Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Chaojie Zhong Date

Long-Term Epigenetic Impact of Chemotherapy in Breast Cancer Patients

By

Chaojie Zhong
Master of Science

Graduate Division of Biological and Biomedical Sciences

Cancer Biology and Translational Oncology

_____
Paula M. Vertino, Ph.D.
Advisor

_____
Arri Eisen, Ph.D.
Committee Member

_____
Adam I. Marcus, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Long-Term Epigenetic Impact of Chemotherapy in Breast Cancer Patients

By

Chaojie Zhong
B.S., Emory University, 2017

Advisor: Paula M. Vertino, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in the Graduate Division of Biological and Biomedical Sciences
Cancer Biology and Translational Oncology
2018

Abstract

Long-Term Epigenetic Impact of Chemotherapy in Breast Cancer Patients

By Chaojie Zhong

Breast cancer is one of the most prevalent cancers among females. Chemotherapy remains an efficacious mainline treatment option, although it is found to be highly toxic and contributes to Cancer Related Fatigue (CRF). The mechanism through which chemotherapy may lower quality of life among breast cancer survivors warrants investigation for its implication in healthcare decisions. Previously, DNA methylation in 3 regions proximal to genes ubiquitin specific peptidase 2 (USP2), SMAD3, and vacuole membrane protein 1 (VMP1) have been associated with prior chemotherapy in breast cancer patients. The difference in DNA methylation was also associated with inflammation, which correlated with CRF. The differential DNA methylation persisted 6 months post treatment as did its association with prior chemotherapy and inflammation. We developed a targeted bisulfite sequencing technique to interrogate specific genomic regions and examined whether significant difference in DNA methylation at these previously identified regions remained 12 months post treatment and whether the difference in DNA methylation was still associated with chemotherapy and inflammation in breast cancer patients from the same cohort. We found that DNA methylation levels at the 3 genomic regions trended lower in patients who received chemotherapy than in those who did not 12 months post treatment, and the difference was statistically significant at 1 region in the gene USP2. DNA methylation at the 3 regions in patients who exhibited higher levels of inflammation and CRF also trended lower 12 months post treatment, but there was no statistically significant association between DNA methylation level and inflammation or CRF. Our findings were consistent with lower levels of DNA methylation at the 3 regions associated with prior chemotherapy, but the difference decayed at 6 months post treatment and further at 12 months post treatment. The results indicate that chemotherapy contributes to persistent CRF in breast cancer patients and imprints the epigenome of peripheral blood mononuclear cells at putative regulatory regions.

Keywords: Breast cancer; Chemotherapy; DNA methylation; Inflammation; Cancer related fatigue; Targeted Bisulfite Sequencing

Long-Term Epigenetic Impact of Chemotherapy in Breast Cancer Patients

By

Chaojie Zhong
B.S., Emory University, 2017

Advisor: Paula M. Vertino, Ph.D.

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in the Graduate Division of Biological and Biomedical Sciences
Cancer Biology and Translational Oncology
2018

Acknowledgements

The Vertino Lab:

Paula M. Vertino, Ph.D.

Benjamin Barwick, Ph.D.

Doris Powell, M.S.

Priya Kapoor, Ph.D.

Cara Shields, B.S.

Elizabeth Zoeller, B.S.

Thesis Committee:

Arri Eisen, Ph.D.

Adam I. Marcus, Ph.D.

Laney Graduate School

GDBBS

CBTO

Emory School of Medicine

Donna Mister

Greg Orloff, Ph.D.

Winship Cancer Institute:

Karen Conneely, Ph.D.

Andrew H. Miller, M.D.

Mylin A. Torres, M.D.

Erwin G. Van Meir, Ph.D.

Canhua Xiao

Table of Contents

Introduction

1. Breast Cancer Overview

Breast cancer is one of the most prevalent types of cancer among females. It occurs in males as well at a much lower rate. The American Cancer Society estimates that in 2017 (Table 1), there were 316,120 new cases of breast cancer diagnosed in American Women and 40610 deaths resulting from breast cancer (American Cancer Society, 2017). It is estimated that there will be 266,120 new cases and 40,920 deaths in 2018 (American Cancer Society, 2018). Not only is breast cancer a common malignancy, it is also a heterogeneous disease that is challenging to target and control. It is often categorized into various subtypes that have clinical implications (Anderson et al., 2014).

Table 1. Estimated New Female Breast Cancer Cases and Deaths by Age, US, 2017

	In Situ Cases		Invasive Cases		Deaths	
Age	**Number**	**%**	**Number**	**%**	**Number**	**%**
<40	1,610	3%	11,160	4%	990	2%
40-49	12,440	20%	36,920	15%	3,480	9%
50-59	17,680	28%	58,620	23%	7,590	19%
60-69	17,550	28%	68,070	27%	9,420	23%
70-79	10,370	16%	47,860	19%	8,220	20%
80+	3,760	6%	30,080	12%	10,910	27%
All ages	**63,410**		**252,710**		**40,610**	

Estimates are rounded to the nearest 10. Percentages may not sum to 100 due to rounding.

Table 1: Estimated new female breast cancer cases and deaths by age, US, 2017 (American Cancer Society, 2017).

Breast cancers are divided into different stages. There are 5 stages in total: 0, I, II, III, and IV. The noninvasive breast cancers, namely ductal carcinoma in situ and lobular carcinoma in situ, belong to stage 0 and have a good prognosis but only approximately 20% of all breast

cancers are detected at this stage. Stage I denotes early-stage invasive cancer, while stage IV denotes advanced cancer with distant metastasis (American Cancer Society, 2018). Invasive breast cancers can be further divided into various histological subtypes based on cell morphology, growth, and structure patterns (Malhotra et al., 2010; Dieci et al., 2014). Some examples include mucinous carcinoma, cribriform carcinoma, and apocrine carcinoma. With the advancement of gene expression profiling technologies, 4 main molecular subtypes have been identified and examined: Luminal A, Luminal B, Her-2 Enriched, and Basal-like (Prat et al., 2015).

However, novel molecular classification remains limited in practice. The clinical treatment of breast cancers primarily focuses on therapeutic groups: the estrogen receptor (ER) positive group, the human epidermal growth factor receptor 2 (Her-2) amplified group, and the triple negative breast cancers (TNBC) (The Cancer Genome Atlas Network, 2012).

2. Breast Cancer Treatment

Clinically, there are 4 overarching therapies used to treat breast cancer: surgery, chemotherapy, radiotherapy, and hormonal therapy (Matsen and Neumayer, 2013). There are novel therapies in clinical trials such as immunotherapy (Ernst and Anderson, 2015), which is not widely available to general patients.

Surgery has been the primary breast cancer treatment option. William Halsted introduced radical mastectomy, which was the prevalent surgical method against breast cancer since the late 19th century until recent decades. Since then, less radical surgeries and breast-conserving procedures that are equivalent to radical mastectomy in terms of efficacy have developed and gained recognition. Surgery also gives access to the tumor itself, allowing for more accurate staging and molecular analysis (Matsen and Neumayer, 2013).

Chemotherapy remains widely used against breast cancer despite cancer heterogeneity. In particular, breast cancers that lack distinct markers, namely the TNBCs, are presented with only chemotherapy options. Chemotherapy is divided into neoadjuvant and adjuvant chemotherapies. Neoadjuvant chemotherapy is administered before surgery, while adjuvant chemotherapy is the same agent administered after surgery. The most common chemotherapy agent regardless of administration timing is anthracycline based. Studies suggest that Basal-like breast cancers, which have the poorest prognosis among all molecular subtypes, are more sensitive to neoadjuvant chemotherapy than Luminal breast cancers are (Carey et al., 2007), although breast cancers that received neoadjuvant chemotherapy may have higher local recurrence after surgery than those that did not receive neoadjuvant chemotherapy (Early Breast Cancer Trialists' Collaborative Group, 2018).

Radiotherapy is also a common treatment option that is used in nearly half of breast cancer patients (Taylor and Kirby, 2015). Radiotherapy uses ionizing radiation to target and eliminate cancer cells. Radiotherapy has been shown to lower both the recurrence and death rate in patients that received breast-conserving surgery (Early Breast Cancer Trialists' Collaborative Group, 2011). The combination of radiation therapy and breast-conserving surgery also proves to be at least equivalent to mastectomy in terms of survival (van Maaren et al., 2016).

Hormonal therapy has been in use to treat breast cancers with distinct markers such as ER. Hormonal therapy such as aromatase inhibitor aims to block ER pathways that drive proliferation, thereby controlling cancer progression. Other targeted therapies, such as Her-2 directed therapy, are also available and under further development (Ahmed et al., 2015). Together, hormonal and targeted therapies present more accurate therapeutic strategies relative to

traditional chemotherapy and radiotherapy, although the reliance on distinct markers also limits their application.

Last but not least, although therapy options are listed as individual topics, in practice, due to the heterogeneous nature of breast cancer and adaptation by the tumor over time, treatment plans must contend with two major issues: 1, various combinations of drugs are necessary to effectively target and control the tumor, and 2, even if the tumor responds well to the treatment at first, it may still develop resistance to the treatment regimen and increases risk of relapse. To summarize, due to the multitude of theoretical and clinical limitations, mainline treatment plans targeting breast cancers still mostly consist of surgery, chemotherapy, radiotherapy, and hormonal and targeted therapies. They are frequently used in combination for best result.

3. Breast Cancer Survivorship

Breast cancer survivorship begins at diagnosis and continues after the end of treatment. Broadly speaking, as long as a patient is alive, the patient can be considered a survivor. The current study will restrict the scope of breast cancer survivorship to the patient wellbeing after treatment, which consists of a number of factors beyond the cancer itself. These factors range from physical symptoms such as pain and infections to emotional issues such as anxiety resulting from diagnosis and surgical procedures such as mastectomy. Breast cancer survivorship can be particularly complicated by the various side effects of therapies which lower quality of life and are difficult to alleviate.

The mainline chemotherapy agents such as anthracycline based ones are known to induce a number of side effects. Notably, doxorubicin may cause cardiomyopathy, which is a highly lethal condition (Chatterjee et al., 2010). Radiotherapy is itself a powerful carcinogen, and can raise the risk of secondary cancers in patients (Burt et al., 2017). Due to the known toxicity of

the prevalent therapy options, some have questioned the necessity of chemotherapy treatment. It has been suggested that depending on the gene-expression profile of patients, harsh chemotherapy agents may not be necessary because some patients may not respond to the drug and exposure to side effects would be unwarranted (Sparano et al., 2015).

All cancer therapies, including endocrine and targeted therapies that show considerably less toxicity than agents such as anthracycline, can still cause systemic side effects in addition to local side effects in particular organs. The immune system is composed of highly replicative blood cells. These cells are susceptible to typical cancer therapeutics targeting cell division or signaling and particularly at risk, resulting in both pro-immunogenic and immune suppressive effects of chemotherapy (Formenti et al., 2010). The dysregulation of the immune system can further lead to inflammation, increased risk of secondary disease, and cancer related fatigue (CRF), an often persistent condition that greatly lowers the quality of life of survivors (Bower, 2015).

4. Cancer Related Fatigue (CRF)

From a clinical perspective, CRF presents a unique problem in cancer research where the question is not about the tumor and its treatment, but the effect of therapy on non-cancerous tissues and the symptoms that arise as a result. Even though CRF is a singular term, it is a collection of symptoms in cancer survivorship characterized by pain, distress, cognitive variables, weakened physical performances, and health status (Figure 1). Within health status, the condition of peripheral blood and time after treatment are important parameters as peripheral blood reflects both genetic and epigenetic changes and inflammation level, while time after treatment shows CRF is a chronic condition that persists after the end of therapy.

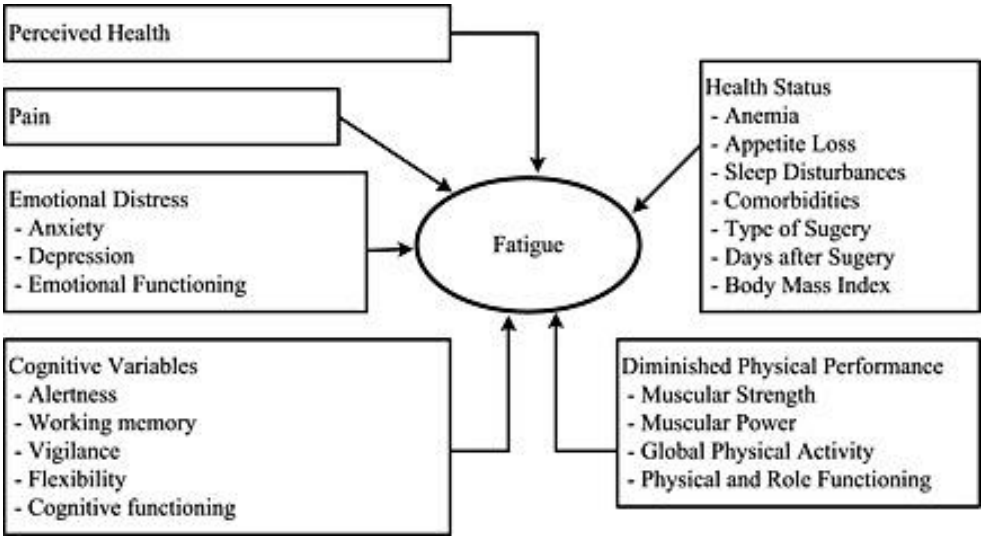

Figure 1: The many contributing factors to fatigue (Bortolon et al., 2014). CRF is a complex condition whose mechanism remains unclear.

It is estimated that during treatment for breast cancer, nearly 70% of patients suffer from CRF. Up to 30% also note persistent CRF months and even years after completing treatment (Bower et al., 2011). Despite the fact that CRF is a common condition, the mechanism of the development and persistence of it is still not well understood. Furthermore, CRF itself cannot be treated effectively, suggesting that CRF may be a condition that is far more complex than its name suggests (Barsevick and Cella, 2010).

Although the genetic aspect of CRF has been under investigation (Bower, 2015), recent work by Smith et al. identified differences in epigenetic changes in DNA extracted from noncancerous peripheral blood mononuclear cells of patients who were treated with chemotherapy versus those who did not, suggesting that epigenetics may play an important role in the persistence of CRF (Smith et al., 2014).

5. Epigenetics

Epigenetics, broadly speaking, is the study of gene expression or cellular phenotype that occurs without changes to the DNA sequence itself. While in classical genetics phenotypic

changes require genetic alterations, epigenetics fills the gap as “the bridge between genotype and phenotype” (Goldberg et al., 2007).

A major molecular mechanism and well characterized example of epigenetics is DNA methylation, the enzymatic addition of a methyl group to DNA performed by DNA methyltransferases (DNMT) on the 5’-carbon of the pyrimidine ring in cytosine (Miller and Grant, 2013). DNA methylation can perform a range of functions within a cell depending on the cellular context (Jones, 2013). Traditionally, DNA methylation was viewed as an important factor in gene expression repression, possibly by blocking transcription from promoters by preventing the binding of transcription factors. Currently, DNA methylation proves to be more complex than a mere expression repressor and it is crucial for cell differentiation and development (Phillips, 2008). For instance, promoter methylation can stabilize gene repression in imprinted genes (Jones, 2013), while the methylation of a gene body may positively correlate with active transcription (Hellman, 2007). In undifferentiated cells, DNA methylation maintains the cell pluripotency, while later in development DNA methylation is involved in maintaining the differentiation state, and different cell lineages exhibit different DNA methylation patterns (Khavari et al., 2010; Roost et al., 2017). It is also observed that during regular cell replication and quiescence, DNA methylation is stable in human primary fibroblasts (Vandiver et al., 2015). Although DNA methylation can be considered stable in different cell types, it is not a static feature. In fact, there appears to be a rapid turnover of DNA methylation in human cells, which also complicates how a certain epigenetic event trigger actually induces changes in DNA methylation levels (Yamagata et al., 2012).

6. Epigenetics and Cancer

Due to its significant involvement in cellular regulation and function, it is not surprising that epigenetics also plays a role in human diseases, including cancers (Baylin, 2011). A common epigenetic modification in breast and other cancers, for example, is the abnormal hypermethylation of Cytosine-phosphate-Guanine (CpG, see Figure 2) islands (CGIs, see Figure 3) (Basse et al., 2014). A CpG dinucleotide consists of a cytosine that precedes a guanine (Esteller, 2008), and CGIs are successions of 500 up to 2,000 base pairs in length that are enriched in CpG dinucleotides mainly located in the proximal promoter region of genes (Parrella, 2010). Hypermethylation of the promoter region leads to inappropriate silencing of tumor suppressors and contributes to the cancer phenotype (Jones, 2002). In fact, some CGIs in promoters of tumor suppressor genes are prone to methylation due to aging, increasing the risk of cancer naturally (Miller and Grant, 2012).

Figure 2: A Cytosine-phosphate-Guanine (CpG) site. The CpG site is shown in the black box. (By Helixitta (Own work) [CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0)], via Wikimedia Commons)

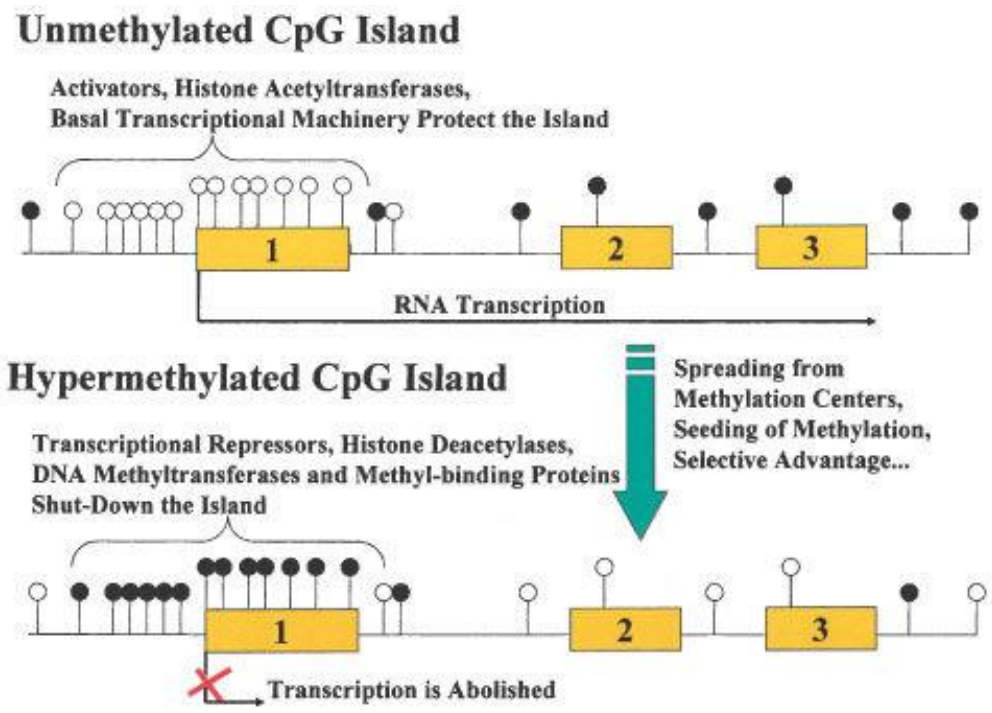

Figure 3: CpG Island (CGI). In this case, the methylation landscape at the promotor region of a gene is shown. Promotor hypermethylation inhibits the transcription of the gene. Yellow boxes: exons; Black circles: methylated CpGs; White circles: unmethylated CpGs (Esteller, 2002).

In concordance with the heterogeneous nature of breast cancers, array data show that the DNA methylation patterns vary greatly between different breast cancers (The Cancer Genome Atlas Network, 2012).

However, since DNA methylation is a ubiquitous modification, the implication of it in breast cancers reaches far beyond the tumor itself. As noted above, radiotherapy and chemotherapy against cancer can be highly toxic, and even less toxic agents such as endocrine therapy and targeted therapy can still have systemic effects on vital parts of the body, such as the immune system. Ionizing radiation is detrimental on a genetic level and anthracycline agents, as mentioned before, can induce tissue damage. One possibility is that cancer therapies impact the non-cancerous tissues on an epigenetic level as well.

7. Thesis Objectives

To identify epigenetic changes related to breast cancer treatment and their relationship to persistent inflammation and CRF in patients, Smith et al. examined DNA methylation in peripheral blood mononuclear cells (PBMCs) isolated from breast cancer patients undergoing treatment at Winship Cancer Institute between 2010 and 2011 after obtaining Emory Institutional Review Board approval. Women between ages 18-75 with stage 0-IIIA breast cancer were eligible. In total, 61 patients were treated with standard breast conserving surgery and lymph node evaluation. Depending on stage and subtype, 22 patients received neoadjuvant (n=15) or adjuvant (n=7) chemotherapy prior to radiotherapy and enrollment. Patients with medical conditions that may interfere with the relationship between fatigue and inflammation were excluded. These conditions include chronic infections, uncontrolled cardiovascular disease, mental issues such as bipolar disorder and schizophrenia, and medications that affect the immune system (Smith et al.,).

Smith et al. discovered that patients who received chemotherapy exhibited significantly lower methylation at 8 CpG sites compared to those who did not receive chemotherapy. Lower methylation at each CpG site was further associated with increased inflammatory markers, and increased inflammatory markers were correlated with fatigue symptom. The comparatively lower methylation level in the chemotherapy treated group, the association between lower methylation level and increased inflammatory markers, and the correlation between increased inflammatory markers and fatigue symptoms all persisted after 6 months with moderate decay (Smith et al., 2014). Interestingly, 6 of the 8 CpG sites identified by Smith et al. are located in or near important regulatory genes such as USP2, SMAD3, VMP1, and miR-21. This provides 3 regions of interest for further investigation.

Importantly, although Smith et al. examined 2 time points, both were cross-sectional studies where DNA methylation levels were aggregated and compared between chemotherapy and non-chemotherapy groups. There were 2 important issues: 1, although the non-chemotherapy group served as a negative control, chemotherapy was already administered prior to the baseline time point, meaning that the chemotherapy group could not compare its post-chemotherapy state to its pre-chemotherapy state to infer any possible causal changes induced by chemotherapy; 2, there could be confounding factors such as prior DNA methylation changes from a different source in the participants that was not directly relevant to chemotherapy.

To follow up on Smith et al., the current study takes advantage of an ongoing longitudinal study designed to address the relationship between therapy, epigenetic changes, inflammation level, and CRF in the same breast cancer patient cohort. Smith et al. has covered the baseline and 6-month post treatment time points, and the current study focuses on the 1-year post treatment time point with 2 primary objectives.

First, the current study seeks to answer 3 research questions:

- Whether comparatively lower methylation level in the chemotherapy treated group observed by Smith et al. persists 1 year after treatment;
- Whether comparatively lower methylation level, if present, remains correlated with chemotherapy status and inflammatory marker levels;
- If there are individual patients with methylation data from all 3 time points, what their methylation trajectories may indicate.

A second goal is to design a procedure that would allow for robust and accurate measurement of DNA methylation levels that can be used in future studies with similar design and purpose.

Methods

1. Selecting Methods

The basic algorithm to select appropriate methods begins with the objective and available material. Based on the purpose and scope of the study, there are well established general approaches (Bowman et al., 2013) to be modified to meet the needs of the study. Some important factors to consider include the specificity of the DNA methylation inquiry, the desired robustness, the amount of starting material available, and last but not least, the cost-effectiveness of the options.

Since the original study by Smith et al. already performed a genome-wide study of methylation and identified 6 particular CpG sites in 3 candidate genes, the current study can proceed directly to focus on these regions of interest and increase the robustness and throughput level of the DNA methylation analysis.

For reference in regards to the typical algorithm used to determine the appropriate methodology, one can see Kurdyukov and Bullock (Kurdyukov and Bullock, 2016).

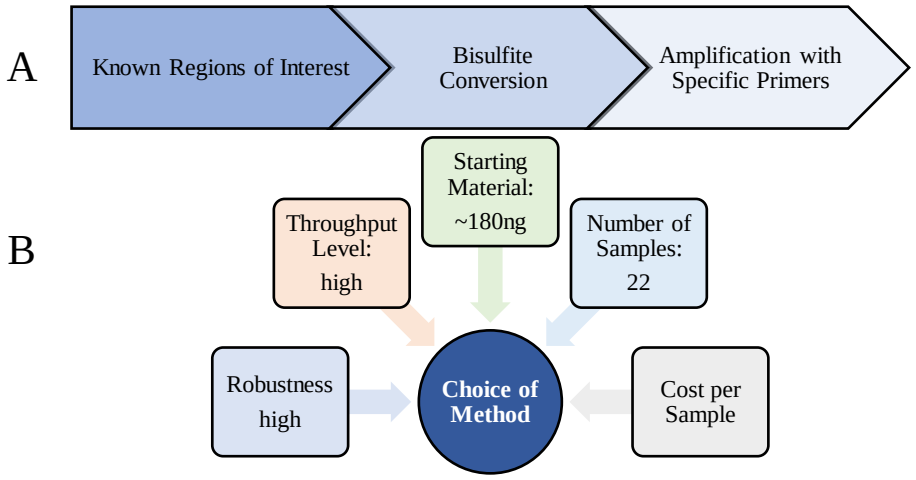

Figure 4: Development of methodology. A. Since the regions of interest are known, there is no need to profile the whole genome and search for differentially methylated areas. The current study can proceed directly to bisulfite treat the regions of interest and amplify them with custom primers. B. To determine the kits necessary for the procedure, a number of factors must be considered to optimize the cost-effectiveness of the method. The current study has a limited amount of available starting material and a relatively small sample size, allowing for robust and high throughput methods that may be too expensive on a larger scale project.

2. Targeted Bisulfite Sequencing

Genomic DNA of peripheral blood mononuclear cells was collected and purified from breast cancer patients by Winship Cancer Institute. The genomic DNA was then bisulfite treated with EpiTect Bisulfite Kit (Qiagen #59104) and amplified for 30 cycles using JumpStart Taq (Sigma P2893) and primers complementary to 3 selected regions of interest covering USP2, SMAD3, and VPM1 (Supplementary Table 1). The amplified product was fragmented with NEB dsDNA Fragmentase (#M0348A) and purified with 1.5x SPRI beads (Kapa #KK8002) following manufacturer's directions. The product was then end-repaired and A-tailed and custom TruSeq compatible sequencing adapters synthesized by IDT (Supplementary Table 2, Barwick et al., 2016) were ligated using the Kapa Hyper Prep Kit (#KK8501). Libraries were amplified for 8 cycles using custom primers and Hifi HotStart ReadyMix polymerase (Kapa #KK8501) (See Figure 5). Adaptor-ligated libraries were quality controlled on an Agilent Bioanalyzer and sequenced on an Illumina HiSeq 4000 using 150bp paired-end sequencing at NYU Genome Technology Center.

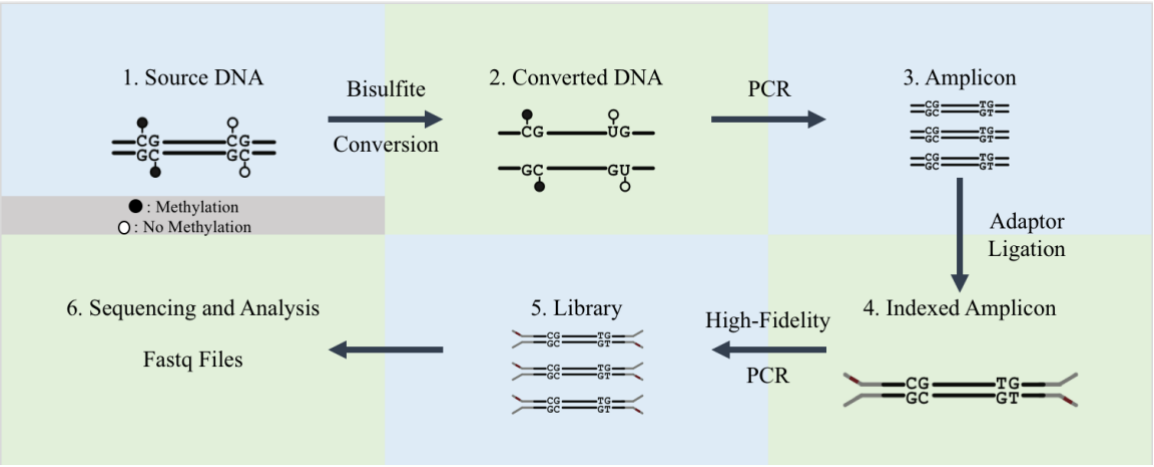

Figure 5: Workflow of Targeted Bisulfite Sequencing. Black circle indicates methylation. White circle indicates no methylation. Unmethylated cytosine is converted to uracil, while methylated cytosine is intact. PCR converts uracil to thymidine. Forked adaptors (adaptor sequence in grey; bar code sequence in red) are ligated to the amplicon for sequencing.

3. Analysis of Targeted Bisulfite Sequencing

The raw sequencing files have low quality reads near the end of the read and adaptor sequences still attached, which should be removed so they do not interfere with the next steps. Fastq sequencing reads were quality trimmed using Trim Galore (Version 0.4.4, available at http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and FastQC (Version 0.11.5, available at https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) (Figure 6 and 7) prior to being mapped to the human genome (GRCh37) using Bismark (Krueger and Andrews, 2011) and Bowtie 2 (Langmead and Salzberg, 2012) (Figure 8). Methylation was called and compiled into percent methylation using Bismark. The annotated coverage files were then visualized with the BeeSwarm package in R (R Core Team, 2017).

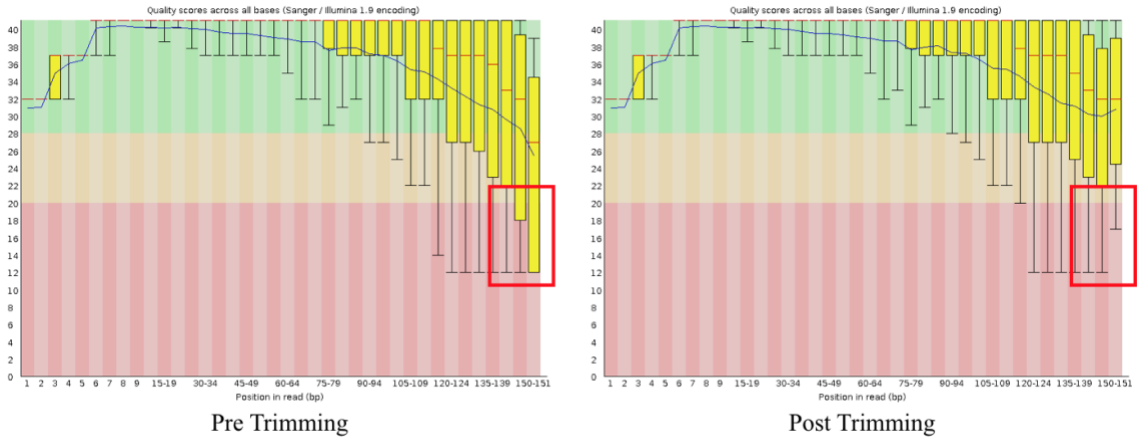

Figure 6: Quality control of raw sequencing files. Reads with low quality scores (<20) can pose threats to future steps including false positives or failure to align reads. The low quality reads that are trimmed are highlighted in the red box. Notice the significant improvement of quality after trimming (vast majority >20).

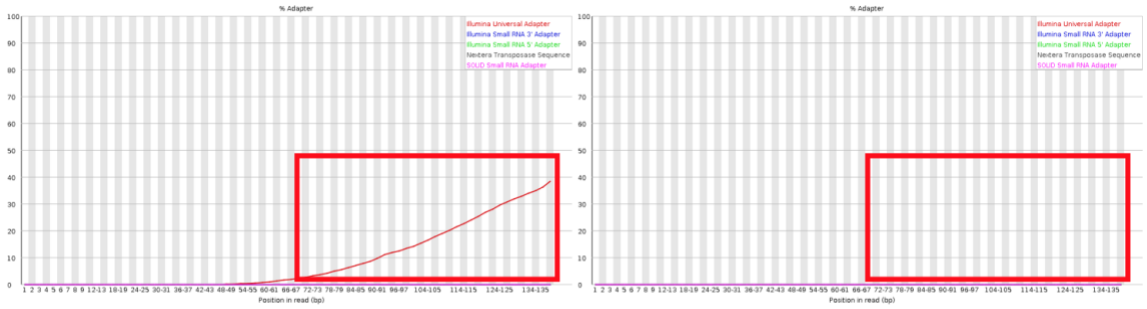

Figure 7: Removing adaptor sequences. Adaptors are artificial sequences that will interfere with read alignment and must be trimmed as well. Here, one can observe that after trimming, all adaptor traces have been removed, as highlighted in the red box (adaptor content ~0%).

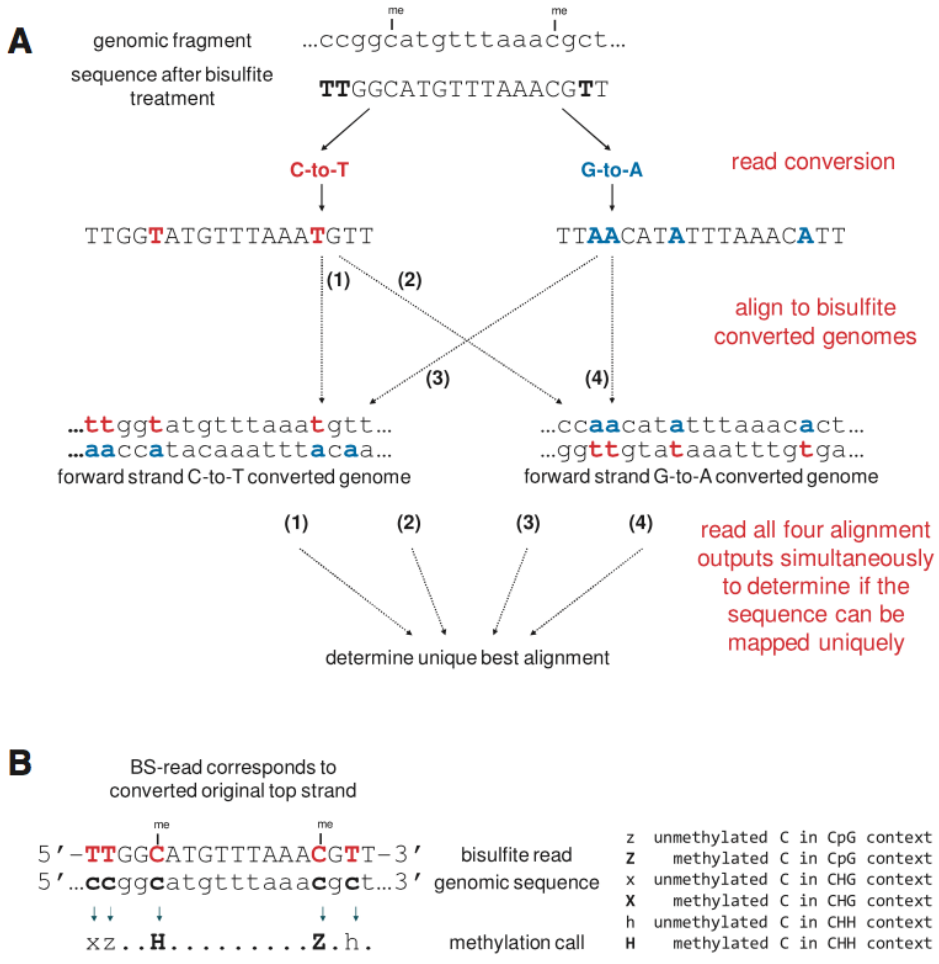

Figure 8: Read alignment and methylation calling. Reads are in silico converted and aligned to bisulfite converted genomes. When the different alignments are compared, one can determine the context and methylation status therein. In the current study, methylation status in CpG context is selected.

4. Statistical Analysis

Non-parametric t-test was used to determine statistical significance due to the small sample size (n=22). To aggregate methylation level by region, average value of percent methylation at each CpG site is calculated. Linear modeling is used to plot the relationship between cytokine concentration and methylation status of cg12054453. All statistical functions are performed in R (R Core Team, 2017). Trajectories are plotted using Excel.

Results

1. CRF Project Overview

Smith et al. identified CpG sites that exhibited significantly lower methylation in the chemotherapy group than in the non-chemotherapy group and associated the lower methylation with increased inflammation markers. Increased inflammation markers were further correlated with fatigue symptoms. Smith et al. observed the same results of associations 6 months post treatment. The current study aims to complete the longitudinal tracking by analyzing data from the same cohort at 1 year post treatment.

In order to accomplish this, multiple sets of data from different sources will be used. It is important to clarify where, when, and how each data set is obtained (Figure 9). Baseline is the time point when all patients (n=61) enrolled into the study. At this point, the chemotherapy group already received either neoadjuvant (n=15) or adjuvant (n=7) chemotherapy, while both the chemotherapy (n=22) and non-chemotherapy (n=39) groups had undergone surgeries already and were waiting for radiotherapy regimen. Among the information collected at this time, surveys and blood work provide the fatigue, inflammation, and DNA methylation assessment. The baseline time point was established by Smith et al. and DNA methylation assessment was completed using Illumina HumanMethylation 450K.

Following the baseline time point is the 6 months post treatment time point. At this point, of the original participants (n=61), a subset (n=39) consented to participate further. The chemotherapy group (n=13) and non-chemotherapy group (n=26) were assessed for fatigue, inflammation, and DNA methylation again through surveys and blood work. The 6 months post treatment time point was established by Smith et al. and DNA methylation assessment was completed using Illumina HumanMethylation 450K.

The current study examines the 1 year post treatment time point. At this point, samples from 22 patients are available, of whom 10 received chemotherapy and 12 did not. However, only 14 of the 22 patients were among the original 61 participants, and only 8 of the patients have complete DNA methylation data from all 3 time points, while missing parts of inflammation marker assessments as well. Of the 8 patients who have complete DNA methylation data, 2 received chemotherapy and 6 did not. Although all the patients belong to the same overarching cohort, the inconsistency in subgrouping may cause errors. Furthermore, the DNA methylation assessment was completed using a new targeted bisulfite sequencing method developed by the lab instead of the Illumina HumanMethylation 450K, which introduces another possible source of error due to inconsistent methodology.

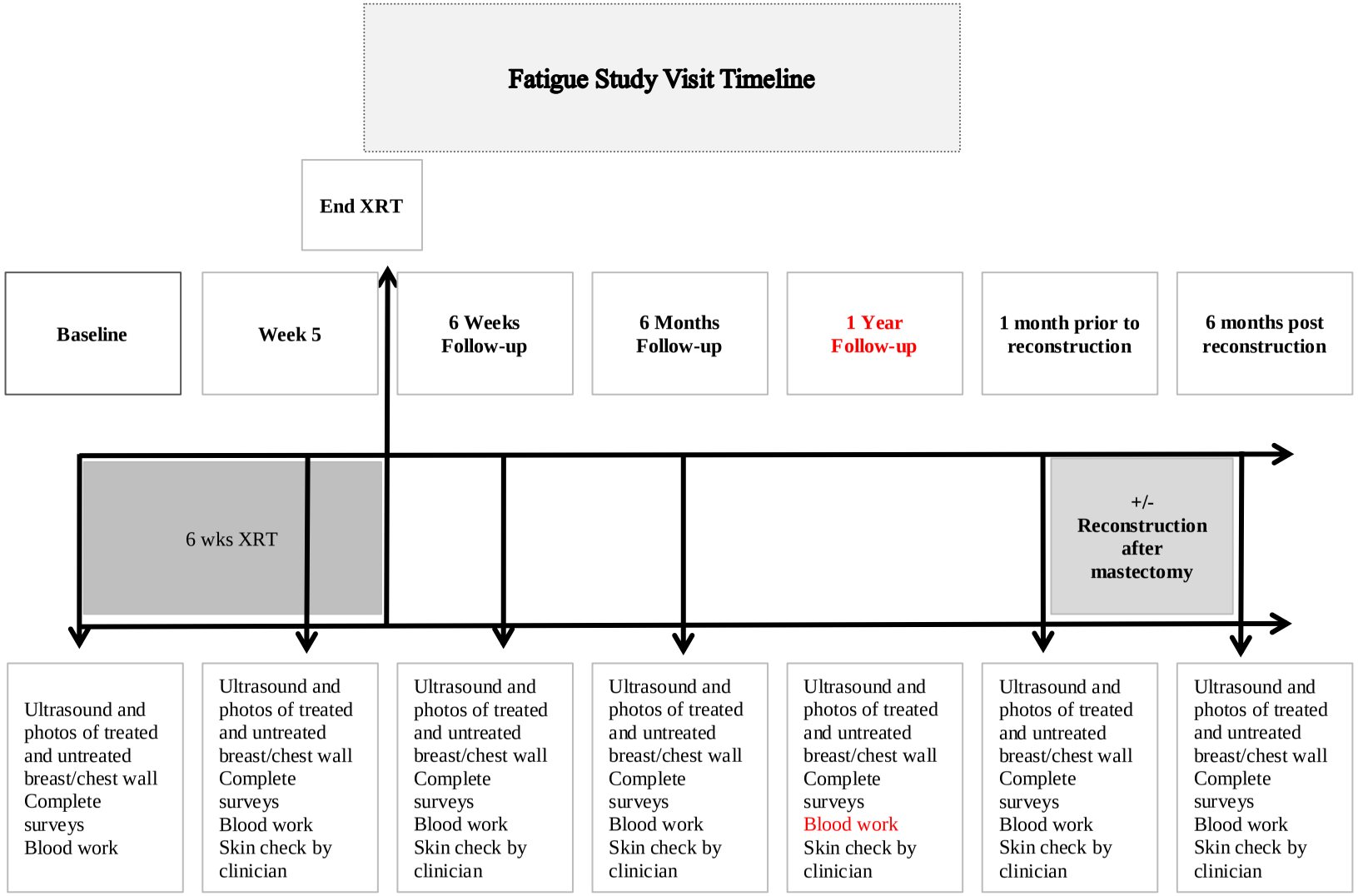

Figure 9: Overview of Cancer Related Fatigue Study with Breast Cancer patients. Previous study by Smith et al. covers the period between baseline and 6-month follow-up. The current study examines the 1-year follow-up time point, which is highlighted in red. The material consists of purified DNA extracted from peripheral blood mononuclear cells, which was drawn at the time of the follow-up. All clinical procedures were completed by Winship Cancer Institute.

2. Regions of Interest

Smith et al. identified 8 CpG sites that showed significant hypomethylation. 6 of the 8 sites are located in 3 genes: ubiquitin specific peptidase 2 (USP2), SMAD3, and vacuole membrane protein 1 (VMP1), also known as TMEM49 (Fig. 10). In particular, the 4 CpG sites located in VMP1 are in close proximity to the reported promotor of miR-21, a miRNA that increases activation of nuclear factor kappa B (NF-κB) and interleukin 6 (IL-6), which are prominent inflammation factors, by downregulating phosphatase and tensin homolog (PTEN) and programmed cell death protein 4 (PDCD4).

Primers were designed to cover adjacent CpGs as well to collect more information on local DNA methylation level, taking advantage of the power of next-generation sequencing. The 6 sites identified by Smith et al. were numbered and matched with both Illumina ID and genomic locations (Table 2). The primers covered 2 CpGs in an intronic region of USP2, 7 CpGs in an intronic region of SMAD3, and 7 CpGs in VMP1, of which CpGs No. 3, 4, and 5 are located within a short exon (Smith et al., 2014).

Illumina ID	Gene	Location	Number
CG26077811	USP2	Chr11:119,232,263	1
CG05438378	SMAD3	Chr15:67,383,736	2
CG16936953	VMP1	Chr17:57,915,665	3
CG12054453	VMP1	Chr17:57,915,717	4
CG01409343	VMP1	Chr17:57,915,740	5
CG18942579	VMP1	Chr17:57,915,773	6

Table 2: Illumina ID and genomic location of the CpGs of interest identified by Smith et al. (Smith et al., 2014)

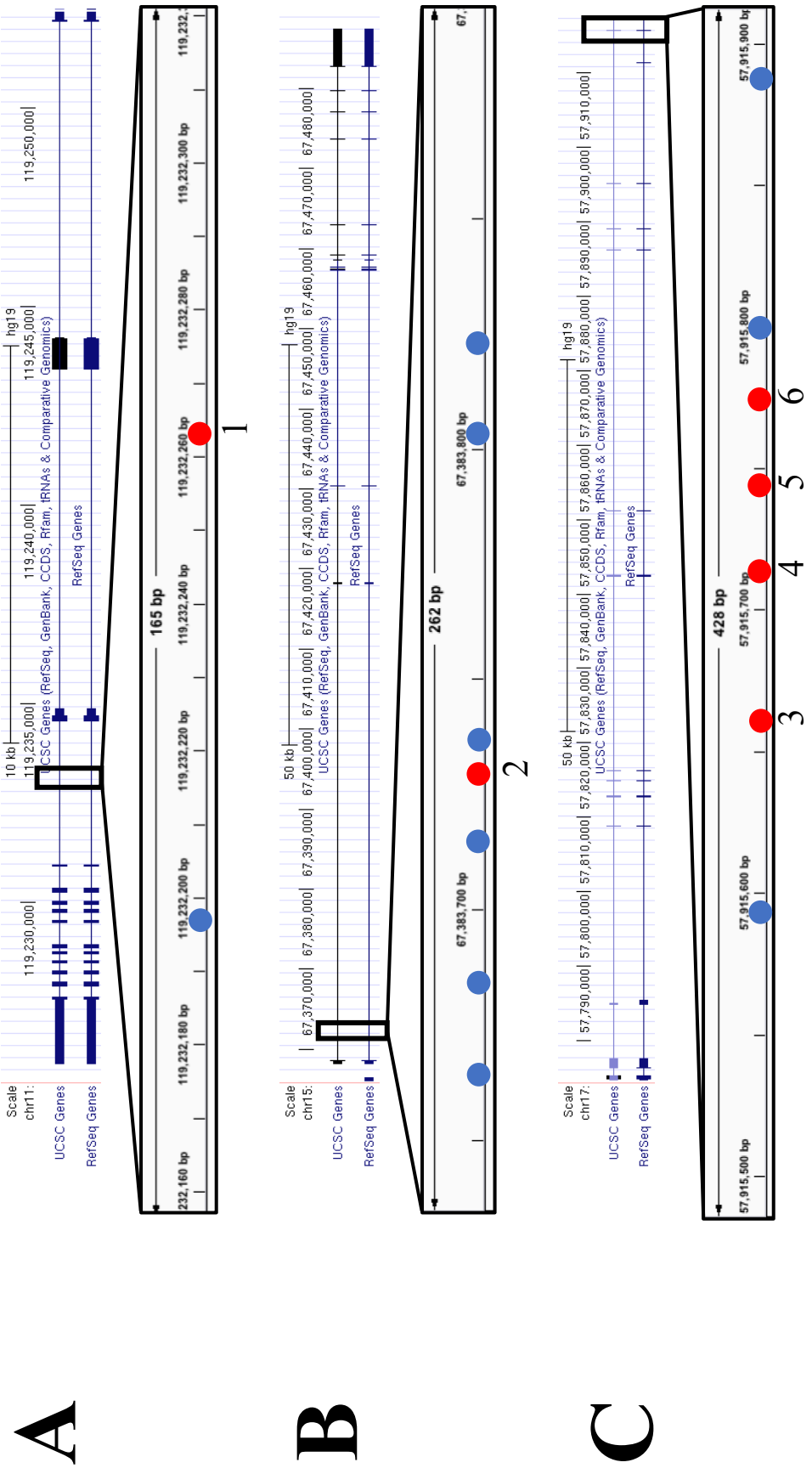

Figure 10: Genomic regions surrounding USP2, SMAD3, and VMP1. Boxed areas are expanded. CpG sites identified by Smith et al. are shown in red and numbered. Additional CpG sites covered by amplicons are shown in blue. A. The region in USP2 (chr11:119,232,157-119,232,321) is in an intron. B. The region in SMAD3 (chr15:67,383,635-67,383,896) is in an intron. C. The region in VMP1 (chr17:57,915,486-57,915,913) covers an exon (chr17:57,915,656-57,915,758). CpG sites 3, 4, and 5 are within the exon.

3. Relationship between methylation level at specific CpG and chemotherapy status

Smith et al. identified CpG sites that exhibited significantly lower methylation in the chemotherapy group than in the non-chemotherapy group at both baseline and 6 months post treatment time points. To determine whether there is still a significant difference in methylation level at specific CpG site between the chemotherapy group and the non-chemotherapy group 1 year post treatment, the methylation level at different CpGs of each patient was calculated, grouped by their chemotherapy status, and then plotted in BeeSwarm plots to visualize the difference in methylation level between the two groups. In total, there are 22 patients, 12 of whom did not receive chemotherapy while 10 did. Out of the 16 CpG sites covered by targeted bisulfite sequencing, 2 sites, both of which lie in an intronic region of USP2, exhibited a significantly lower methylation level in the chemotherapy group (Figure 11). Although the majority of these sites do not show statistically significant difference in methylation, overall, there is a trend towards lower average methylation level in the chemotherapy group than the non-chemotherapy group.

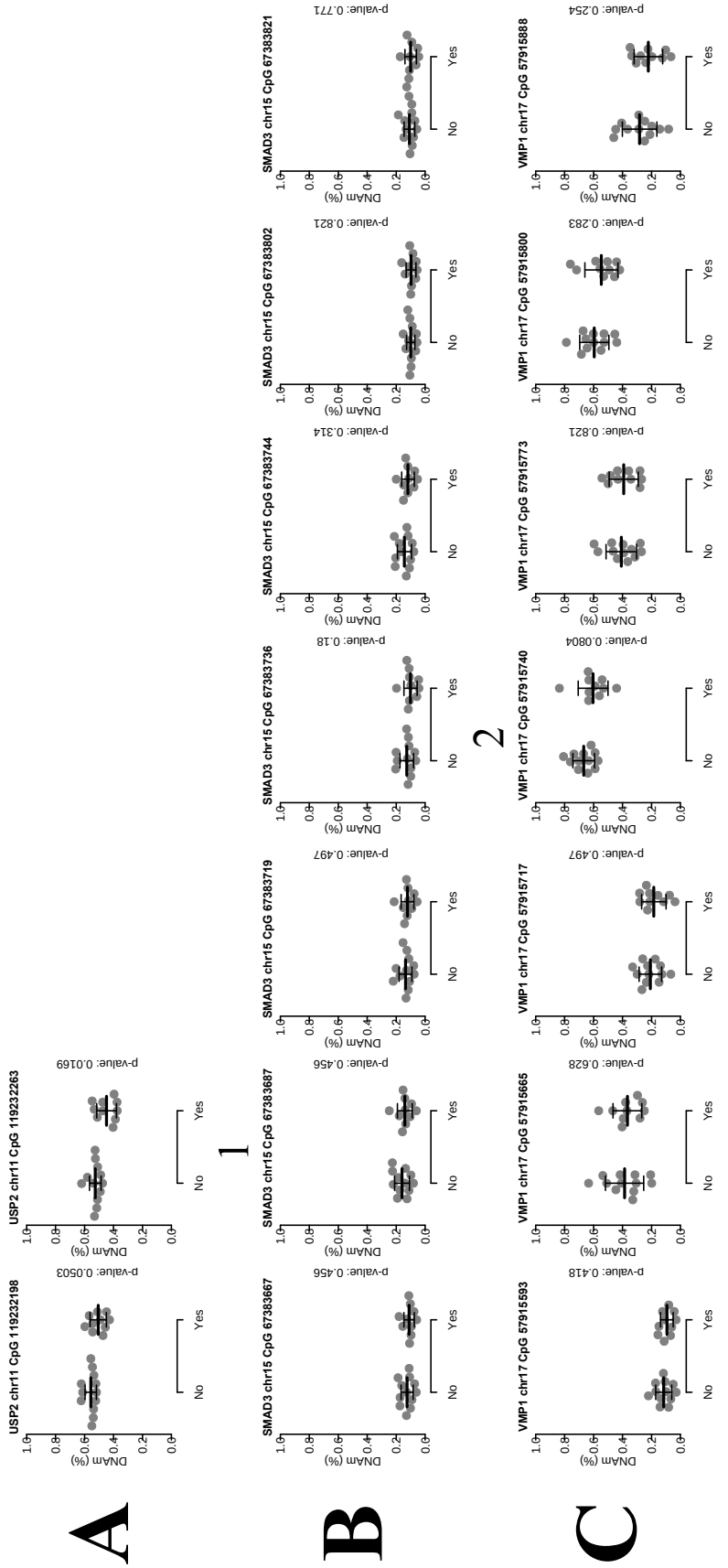

Figure 11: Methylation level at specific CpG sites grouped by patient chemotherapy status (No: did not receive chemotherapy, n=12; Yes: received chemotherapy, n=10). Non-parametric T-Test is used due to small sample size (n=22) and p-value is shown on the right. A. 2 sites in USP2. B. 7 sites in SMAD3. C. 7 sites in VMP1.

4. Aggregated methylation level of the 3 regions of interest grouped by patient chemotherapy status

To examine whether the methylation level of the entire region surrounding each gene might differ depending on the chemotherapy status, the methylation levels at different CpG sites within the same gene were grouped by chemotherapy status and aggregated by calculating the average of the methylation levels of CpGs within the same gene. Overall, the region in USP2, covering 2 CpGs, exhibited a significantly lower methylation level in the chemotherapy group relative to the non-chemotherapy group. The other 2 regions, each covering 7 CpGs, although not statistically significant, still showed moderately lower methylation level in the chemotherapy group compared to the non-chemotherapy group (Fig. 12).

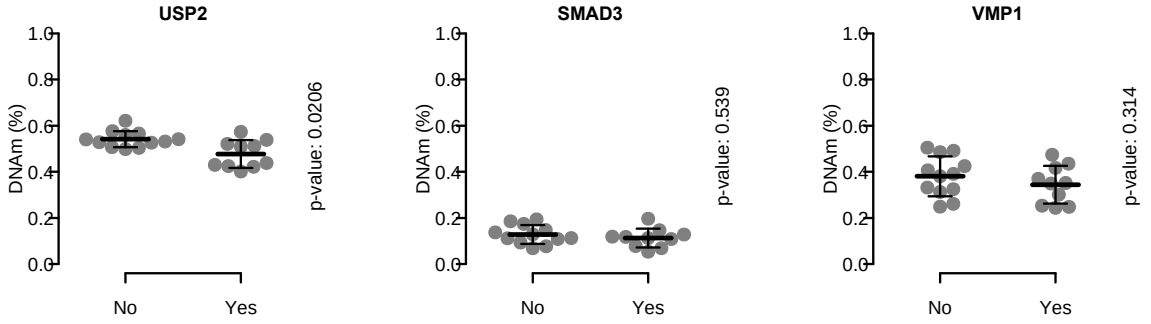

Figure 12: Aggregated methylation level of the 3 regions of interest grouped by patient chemotherapy status (No: did not receive chemotherapy, n=12; Yes: received chemotherapy, n=10). Non-parametric T-Test is used due to small sample size (n=22) and p-value is shown on the right. A. 2 sites in USP2. B. 7 sites in SMAD3. C. 7 sites in VMP1.

5. Relation between methylation level at specific CpG sites and inflammation or CRF

Smith et al. associated the lower methylation at CpGs with increased inflammation markers. Increased inflammation markers were further correlated with fatigue symptoms. To determine whether difference in methylation is still associated with inflammation or CRF 1 year post treatment, methylation at different CpGs of each patient was calculated, grouped by patient fatigue score and inflammation level ranking, and plotted in BeeSwarm plots to visualize the difference in methylation level between the two groups. The ranking was performed by Dr. Mylin Torres's group. The top 11 cases with higher inflammation level and fatigue score were designated the "High" group, while the other 11 subjects with lower inflammation level and fatigue score were designated the "Low" group.

Of the 16 CpG sites, none were statistically significantly correlated with inflammation level or CRF. However the methylation level of CpGs in the High group was consistently lower than that in the Low group (Figure 13).

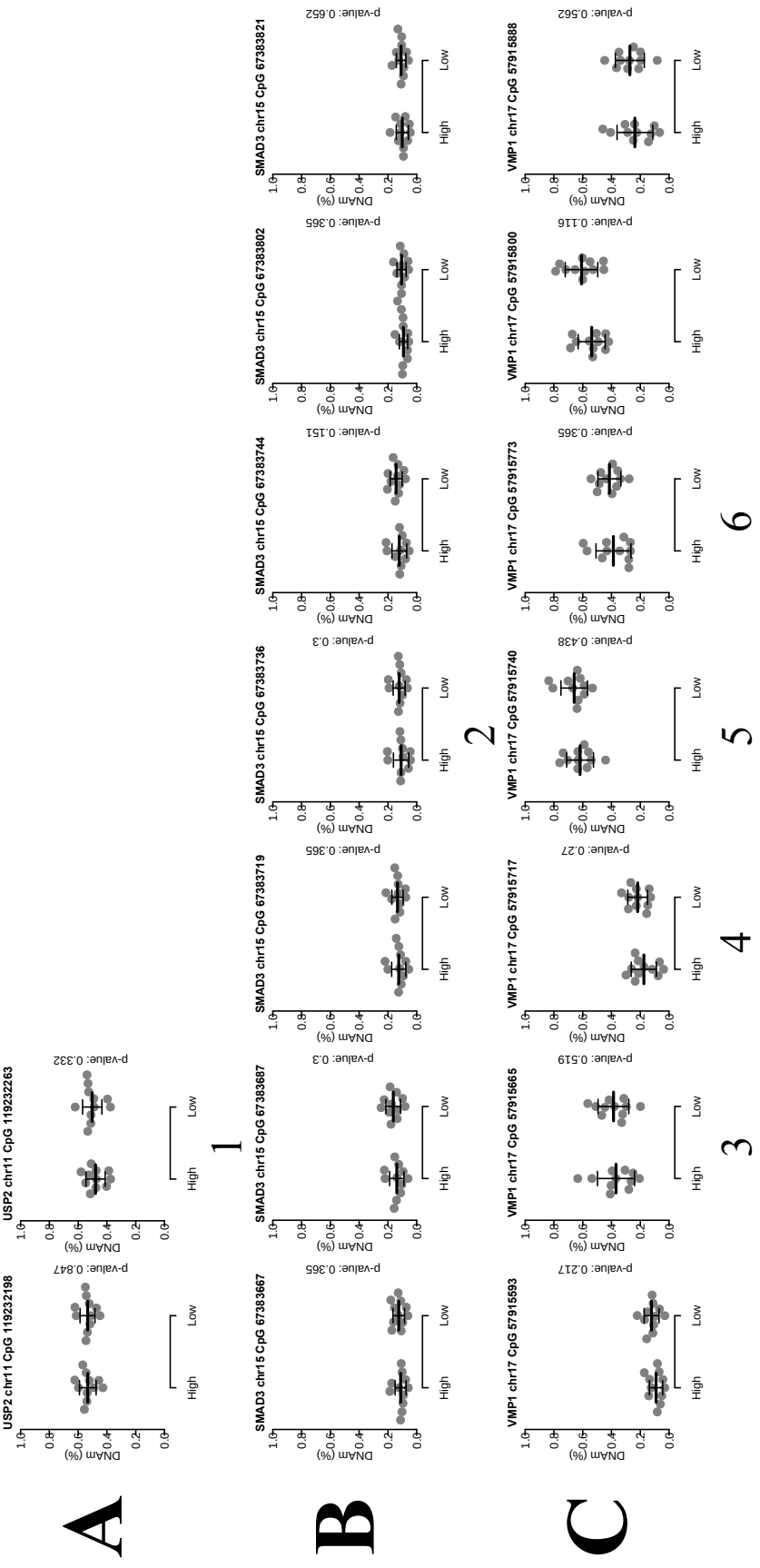

Figure 13: Relation between methylation level at specific CpG sites inflammation or CRF (High: high fatigue score and inflammation level, n=11; Low: low fatigue score and inflammation level, n=11). Non-parametric T-Test is used due to small sample size (n=22) and p-value is shown on the right. A. 2 sites in USP2. B. 7 sites in SMAD3. C. 7 sites in VMP1.

6. Aggregated methylation level of the 3 regions of interest grouped by patient inflammation level and CRF ranking

To examine whether the methylation level of the entire region surrounding each gene might differ depending on inflammation level and CRF ranking, the methylation levels at different CpG sites within the same gene were grouped according to the High versus Low groups and aggregated by calculating the average of the methylation levels of the CpG sites within the same gene. Overall, there was no significant difference in levels of methylation across these genes between the High and Low groups. Nevertheless, there was a trend towards a consistently lower level of methylation in the High group than in the Low group, even though there was no statistical significance (Fig. 16).

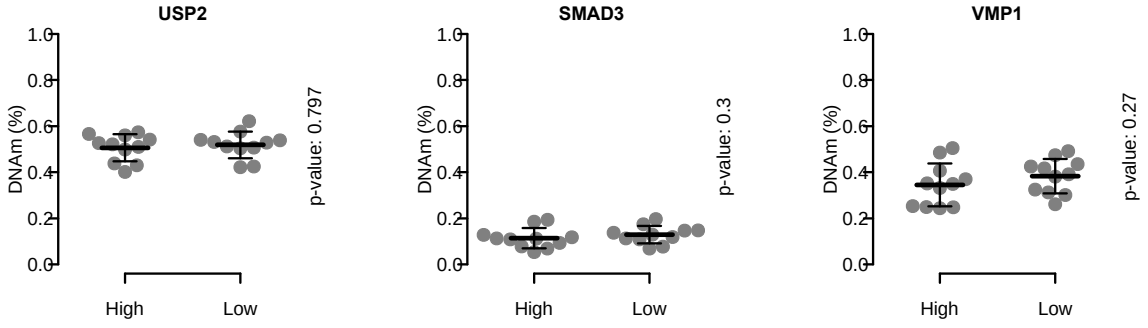

Figure 14: Aggregated methylation level of the 3 regions of interest grouped by patient relative fatigue and inflammation levels (High: high fatigue score and inflammation level, n=11; Low: low fatigue score and inflammation level, n=11). Non-parametric T-Test is used due to small sample size (n=22) and p-value is shown on the right. A. 2 sites in USP2. B. 7 sites in SMAD3. C. 7 sites in VMP1.

7. The temporal changes of methylation levels at different CpG sites

Smith et al. examined the baseline and 6 month post treatment time points. Both were cross-sectional studies where DNA methylation levels were aggregated and compared between chemotherapy and non-chemotherapy groups. So far, the current study has addressed the relation between methylation difference and chemotherapy status as well as the relation between methylation difference and inflammation level or CRF at 1 year post treatment by aggregating individual patients for group comparisons. In order to better understand the impact of chemotherapy, it is necessary to follow methylation level trajectory of individual patients.

Of the 22 patients available at 1 year post treatment, 8 have complete methylation profiles from all 3 time points: baseline, 6 month post treatment, and 1 year post treatment. Their data are plotted into trend-lines (Figure 15).

First, even though the sample size (n=8) is small, the baseline methylation level of the patients is consistent with the findings of Smith et al., where patients who received chemotherapy (in red) showed significantly lower levels of methylation at the 6 CpG sites identified by Smith et al.

Secondly, among those patients who received chemotherapy and for whom samples were available at all time points (n=2), the methylation level was uniformly higher at 6 months post treatment than at baseline, but in most cases returned to baseline, while other patients (n=6) were trending in different directions. At CpG Chr17:57,915,665, however, the methylation level of the patients who received chemotherapy did not return to baseline and continued to increase from 6 months post treatment.

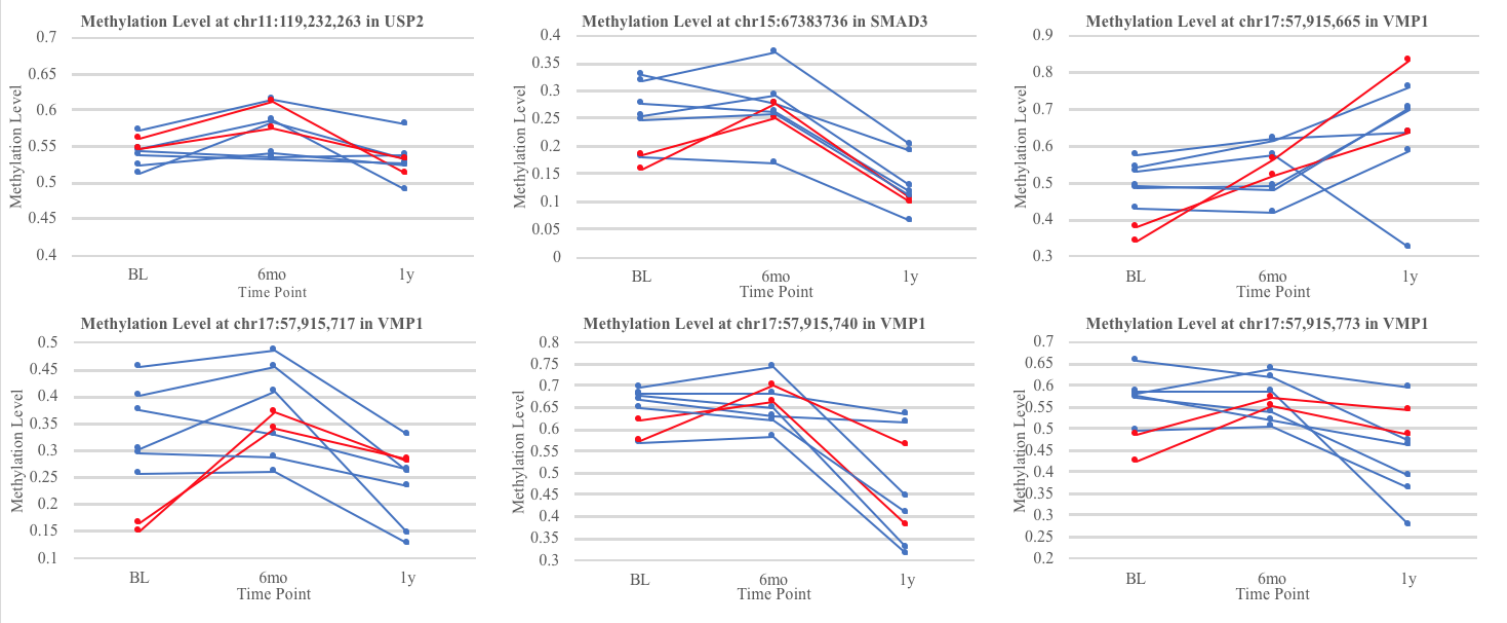

Figure 15: The temporal changes of methylation levels at different CpG sites. There are 8 patients whose methylation data sets are complete. The 1-year follow-up data are collected from a different platform (Targeted Bisulfite Sequencing) than the baseline and 6-month follow-up data (Illumina HumanMethylation 450K). Future study requires consistent methodology in order to be more precise.

8. Relationship between methylation level at specific CpG sites and cytokine level

In Smith et al., lower methylation at each identified CpG site was associated with higher interleukin 6 (IL-6) and plasma soluble tumor necrosis factor receptor 2 (sTNFR2), while sTNFR2, but not CpG methylation status, was correlated with fatigue (Smith et al., 2014).

To examine the relation between methylation level at CpGs and cytokine level at 1 year post treatment, methylation level at each CpG and the cytokine level were plotted using the linear regression model. At 1 year post treatment, lower methylation at each identified CpG site was no longer associated with higher IL-6 or sTNFR2 levels (For IL-6, see Figure 16; For sTNFR2, see Figure 17), although sTNFR2 was still correlated with fatigue. It is important to note the population difference. Only 10 patients from the 22 available at 1-year follow-up had available circulating cytokine levels.

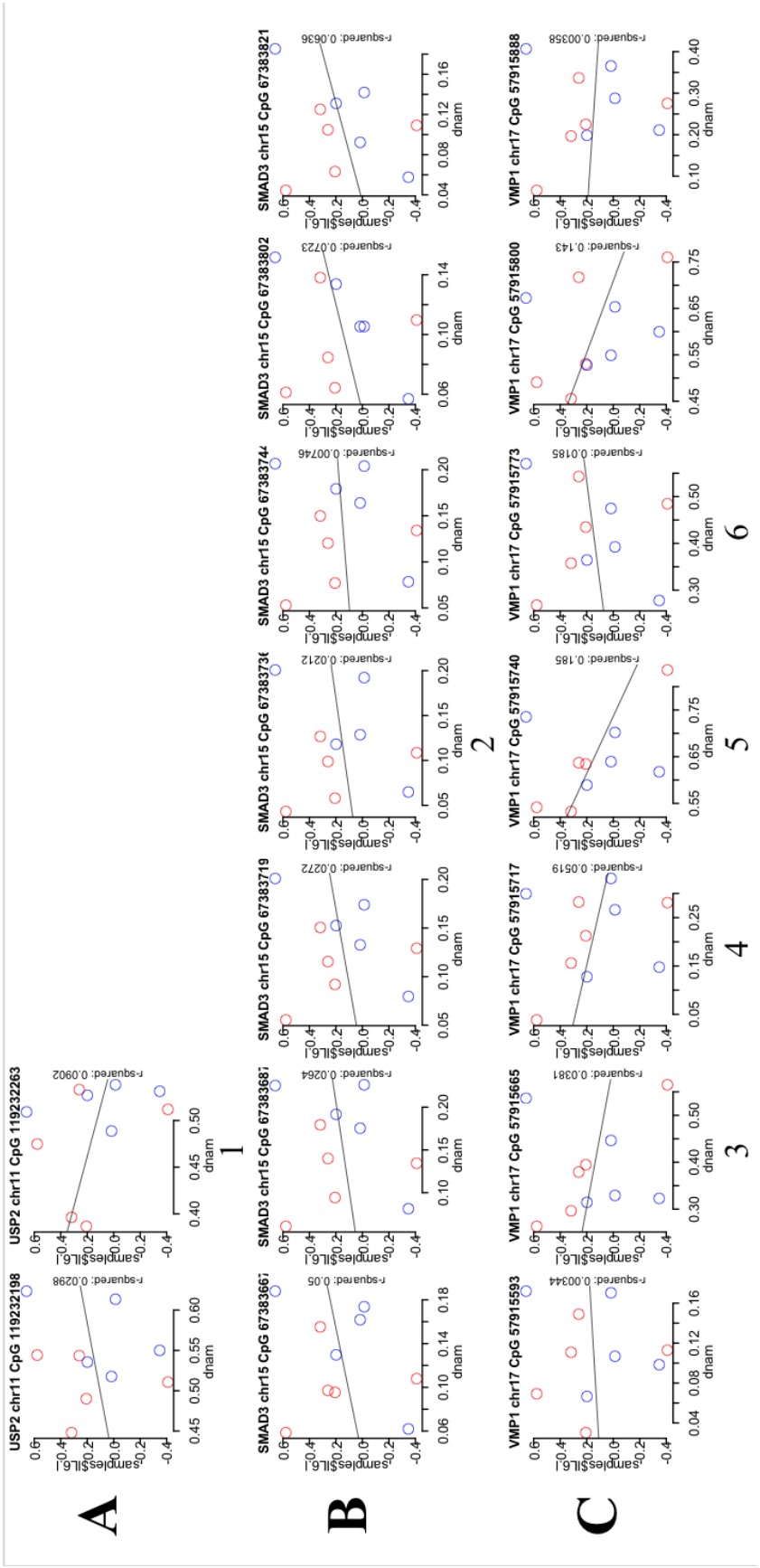

Figure 16: Relationship between methylation level CpGs and patient IL-6 level (n=10) at 1 year post treatment. Red circles: received chemotherapy (n=5); Blue circles: did not receive chemotherapy (n=5).

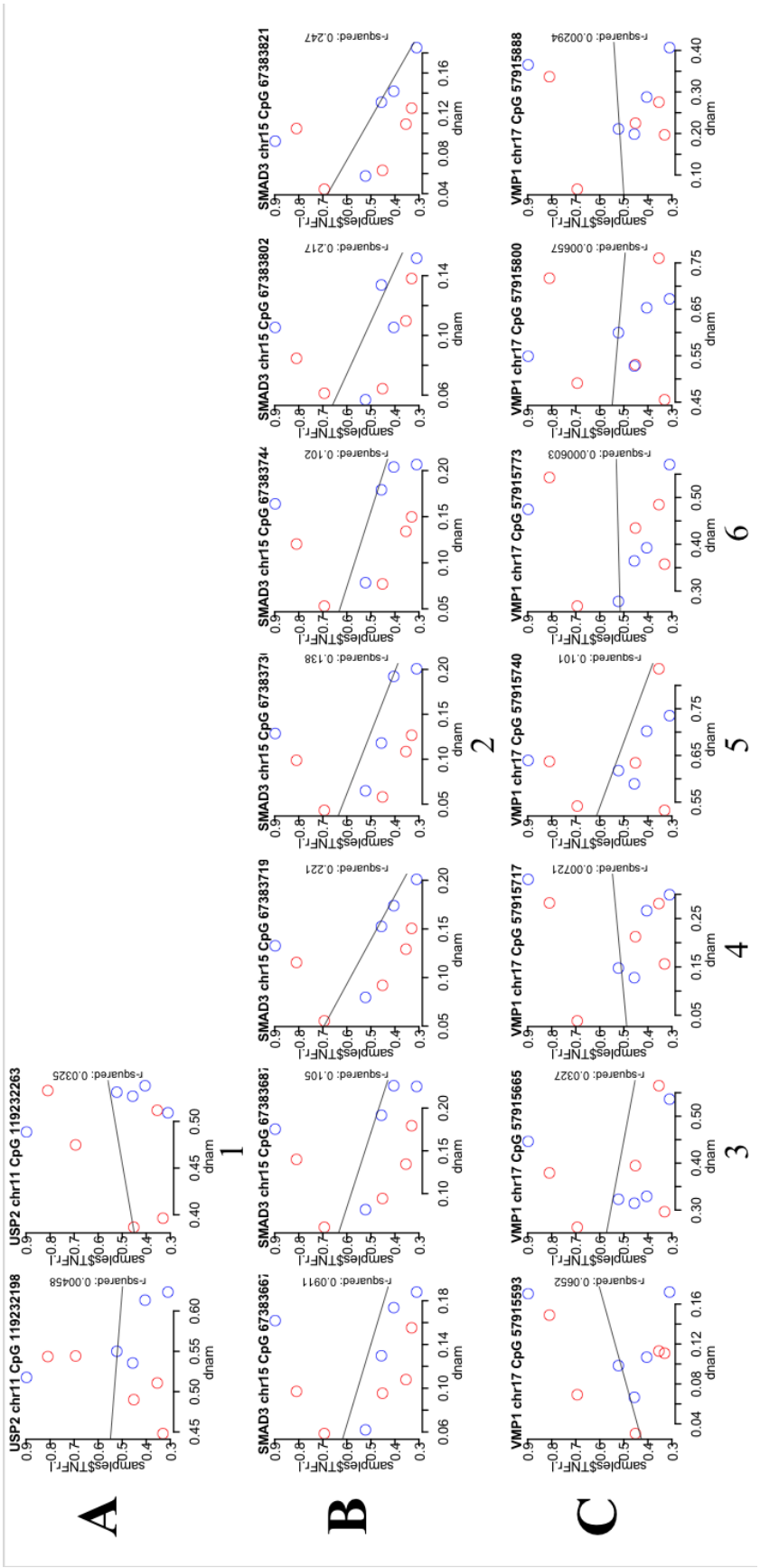

Figure 17: Relationship between methylation level CpGs and patient sTNFR2 level (n=10) at 1 year post treatment.
Red circles: received chemotherapy (n=5); Blue circles: did not receive chemotherapy (n=5).

9. Methylation differences between chemotherapy vs. non-chemotherapy groups

To compare the difference in methylation level at the 6 selected CpG sites between the chemotherapy group and the non-chemotherapy group at all 3 time points, Table 4 from Smith et al. was modified to include data from the current study and shown in Table 3. The methylation level was still consistently lower in the chemotherapy group than the non-chemotherapy group. However, at most of the sites, the difference was no longer statistically significant at 1 year post treatment. Only CG26077811 (CpG Chr11:119,232,263) in USP2 shows a significant decrease of methylation at 1 year post treatment between patients who received chemotherapy and those who did not.

CpG Site*	Gene	Baseline** (n=61)		6 Months Post Treatment* (n=39)		1 Year Post Treatment (n=22)	
		Δβ	p	Δβ	p	ΔMe***	p
cg26077811	USP2	-0.08	3.2×10^{-6}	-0.028	0.045	-0.078	0.017
cg05438378	SMAD3	-0.10	3.4×10^{-6}	-0.027	0.12	-0.026	0.18
cg18942579	VMP1	-0.22	3.6×10^{-15}	-0.060	0.0027	-0.016	0.82
cg12054453	VMP1	-0.19	3.9×10^{-9}	-0.088	8.4×10^{-4}	-0.024	0.50
cg16936953	VMP1	-0.21	1.0×10^{-10}	-0.061	0.014	-0.019	0.63
cg01409343	VMP1	-0.18	1.5×10^{-12}	-0.042	0.031	-0.065	0.08

*Smith et al. labeled CpGs using Illumina IDs. The current study uses genomic locations. See Table 2 for matching information.
**Data from Smith et al.
***Targeted bisulfite sequencing data do not provide Δβ. The difference is calculated and presented as ΔMe.

Table 3: Methylation differences (Δβ and ΔMe) between patients who received chemotherapy and those who did not at baseline (n=39), 6-months follow-up (n=39), and 1-year follow-up (n=22).

Discussion

1. Interpretation

Smith et al. discovered that patients who received chemotherapy exhibited significantly lower methylation at 8 CpG sites compared to those who did not receive chemotherapy, 6 of these 8 being located in important regulatory genes USP2 (1), SMAD3 (1), and VMP1 (4). Lower methylation at each CpG site was further associated with increased inflammatory markers, which, but not CpG methylation status, was correlated with fatigue symptom. The comparatively lower methylation level in the chemotherapy treated group, the association between lower methylation level and increased inflammatory markers, and the correlation between increased inflammatory markers and fatigue symptom all persisted after 6 months with moderate decay (Smith et al., 2014).

The current study designed new targeted bisulfite sequencing method to focus on specific regions of interest and designed primers covering 10 CpGs adjacent to the 6 CpGs identified by Smith et al. to collect more information on local DNA methylation level. The current study then analyzed the methylation level at 1 year post treatment and compared it to available methylation and cytokine data from Smith et al. and additional inflammation and CRF data from Dr. Mylin Torres's group.

The current study aimed to address 3 research questions:

- Whether comparatively lower methylation level in the chemotherapy treated group observed by Smith et al. persists 1 year after treatment;
- Whether comparatively lower methylation level, if present, remains correlated with chemotherapy status and inflammatory marker levels;

- If there are individual patients with methylation data from all 3 time points, what their methylation level trend line may indicate.

1. Comparatively Lower Methylation Level in the Chemotherapy Treated Group Persisted

The average methylation level of the CpGs among patients who received chemotherapy is uniformly lower than the average methylation level of the same CpGs among patients who did not receive chemotherapy. The presence of a clear trend is consistent with the findings of Smith et al., who observed lower methylation level at both baseline and 6 months post treatment time points among the patients who received chemotherapy.

2. The Relation between Comparatively Lower Methylation Level and Relevant Factors

a. Chemotherapy Status

Smith et al. observed that at 6 months post treatment, the difference in methylation level between patients who were treated with chemotherapy and those who were not underwent decay. The result from the current study is consistent with the trend of decay as the methylation level was uniformly lower among the patients who received chemotherapy, but the differences are no longer statistically significant at all but 1 site (CpG Chr11:119,232,263).

There is a number of reasons how this decay might happen. Research indicates that lifestyle and environmental factors such as diet, physical activity, body weight, and smoking or exposure to smoking can all impact DNA methylation (Lim and Song, 2012). Because the study is based on human subjects, not all if any of the lifestyle or environmental factors can be controlled. The fact that clear trends of methylation difference and decay are still present despite the possible interfering factors suggests the selected CpGs warrant further investigation.

b. Inflammation Level and CRF

Smith et al. found that lower methylation level at each CpG site was associated with increased inflammatory markers that was further correlated with fatigue symptom. The current study did not observe the association between lower methylation level at each CpG site and increased inflammatory markers at 1 year post treatment.

Furthermore, the current study matched methylation level at each CpG site with a combined inflammation and CRF ranking and observed that among patients who had comparatively higher inflammation and CRF, methylation level at every CpG site was uniformly lower compared to those who had comparatively lower inflammation and CRF, although the difference in methylation was not statistically significant.

Both of the findings above, again, are consistent with theory of decay.

3. Methylation Changes in Individuals over Time

The current study aligned available methylation data of 8 patients from all 3 time points (baseline, 6 months post treatment, and 1 year post treatment) and provided a longitudinal overview of the methylation trajectory. The trend lines of individuals provided a new perspective as so far the study mainly focused on cross-section comparisons between aggregated groups.

Based on the study design, the patients who did not receive chemotherapy may appear as a steady negative control with an underlying assumption that there should not be many changes. However, trend lines of individuals from this group (n=6) suggest a large degree of variation among individuals and demonstrate that other factors may influence DNA methylation after treatment.

Additionally, all patients who received chemotherapy and had complete methylation data from all time points (n=2) exhibited the same pattern at all 6 CpGs. Their methylation level

increased at 6 months post treatment and then decreased at 1 year post treatment at all sites except for CpG Chr17:57,915,665, where the methylation level continued to increase after 6 months post treatment. Possible explanations include: 1, chemotherapy caused temporally and reversible changes to methylation at these CpGs; 2, CpG Chr17:57,915,665 may have special significance as 4 of the 6 patients exhibited marked increase in methylation at this site at 1 year post treatment compared to the uniform decline in methylation at all other sites over the same period; 3, chemotherapy may have influenced a population of progenitor cells that gave rise to the PBMCs. Hematopoietic Stem Cells (HSCs) have various lifespans between 10 months and 60 months (Sieburg et al., 2011), and changes to its DNA methylation level may explain the methylation trajectories in its progenies shown by the current study.

2. Limitation

The current study has a number of limitations.

First of all, the current study used a different method to measure DNA methylation than Smith et al. Smith et al. used Illumina HumanMethylation 450K while the current study used Targeted Bisulfite Sequencing. They have 2 major differences. Illumina HumanMethylation 450K provides limited coverage at each CpG but does not over amplify a single allele, while Targeted Bisulfite Sequencing provides much greater coverage with over 100,000 reads at each CpG but risks over amplifying a single allele. Both methods are robust and reliable, although the current study did not compare them in a controlled experiment.

Secondly, datasets for individual patients are incomplete for various reasons the current study cannot control. The previous study had good sample size (n>30) and relatively complete data covering fatigue score and multiple circulating cytokines at baseline. However, the current study had a less than optimal sample size (n=22), which was exacerbated by the missing cytokine

and methylation data of different patients from different time points. Although all patients who participated in the overarching project belong to the same cohort, the inconsistency in data availability makes attempts at simulating a longitudinal study difficult (see Figure 15, Figure 17, and Figure 18).

Additionally, because the study was based on human subjects, factors that could impact the immune system such as lifestyle, disease history outside of the study duration, exposure to chemicals could not be controlled.

Furthermore, the overarching project on CRF set the baseline time point at the end of chemotherapy, which was after the administration of chemotherapy, while chemotherapy was the principal effector in question. Ideally, baseline time point should be set prior to the administration of chemotherapy or any form of treatment, so researcher could compare pre-treatment and post-treatment methylation levels, inflammation levels, and CRF evaluations.

Finally, peripheral blood mononuclear cells are a heterogeneous collection of cells that includes several blood cell types such as lymphocytes and monocytes. Currently, the methylation level at different CpGs is based on the entire PBMC population covered by the sample. It is possible that DNA methylation in different cell types might be disproportionally influenced by chemotherapy, but neither the cell type nor the proportion could be covered by the current study. The current focus on well differentiated cell types also means that potential upstream changes in the progenitor cells are overlooked or undetected.

3. Future Direction

First and foremost, a verification study is necessary to compare the Illumina HumanMethylation 450K and Targeted Bisulfite Sequencing methods. Although it is unlikely to

have caused errors that would significantly change the findings, any potential differences in their performances will help future decision-making on methodology.

Here I propose a study design (Figure 18) that will address the limitations and more importantly, the unanswered questions.

Methodologically, since there are clearly defined regions of interest within the genes USP2, SMAD3, and VMP1 already, the new study should use a single consistent method over its course.

Regarding the material, the new study should recruit a greater number of participants to account for possible dropouts. Ideally, the study should have more than 30 patients, of whom half have received chemotherapy, with full methylation profiles over the course of the study.

In addition, although it may be difficult to obtain hematopoietic stem cells for study, efforts should still be made to cover, for example, the myeloid and lymphoid progenitors that may pass epigenetic configurations during differentiation, to aid the search for the origin of any methylation changes and to better understand the mechanism of CRF. Research suggests that DNA methylation plays a role in genomic imprinting and is heritable across cell divisions (Plass and Soloway, 2002; Trerotola et al., 2015). It will be important to collect methylation level at CpG sites of interest from blood cells at different stages of differentiation to locate the primary change and trace the possible inheritance of the change in methylation.

Furthermore, the new study will set the baseline time point prior to any form of treatment, so that both pre-treatment and post-treatment methylation levels, inflammation levels, and CRF evaluations will be available for comparison.

Finally, if the study design above confirms the significance of particular CpG sites, one can search for transcription binding sites in close proximity and determine whether the CpG site

methylation has potential impact on cellular functions, which will serve as a step-stone to uncovering the molecular mechanism of CRF.

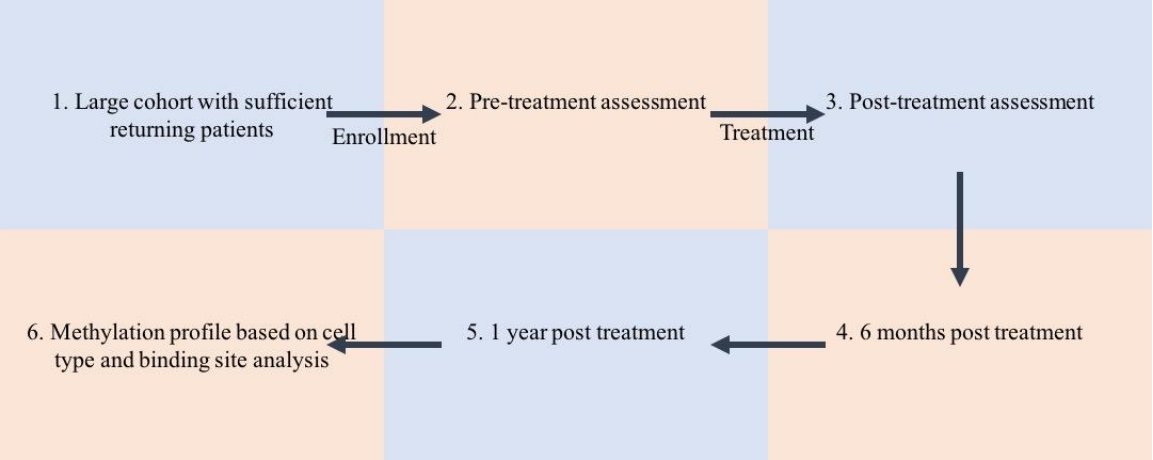

Figure 18: Proposed Study Structure. 1: The new study will aim to secure 30 patients with complete data from all time points; 2: the new study will assess pre-treatment methylation level, inflammation, and CRF in all patients; 3, 4, and 5: the new study will assess post-treatment methylation level, inflammation, and CRF shortly after treatment, 6 months after treatment, and 1 year after treatment; 6, the new study will account for cell types and search for binding sites in regions of interest and examine whether differential methylation at the specific CpGs is linked with regulatory activity.

Reference

(To be generated from EndNote)

Ahmed, S., Sami, A., & Xiang, J. (2015). HER2-directed therapy: current treatment options for HER2-positive breast cancer. Breast Cancer, 22(2), 101-116. doi:10.1007/s12282-015-0587-x

Anderson, W. F., Rosenberg, P. S., Prat, A., Perou, C. M., & Sherman, M. E. (2014). How Many Etiological Subtypes of Breast Cancer: Two, Three, Four, Or More? J Natl Cancer Inst, 106(8). doi:10.1093/jnci/dju165

Asselain, B., Barlow, W., Bartlett, J., Bergh, J., Bergsten-Nordström, E., Bliss, J., . . . Zujewski, J. A. Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials. Lancet Oncol, 19(1), 27-39. doi:10.1016/S1470-2045(17)30777-5

Barsevick A, Cella D. NCI Symptom Management and Quality of Life Steering Committee. Cancer-related Fatigue. 2010 Clinical Trial Planning Meeting (CTPM) Executive Summary, April 13-14, 2010. Available at: https://www.cancer.gov/about-nci/organization/ccct/steering-committees/ncorp/symptom-management/sxqol-ctpm-execsum-2010.pdf

Barwick, B. G., Scharer, C. D., Bally, A. P. R., & Boss, J. M. (2016). Plasma cell differentiation is coupled to division-dependent DNA hypomethylation and gene regulation. Nature Immunology, 17(10), 1216–1225. http://doi.org/10.1038/ni.3519

Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E., & Gianni, L. (2016). Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. Nat Rev Clin Oncol, 13(11), 674-690. doi:10.1038/nrclinonc.2016.66

Bortolon, C., Krikorian, A., Carayol, M., Brouillet, D., Romieu, G., & Ninot, G. (2014). Cancer-related fatigue in breast cancer patients after surgery: A multicomponent model using partial least squares - Path modeling (Vol. 23).

Bower, J. E. (2014). Cancer-related fatigue: Mechanisms, risk factors, and treatments. Nat Rev Clin Oncol, 11(10), 597-609. doi:10.1038/nrclinonc.2014.127

Bower, J. E., Ganz, P. A., Irwin, M. R., Kwan, L., Breen, E. C., & Cole, S. W. (2011). Inflammation and Behavioral Symptoms After Breast Cancer Treatment: Do Fatigue, Depression, and Sleep Disturbance Share a Common Underlying Mechanism? J Clin Oncol, 29(26), 3517-3522. doi:10.1200/jco.2011.36.1154

Bowman, S. K., Simon, M. D., Deaton, A. M., Tolstorukov, M., Borowsky, M. L., & Kingston, R. E. (2013). Multiplexed Illumina sequencing libraries from picogram quantities of DNA. BMC Genomics, 14, 466. doi:10.1186/1471-2164-14-466

Burt, L. M., Ying, J., Poppe, M. M., Suneja, G., & Gaffney, D. K. (2017). Risk of secondary malignancies after radiation therapy for breast cancer: Comprehensive results. Breast, 35, 122-129. doi:10.1016/j.breast.2017.07.004

Carey, L. A., Dees, E. C., Sawyer, L., Gatti, L., Moore, D. T., Collichio, F., . . . Perou, C. M. (2007). The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. Clin Cancer Res, 13(8), 2329-2334. doi:10.1158/1078-0432.ccr-06-1109

Chatterjee, K., Zhang, J., Honbo, N., & Karliner, J. S. (2010). Doxorubicin Cardiomyopathy. Cardiology, 115(2), 155-162. doi:10.1159/000265166

Dieci, M. V., Orvieto, E., Dominici, M., Conte, P. F., & Guarneri, V. (2014). Rare Breast Cancer Subtypes: Histological, Molecular, and Clinical Peculiarities. Oncologist, 19(8), 805-813. doi:10.1634/theoncologist.2014-0108

Early Breast Cancer Trialists' Collaborative Group (EBCTCG). (2011). Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10 801 women in 17 randomised trials. Lancet, 378(9804), 1707–1716. http://doi.org/10.1016/S0140-6736(11)61629-2

Ernst, B., & Anderson, K. S. (2015). Immunotherapy for the treatment of breast cancer. Curr Oncol Rep, 17(2), 5. doi:10.1007/s11912-014-0426-9

Esteller, M. (2002). CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. Oncogene, 21, 5427. doi:10.1038/sj.onc.1205600

Formenti, S. C., & Demaria, S. (2009). Systemic effects of local radiotherapy. Lancet Oncol, 10(7), 718-726. doi:10.1016/s1470-2045(09)70082-8

Khavari, D. A., Sen, G. L., & Rinn, J. L. (2010). DNA methylation and epigenetic control of cellular differentiation. Cell Cycle, 9(19), 3880-3883. doi:10.4161/cc.9.19.13385

Krueger, F., & Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics, 27(11), 1571–1572. http://doi.org/10.1093/bioinformatics/btr167

Kurdyukov, S., & Bullock, M. (2016). DNA Methylation Analysis: Choosing the Right Method. Biology (Basel), 5(1). doi:10.3390/biology5010003

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4), 357–359. http://doi.org/10.1038/nmeth.1923

Lim, U., & Song, M. A. (2012). Dietary and lifestyle factors of DNA methylation. Methods Mol Biol, 863, 359-376. doi:10.1007/978-1-61779-612-8_23

Malhotra, G. K., Zhao, X., Band, H., & Band, V. (2010). Histological, molecular and functional subtypes of breast cancers. Cancer Biol Ther, 10(10), 955-960. doi:10.4161/cbt.10.10.13879

Matsen, C. B., & Neumayer, L. A. (2013). Breast cancer: a review for the general surgeon. JAMA Surg, 148(10), 971-979. doi:10.1001/jamasurg.2013.3393

Miller, J. L., & Grant, P. A. (2013). The role of DNA methylation and histone modifications in transcriptional regulation in humans. Subcell Biochem, 61, 289-317. doi:10.1007/978-94-007-4525-4_13

Phillips, T. (2008) The role of methylation in gene expression. Nature Education 1(1):116

Plass, C., & Soloway, P. D. (2002). DNA methylation, imprinting and cancer. Eur J Hum Genet, 10(1), 6-16. doi:10.1038/sj.ejhg.5200768

Prat, A., Pineda, E., Adamo, B., Galvan, P., Fernandez, A., Gaba, L., . . . Munoz, M. (2015). Clinical implications of the intrinsic molecular subtypes of breast cancer. Breast, 24 Suppl 2, S26-35. doi:10.1016/j.breast.2015.07.008

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roost, M. S., Slieker, R. C., Bialecka, M., van Iperen, L., Gomes Fernandes, M. M., He, N., . . . Chuva de Sousa Lopes, S. M. (2017). DNA methylation and transcriptional trajectories during human development and reprogramming of isogenic pluripotent stem cells. Nat Commun, 8(1), 908. doi:10.1038/s41467-017-01077-3

Sieburg, H. B., Rezner, B. D., & Muller-Sieburg, C. E. (2011). Predicting clonal self-renewal and extinction of hematopoietic stem cells. Proc Natl Acad Sci U S A, 108(11), 4370-4375. doi:10.1073/pnas.1011414108

Sparano, J. A., Gray, R. J., Makower, D. F., Pritchard, K. I., Albain, K. S., Hayes, D. F., . . . Sledge, G. W. (2015). Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. New England Journal of Medicine, 373(21), 2005-2014. doi:10.1056/NEJMoa1510764

Taylor, C. W., & Kirby, A. M. (2015). Cardiac Side-effects From Breast Cancer Radiotherapy. Clin Oncol (R Coll Radiol), 27(11), 621-629. doi:10.1016/j.clon.2015.06.007

The Cancer Genome Atlas Network. (2012). Comprehensive molecular portraits of human breast tumors. Nature, 490(7418), 61–70. http://doi.org/10.1038/nature11412

Trerotola, M., Relli, V., Simeone, P., & Alberti, S. (2015). Epigenetic inheritance and the missing heritability. Hum Genomics, 9(1). doi:10.1186/s40246-015-0041-3

van Maaren, M. C., de Munck, L., de Bock, G. H., Jobsen, J. J., van Dalen, T., Linn, S. C., . . . Siesling, S. 10 year survival after breast-conserving surgery plus radiotherapy compared with mastectomy in early breast cancer in the Netherlands: a population-based study. Lancet Oncol, 17(8), 1158-1170. doi:10.1016/S1470-2045(16)30067-5

Vandiver, A. R., Idrizi, A., Rizzardi, L., Feinberg, A. P., & Hansen, K. D. (2015). DNA methylation is stable during replication and cell cycle arrest. Sci Rep, 5, 17911. doi:10.1038/srep17911

Yamagata, Y., Szabó, P., Szüts, D., Bacquet, C., Arányi, T., & Páldi, A. (2012). Rapid turnover of DNA methylation in human cells. Epigenetics, 7(2), 141-145. doi:10.4161/epi.7.2.18906

Supplementary Tables

Supplementary Table 1

Primer	Gene	Sequence
B914	USP2.F	GGGATTGGGAGTAGGTAATTTTGGA
B915	USP2.R	CCAACTAACAACCTCATAAACTT
B918	SMAD3.F	GAGTTTGATAGGAGAGATAGGATTT
B921	SMAD3.R	AACCAATCCCTATACTTTCATTTCTA
B884	VMP1.F	TAAAAAGGGGTTATAGAATTTTAGTAG
B885	VMP1.R	ATTTCCCAATAAAACATACACAAAC

Supplementary Table 2

Adaptors:
ACACTCTTTCCCTACACGACGCTCTTCCGATC*T
Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTCAC

Index primers:
CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T

Universal primer:
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T

See adaptors and primers previously used by Barwick et al. (Barwick et al., 2016)