS

A. MATERIALS

1) TOOLS:

- Eppendorf-pipettes (1µL - 1000µL)
- Qiagen PyroMark Q24 Pyrosequencer
- Tecan plate reader

2) CHEMICALS:

- PCR reagents for all reactions (provided by institute)
 - MgCl
 - dNTPs
 - Hot Star Taq Polymerase
 - ddH₂O
 - Buffer 10x
- Thermo GeneRuler 100 bp Ladder SM0321 (For all gels in supplement, 1 step = 100 bp)
- cell culture media (provided by the institute)
- DMSO, DAC (provided by the institute)

	medium	DAC (20mM)	DMSO
U	5ml	-	-
DMSO	5ml	-	0.125µl
0.1 DAC	4ml	1ml 0.5 DAC	0.1µl
0.5 DAC	20ml	0.5µl	-

Table I: Recipes for the treatment media. Abbrev: U = untreated, DMSO = DMSO control group, 0.1 DAC = 0.1µM DAC, 0.5 DAC=0.5µM DAC

B. METHODS

1) SEQUENCING:

a) Cell culture:

Seed human lung cancer cell line (H1299, 2.5×10^5 cells/well) and grow them for 24h in the incubator. The cells are grouped and incubated with the according treatment medium for 24h (Tab. I). Harvest cells with trypsin according to the lab SOP, and isolate gDNA with the QIAamp DNA Blood Mini Kit (Cat. 51104) manual.

b) BT conversion:

The isolated DNA is diluted and treated with HSO_3^- converion reagent under acidic conditions. The unmethylated cytosines will therefore be deaminated into uracil, while 5-methylcytosine is protected by the methylgroup. In the PCR (dna meth kit) reaction with converted DNA (BT-DNA) as template, thymines are incorporated for uracils. The resulting product contains thymines at formerly unmethylated cytosines and cytosines for 5-methylcytosines. The PCR product is collected, washed, desulphonated and eluted. Concentrations are assessed by nanodrop measurement.

c) Sall-3 PCR:

Sall-3 PCR is performed to as a quality control for conversion. The primers contain long T sequences that bind very specifically to BT-converted DNA.

Primer 5'→3', Expected fragment size: 200 bp.

18 Sall3-fw-b1 GTTGCGGTTGGTTTTGT

19 Sall3-rv-b1 ACCCTTACCAATCTCTAACTTTC

d) LINE1 PCR:

LINEs are a functional retroelements which are spread over the human genome. Methylation of these sequences is one mechanism to silence them and secure genomic stability. Application of hypomethylating agents like DAC, can lead to global demethylation hence reactivate LINE1 elements. Sequencing the methylation of LINE1 elements is a control for global hypomethylation.

Primer 5'→3'

32 LINE1-fw-b2 TTTTGAGTTAGGTGTGGATATA

33 LINE1-rv-b2 BIO-AAAATCAAAAAATTCCCTTTC

e) p16 PCR:

p16 PCR is performed for sequencing in order to assess the methylation levels in the promotor region.

Primer 5'→3'

97 p16-fw-1-b1 GAGGGGTTGGTTGGTTATTAGA

98 p16-rv-1-b1 BIO-TACAAACCCTCTACCCACCTAAAT

All Agarose gels are provided in the Supplement.

f) Pyrosequencing:

Pyrosequencing is a sequencing by synthesis method that requires a single strand template DNA for nucleotide incorporation. To isolate exactly that sequence from the PCR products, one of the primers is biotinylated (LINE1 rev 33, p16 rev 98). Isolate the elongated strands via streptavidin-coated sepharose beads, shake and add appropriate sequencing primers (34, 99). The PyroMark cartridge and the preparation of the template DNA are performed according to the SOP of the lab.

SQ-Primer 5'→3'

34 LINE1-SEQ-b2 AGTTAGGTGTGGATATAGT

99 p16-SEQ-1-b1 GGTTATTAGAGGGTGGGG

Pyrosequencing indirectly quantifies the pyrophosphate generated during elongation by the DNA polymerase. Sulfurylase generates ATP, serving the luciferase as energy for conversion of luciferin to oxyluciferin and light ($\text{APS} \xrightarrow{\text{PP}_i} \text{ATP}$, $\text{luciferin} + \text{O}_2 \xrightarrow{\text{ATP}} \text{Oxyluciferin} + \text{CO}_2 + \text{Light}$ [simplified]). Since only one nucleotide is provided per round, Apyrase degrades remaining dNTPs, in order to prevent false positive signals. The emitted signal is detected and directly proportional to the number of integrated dNTPs.

2) LUCIFERASE EXPRESSION ASSAY:

a) Cell culture:

Seed human embryonic kidney cells (HEK293T, $1, 2 \times 10^5$ cells/well) and incubate them overnight. Transfect them with the different plasmids, as described in Table II.

b) Luciferase readout:

Cells are lysed, and the constructs are purified for detection. A dual reporter system is employed whereby the firefly plasmid (pGL4 SV40) serves as a transfection control which is always expressing and utilizing luciferin as a substrate. Two different reporter constructs are cloned into the pCpGfreeLucia-p16 vector (CpG free), and methylated. Only clone 1 spans a CpG island while construct 5 reporter constructs are cloned into the pCpGfreeLucia-p16 vector (CpG free), and methylated. Only clone 1 spans a CpG island while construct 5 is not can not be methylated (Fig. S7). An empty vector (pCpGfree) serves as a negative control and as compensation for the two constructs. Construct 5 is shorter and therby lighter than the first, hence the empty Vector is added to C5 in order to balance the DNA molarity (Tab II, Fig. S7).

	construct 1	construct 5	negative control
constructs	450 ng	348.8 ng	0 ng
firefly pGL4	25 ng	25 ng	25 ng
empty Vector	0 ng	101.2 ng	450 ng

Table II: Construct 1 and construct 5 span differnt regions. Resulting groups and abbreviations are described in Figure 2.

RESULTS

C. Pyrosequencing

The data generated by pyrosequencing was mostly not credible, due to an exceedance of the threshold for the internal

control. Thereby the system deliberately injects a Cytosine instead of a Thymine, outside a CpG context. Since Cytosine can not be methylated in this sequence context, a signal represents an ineffective BT conversion. All bars in Figure 1, framed in red represent unreliable data, while orange framed ones have higher credibility and black framed data is trustworthy. The methylation standards for the p16 promoter (Fig. 1A) are fitted using a polynomial equation of third order (Fig. 1B). Only the first two CpG sites are used for fitting, because the 0% and 20% standard only produced useable data at those sites. With the extracted equation (Eq. 1), x values are calculated using the Wolfram Alpha webtool [5].

$$y = 1 \times 10^{-5} x^3 + 0.0028x^2 + 0.2587x + 3.6542 \quad (1)$$

The results are consequently plotted in order to compensate for PCR biased amplification (Fig. 1E). The data for the LINE1 (cell culture 2) construct passed the internal control and is visualized in Figure 1C. Cell culture 1 is not sequenced because, a DNA product was visible in the gel. Out of curiosity and the availability of space, the negative control is sequenced and shows clearly a DNA amplicon (The gel As a comparison, reliable sequencing data from the seminar 2017 is provided in Figure S1, for later discussion of the results).

D. Expression assay

The results from the dual reporter system show a mean firefly signal of 276244.75 counts/sec (± 68495.73) compared to the signal from the 2017 data set of 899511.4 counts/sec (± 19813.4) (Fig. 2A,D). Lucia/Firefly proportions are plotted for the individual groups in Figure 2C,E where the control group is used to normalize signals (Fig. 2C,F).

III. DISCUSSION

The high uncertainty of the pyrosequencing results can not be traced back to a specific error, however some hypothesis are contingent. If the BT conversion was not efficient, the standards should still produce robust data because they were not prepared in the course. Consequently it is suspected that an enzyme in the reaction mix was not working optimally, also indicated by false positive peaks on other sites in the sequence. In addition, the standards are skewed compared to the data from 2017. The maximum standard (100%) produces a relative methylation level of 83.4% (2017) but only 62.7% (all CpG sites)/ 70.51 % (first two CpG sites) in our data. As a consequence, the relative methylation levels of some p16 samples exceeded 100% methylation. Nevertheless, the data for p16-ZK1 contains expected trends, with no methylation difference between DMSO and untreated, lower methylation for the 0.1 DAC and lowest methylation for the 0.5 DAC group. LINE1 also represents this pattern, reinforced by reliable, significant results (Fig. 1C). In contrast p16-ZK2 shows higher methylation of 0.5 DAC, than 0.1 DAC. This might be due to bad cellular health of the 0.5 DAC-ZK2 group, observed in the microscope, before harvest. Patient A has a relative methylation level (corrected) of 18.6% while Patient B has 56.9%. Since our corrected values are overestimated,

the promotor of Patient A should be mostly hypomethylated while in Patient B, \approx half of the CpGs should be methylated.

The expression analysis yielded \approx 3.3 fold lower transfection efficiency with a \approx 3.5 fold higher standard deviation, compared to the 2017 data. Additionally, the negative control has high Lucia expression, which leads to the assumption that the empty vector was interchanged with an unmethylated Lucia construct. As a consequence the normalized data show only a fraction of the expression of the negative control. In comparison, the results from 2017 show a strong signal from the unmethylated construct one, and almost no response from the other groups. Accordingly, the expression of this construct is multiple times higher than the negative control. The transcription of two constructs 5 is not equal, even though it spans no CpGs. The origin of this difference is unclear, and might be due to inaccurate working steps. In contrast, the 2017 data displays not distinct expression of the construct 5 groups.

APPENDIX SUPPLEMENT

All gel pictures in the Supplement are cast with 1.5% Agarose in Borate Buffer, stained with Midori green. Gel runs are run at 120-150 V.

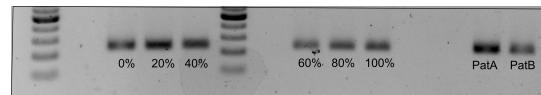


Figure S3: Sal3 size control agarose gel (1,5%), \approx 200-300bp.

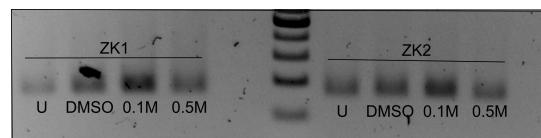


Figure S4: LINE1 size control agarose gel (1,5%), \approx 200bp.

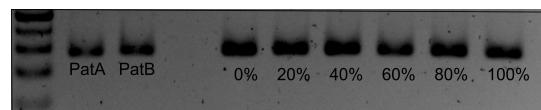


Figure S5: p16 size control agarose gel (1,5%), \approx 200-300bp.

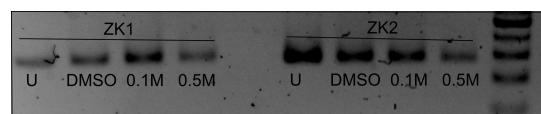


Figure S6: p16 size control agarose gel (1,5%), \approx 200-300bp.

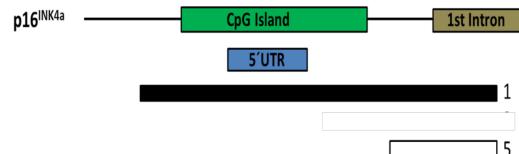


Figure S7: Visualisation of construct 1 and 5, showing the CpG island.

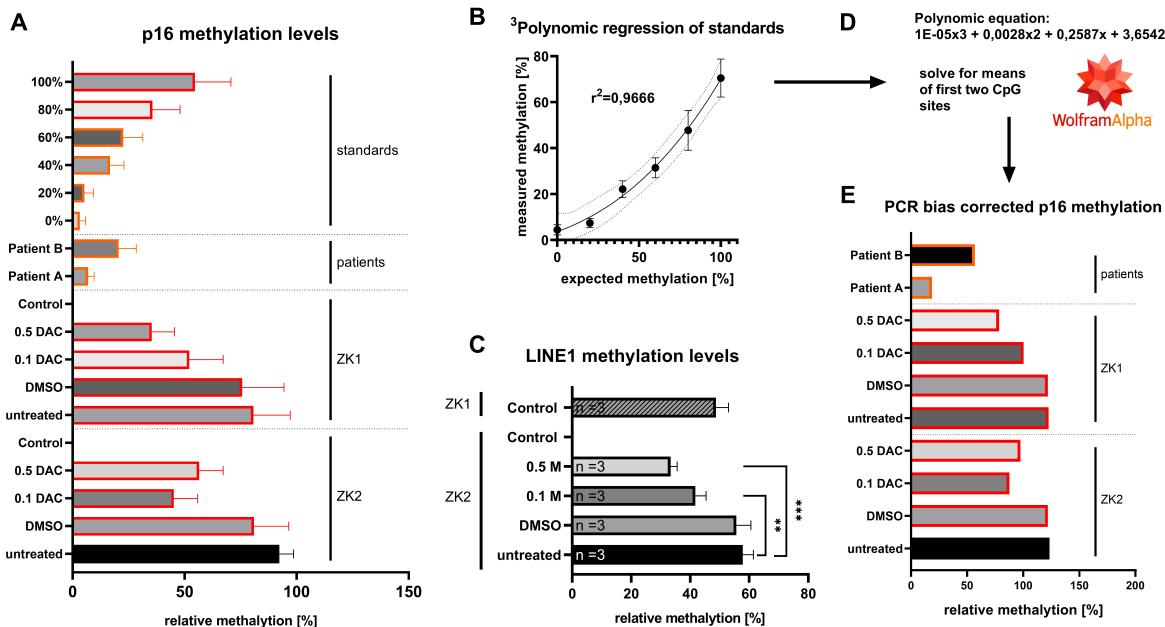


Figure 1: Pyrosequencing data of p16 and LINE1. **A:** Methylation levels of the p16 promoter after BT conversion. **B:** Polynomic regression of p16 methylation standards. **C** LINE1 methylation levels after BT conversion. **D:** Equation of the fitted curve from **B** used to calculate corrected methylation levels for p16 data (first two CpG sites). **E:** Corrected methylation levels of p16 promoter. Abbrev: 0.1 DAC = $0.1\mu\text{M}$ decitabine, 0.5 DAC = $0.5\mu\text{M}$ decitabine, ZK1 = cell culture 1, ZK2 = cell culture 2. Grouped comparisons performed via one-way ANOVA in Graphpad Prism 8 [4]. Results are considered significant when $p\text{-value} < 0.05$.

REFERENCES

- [1] M. D. Crowther, G. Dolton, M. Legut, M. E. Caillaud, A. Lloyd, M. Attaf, S. A. Galloway, C. Rius, C. P. Farrell, B. Szomolay, *et al.*, “Genome-wide crisper-cas9 screening reveals ubiquitous t cell cancer targeting via the monomorphic mhc class i-related protein mr1,” *Nature Immunology*, vol. 21, no. 2, pp. 178–185, 2020.
- [2] A. Risch, “Epigenetics of cancer,” 2019.
- [3] F. Wolff, M. Leisch, R. Greil, A. Risch, and L. Pleyer, “The double-edged sword of (re) expression of genes by hypomethylating agents: from viral mimicry to exploitation as priming agents for targeted immune checkpoint modulation,” *Cell Communication and Signaling*, vol. 15, no. 1, p. 13, 2017.
- [4] M. L. Swift, “Graphpad prism, data analysis, and scientific graphing,” *Journal of chemical information and computer sciences*, vol. 37, no. 2, pp. 411–412, 1997.
- [5] “Wolframalpha,” Wolfram Alpha LLC. [Online]. Available: <https://www.wolframalpha.com/> [Accessed: 2019-12-19]

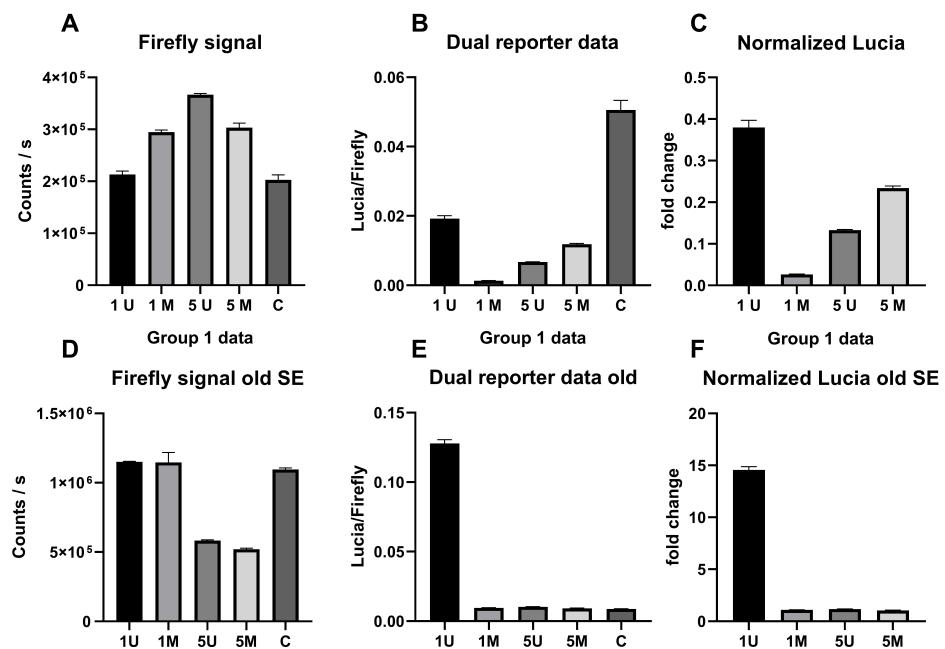


Figure 2: Transcription activity measurement via a dual reporter system of firefly and lucia. **A,B,C** represent measured data, **D,E,F** shows data from the seminar 2017.

A: Firefly signal of HEK293 cells confirms transfection. B: Lucia expression controlled for variations in transfection shown in A. C: Data was normalized against expression of the empty control vector. Description of sub-figures follows the same scheme ($A \leftrightarrow D, B \leftrightarrow E, C \leftrightarrow F$). Abbrev: 1U = Construct 1 unmethylated, 1M = construct 1 methylated, 5U = construct 5 unmethylated, 5M = construct 5 methylated, C = control.

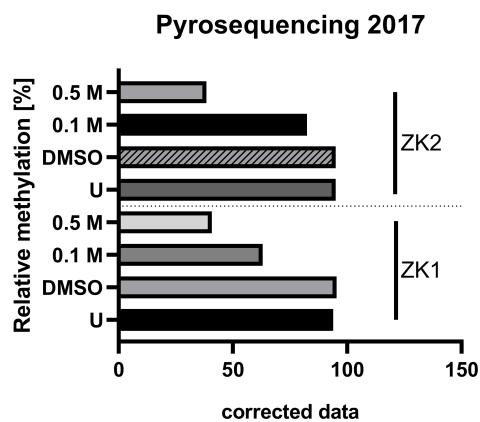


Figure S1: Pyrosequencing results (2017) for p16 constructs.

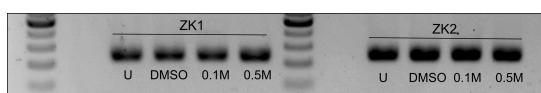


Figure S2: Sall3 size control agarose gel (1,5%), \approx 200-300bp.