

DREAMing: a simple and ultrasensitive method for assessing intratumor epigenetic heterogeneity directly from liquid biopsies

Thomas R. Pisanic II¹, Pornpat Athamanolap², Weijie Poh³, Chen Chen^{3,4}, Alicia Hulbert³, Malcolm V. Brock³, James G. Herman^{3,*} and Tza-Huei Wang^{1,2,3,5,*}

¹Johns Hopkins Institute for NanoBioTechnology, Baltimore, MD 21218, USA, ²Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21218, USA, ³Cancer Biology Program, The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, MD 21287, USA, ⁴Department of Thoracic Surgery, The Second Xiangya Hospital of Central South University, Changsha, Hunan 410011, China and ⁵Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218, USA

Received May 01, 2015; Revised July 23, 2015; Accepted July 25, 2015

ABSTRACT

Many cancers comprise heterogeneous populations of cells at primary and metastatic sites throughout the body. The presence or emergence of distinct subclones with drug-resistant genetic and epigenetic phenotypes within these populations can greatly complicate therapeutic intervention. Liquid biopsies of peripheral blood from cancer patients have been suggested as an ideal means of sampling intratumor genetic and epigenetic heterogeneity for diagnostics, monitoring and therapeutic guidance. However, current molecular diagnostic and sequencing methods are not well suited to the routine assessment of epigenetic heterogeneity in difficult samples such as liquid biopsies that contain intrinsically low fractional concentrations of circulating tumor DNA (ctDNA) and rare epigenetic subclonal populations. Here we report an alternative approach, deemed DREAMing (Discrimination of Rare EpiAlleles by Melt), which uses semi-limiting dilution and precise melt curve analysis to distinguish and enumerate individual copies of epiallelic species at single-CpG-site resolution in fractions as low as 0.005%, providing facile and inexpensive ultrasensitive assessment of locus-specific epigenetic heterogeneity directly from liquid biopsies. The technique is demonstrated here for the evaluation of epigenetic heterogeneity at p14^{ARF} and *BRCA1* gene-promoter loci in liquid biopsies obtained from patients in association with non-small cell lung cancer (NSCLC)

and myelodysplastic/myeloproliferative neoplasms (MDS/MPN), respectively.

INTRODUCTION

Advances in molecular diagnostic and sequencing technologies continue to shed light on the manifold genetic and epigenetic alterations that occur during the initiation, progression and evolution of cancer. Carcinogenesis is thought to arise from genomic instability due to the gradual accumulation of these alterations that ultimately leads to the development of neoplastic cell populations (1,2). Perhaps the most well-studied epigenetic alteration associated with cancer is DNA hypermethylation (3). Many forms of cancer exhibit DNA hypermethylation within CpG islands of tumor suppressor gene promoters that is distinct from healthy tissue and has likewise been shown to act as a promising biomarker for cancer diagnostics, prognostics and predicting drug response (4,5). Aberrant methylation occurs early during carcinogenesis (3) in a progressive and stochastic manner (4,6), producing daughter cell populations with heterogeneously-methylated epialleles and epigenetic phenotypes (7,8). These cells can later give rise to subclones that contribute to cancer heterogeneity, plasticity and drug resistance that can often lead to complications in therapeutic interventions and corresponding adverse clinical outcomes (9–15). The assessment of tumor heterogeneity may thus provide key advantages in the early detection of cancer, as well as predicting and monitoring responses to targeted chemotherapies (4,9).

The evaluation of tumor heterogeneity requires an ability to both sample and discriminate between the subclonal populations that constitute cancerous tissue at primary and metastatic sites. Toward this end, liquid biopsies of peripheral blood from cancer patients have been shown to con-

*To whom correspondence should be addressed. Tel: +1 412 623 7769; Fax: +1 412 623 7768; Email: hermanj3@upmc.edu
Correspondence may also be addressed to Tza-Huei Wang. Tel: +1 410 516 7086; Fax: +1 410 516 7254; Email: thwang@jhu.edu

tain circulating tumor DNA (ctDNA) (16) and rare epigenetic subclonal populations (7,17) with molecular alterations derived from [cancerous] cell populations throughout the entire body and have consequently been heralded as an ideal means of sampling overall intratumor heterogeneity. Nonetheless, tumor-specific genetic and epigenetic alterations within liquid biopsies are often only present at extraordinarily low fractional concentrations (<0.1%) and their detection remains extremely challenging (18–20). Additionally, as epigenetic silencing (and resulting drug resistance) can be effected by the methylation of even only a few CpG sites (21), the ability to detect variants and correspondingly potentially drug-resistant epiallelic clones at ultra-high resolution is also of particular importance.

While there exist numerous methods for the assessment of methylation in genomic DNA (5), they are fundamentally limited in their ability to discriminate rare epiallelic variants. For example, commonly-employed locus-specific techniques such as methylation-specific PCR (MSP) (22) and its real-time analogs, quantitative MSP (qMSP) (23) and MethyLight (24), are routinely used to detect methylation of specific loci in fractions as low as 0.01%, but can only provide semi-quantitative information regarding a specific [typically fully-methylated] methylation pattern. On the other hand, genome-wide approaches including array-based analyses, such as Infinium BeadArray methods (25), and bisulfite sequencing (26,27) provide broad views of DNA methylation and epigenetic heterogeneity across the genome, but despite tremendous advances in sequencing technology are still not capable of directly quantifying heterogeneously-methylated epialleles in fractions below 0.1% without extraordinary time and expense (28–30). These constraints have thus far precluded the widespread assessment of epigenetic heterogeneity in challenging samples such as liquid biopsies and thereby limited studies primarily to the evaluation of samples containing high fractional concentrations of epiallelic variants such as standard tissue biopsies (31–34) or cell line models (10,35).

Here, we present an alternative approach called DREAMing (Discrimination of Rare EpiAlleles by Melt) that provides a facile and inexpensive means ideally suited to the detection and assessment of epigenetic heterogeneity of ultra-rare epiallelic variants such as those found in liquid biopsies. DREAMing employs some of the basic principles of methylation-sensitive high resolution melt (MS-HRM) (36), but unlike bulk MS-HRM, it is directly quantitative and aimed at the detection and enumeration of ultra-rare epiallelic variants down to the single copy level. DREAMing relies upon two key underlying concepts. The first is that DNA samples can be diluted to a so-called ‘quasi-digital’ level such that no more than two epiallelic species of a given locus exist per reaction volume. The second is that, in a single reaction volume, the HRM profiles of amplicons derived from bisulfite-treated (BST) heterogeneously-methylated epialleles can be readily distinguished from unmethylated epialleles and their respective melt temperatures (T_m) used to quantify the methylation density of the original templates. Thus by plotting the T_m heterogeneity of individual epialleles, the overall epigenetic heterogeneity of the sample can be directly inferred.

MATERIALS AND METHODS

Primer and locus selection

Methylation-independent primer design was adapted from previously published design criteria (36). Briefly, the design constraints were as follows: (i) target regions < 150 bp (shorter is generally better); (ii) inclusion of one or two CpG dinucleotide targets in each primer as far toward the 5′ end of the respective primer as possible; (iii) primer melting temperatures near 65°C and within 1°C of each other; (iv) specificity check through Bisearch (37); (v) no significant hairpin or primer dimer formation; (vi) single-peak melt profiles for both fully methylated and fully unmethylated bisulfite-treated sequences, as determined through uMELT (38). All primers were obtained from Integrated DNA Technologies (IDT).

Extraction and bisulfite treatment of genomic DNA

Unmethylated control genomic DNA was obtained from either DNA extracted from DNMT1/DNMT3b double-knockout (DKO) cells (39), human male genomic DNA (Promega) or Epitect unmethylated control DNA (Qiagen). Enzymatically CpG-methylated HeLa genomic DNA (New England BioLabs) was used as a fully-methylated control.

All genomic DNA was processed according to the ‘Methylation-on-Beads’ (MOB) bisulfite conversion technique as previously published (40). Briefly, for samples requiring extraction, cells or plasma were first digested in a solution containing 3 ml of Buffer AL (Qiagen) and 1 ml of Proteinase K (Invitrogen). DNA was then extracted via precipitation by isopropanol, then purified and washed by a series of magnetic decantation steps. The resulting DNA was bisulfite converted using reagents contained within the EZ DNA Methylation Kit (Zymo Research) according to the MOB protocol, washed by magnetic decantation and eluted into a final volume of 100 μ l.

MS-HRM and DREAMing assay conditions and optimization

Each candidate primer pair underwent a multi-step assay condition optimization process. Master mix containing 16.6 mM $(\text{NH}_4)_2\text{SO}_4$, 67 mM Tris pH 8.8, 10 mM β -mercaptoethanol, 1X Evagreen dye (Biotium), 10 nM fluorescein, 200 μ M of each deoxynucleotide triphosphate (dNTP) and 0.04 U/ μ l of Platinum Taq polymerase (ThermoFisher Scientific) was mixed with magnesium chloride to yield a final working magnesium concentration between 1.7 mM to 6.7 mM. Final reaction volumes for all assays were 25 μ l.

A PCR condition matrix was then created by varying the quantitative PCR (qPCR) annealing temperature and master mix magnesium concentration. qPCR and HRM were performed using a CFX96 Touch Real-time PCR Detection System (Bio-Rad) and analyzed using the accompanying stock software, CFX Manager. For all MS-HRM and DREAMing assays, the qPCR cycling conditions were as follows: 5 min at 95°C, followed by 60 cycles of (95°C for 30 s, 30 s at T_A and a 30-s extension step at 72°C, where T_A is the primer annealing temperature). HRM was performed

immediately after amplification with a temperature range of 60°C–90°C at 0.2°C increments and a 10-s hold time before each measurement. Primer pairs and their respective PCR assay conditions were evaluated through analysis of cycle threshold (Ct) values and melt curve profiles according to two primary criteria: difference in the Ct values and ability to produce clearly distinguishable single-peak melting temperature differences in unmethylated versus methylated controls. Primer pairs and conditions that produced consistent and large differences in Ct value as well as large, distinct differences in melting temperature between the two controls were selected for use. The final, optimized, assay conditions were 2.7 mM Mg²⁺ and T_A = 65.2°C for *BRCA1* and 62.2°C for p14^{ARF}.

Following primer selection and assay condition optimization, analytic sensitivity was evaluated through serial dilution of BST methylated control DNA in negative-control BST unmethylated genomic DNA (NC-BSTDNA). Genomic DNA copy numbers were calculated according to beta-actin standards (see below).

DREAMing background determination and model assay

Baseline assay signal was evaluated via DREAMing on BST DKO, human male, Epitect and synthetic DNA. Approximately 20 000 genomic copy equivalents were diluted into a 96-well microtiter plate. Background methylation signals for BST DKO, human male and Epitect DNA were an average epiallelic fraction of 0.02% for p14^{ARF} and 0.005% for *BRCA1*. Here, the epiallelic fraction of a given sample is defined as the ratio of all fully and heterogeneously-methylated epialleles divided by the total number of epialleles. Background epiallelic fraction signal for synthetic-equivalent BST unmethylated DNA was 0.000% for both p14^{ARF} and *BRCA1* assays. The difference between the baseline methylation level of synthetic and BST Epitect is ostensibly attributable to incomplete bisulfite conversion, an issue common to all bisulfite-based methylation analysis and has been shown to be a function of the conversion protocol, as well as other potential factors including sequence complexity and secondary structures (41).

For the model p14^{ARF} DREAMing assay, a manufacturer-estimated eight copies of a synthetic DNA oligonucleotide equivalent to a BST fully-methylated locus were diluted into an estimated 5000 genomic equivalents of NC-BSTDNA and distributed amongst 87 microtiter wells (the nine remaining wells were used for controls). Wells containing only unmethylated epialleles exhibited melt peaks with melt temperatures under 79°C, while wells containing a copy of the synthetic oligonucleotide were identified by distinct secondary melt peaks at 83.8°C ± 0.2°C. For the *BRCA1* model assay, a manufacturer-estimated 4.2 copies each of synthetic equivalents of BST 20% and 100% methylation density *BRCA1* loci were diluted into approximately 5000 genomic equivalents of NC-BSTDNA and distributed amongst 84 wells (the remaining 12 wells were used for controls). qPCR and HRM were performed as with the bulk assay. Resulting HRM curves were then classified according to melt temperature and assigned a methylation density based upon calculated 80% confidence intervals (see below).

DREAMing for detection of p14^{ARF} epialleles in lung-CT-scan-positive plasma samples

Peripheral blood and tumor biopsy samples from patients that had undergone CT-screening for lung cancer were collected at Johns Hopkins Hospital under research study J06115. All patients provided informed consent according to HIPAA regulation. Patients were independently diagnosed according to standard histological examination of biopsies. Blood samples were collected and transferred into 15 ml conical tubes containing 4 ml of Ficoll buffer. These conical tubes were spun for 10 min at 3000 rpm, and the plasma then transferred into 1.5 ml microcentrifuge tubes. These samples were stored at -80°C prior to isolation of cfDNA via Methylation-on-Beads (MOB), as described above. NSCLC samples were assessed for *CDKN2A* (p16) promoter methylation using a standard MethyLight assay (24) with 300 nM forward primer, 5'- TTA TTA GAG GGT GGG GCG GAT CGC -3', 300 nM reverse primer, 5' - GAC CCC GAA CCG CGA CCG TAA -3', spanning a 151-bp region (chr9:21,974,757–21,974,907), and 100 nM probe, 5'- \56-FAM\ ACT ACT CCC \ZEN\ CRC CRC CRA CTC CAT AC \3IABkFQ\ -3'. Cycling conditions were 95°C for 5 min, followed by 60 cycles of (95°C for 30 s, 60°C for 30 s and 72°C for 30 s). These samples were also assessed for *CDKN2A* (p14^{ARF}) promoter methylation by qMSP with Evagreen dye using a touch-down qMSP approach (42) using primers adapted from previous studies (43) with 300 nM forward primer, 5'- GTG TTA AAG GGC GGC GTA GC -3' and 300 nM reverse primer, 5'- AAA ACC CTC ACT CGC GAC GA -3', spanning a 123-bp region (chr9:21,994,263–21,994,385). Cycling conditions were 95°C for 5 min, followed by 5 cycles of (95°C for 30 s, 69.5°C for 30 s and 72°C for 30 s), 5 cycles of (95°C for 30 s, 67.5°C for 30 s and 72°C for 30 s), 5 cycles of (95°C for 30 s, 65°C for 30 s and 72°C for 30 s), 5 cycles of (95°C for 30 s, 62.5°C for 30 s and 72°C for 30 s) and 30 cycles of (95°C for 30 s, 60°C for 30 s and 72°C for 30 s). Sequences for beta-actin recognizing both methylated and unmethylated templates were: actin-sense 5'- TAG GGA GTA TAT AGG TTG GGG AAG TT -3'; actin-antisense 5'- AAC ACA CAA TAA CAA ACA CAA ATT CAC -3', spanning a 104-bp region (chr7:5,532,168–5,532,271). Cycling conditions were 95°C for 5 min, followed by 50 cycles of (95°C for 5 s, 60°C for 30 s and 72°C for 30 s).

For the NSCLC plasma samples, beta-actin standards were used to estimate overall circulating DNA copy numbers. Many samples exhibited weak p16 MethyLight and p14^{ARF} qMSP amplification profiles and/or did not show amplification in all replicates. Thus, for Ct values greater than 35, methylated-epiallele copy numbers were estimated by counting the fraction of replicates that exhibited *any* amplification, however slight, and dividing by the total number of replicates. Each replicate contained 2% of extracted and BST DNA from a 2 ml plasma sample, thus representing circulating DNA from roughly 40 µl of plasma per replicate. This was then used to estimate the number of copies per volume of plasma for both qMSP and DREAMing assays. Wells in the p14^{ARF} DREAMing assay were considered positive if the melt profile exhibited a peak with a melt temperature of 79°C or higher, which corresponds to a >96% con-

confidence interval based upon the average (78.7°C) and standard deviation (0.129°C) determined by DREAMing assay of 96 wells of NC-BSTDNA.

Pyrosequencing of DREAMing products

In order to perform pyrosequencing, one of the primers was replaced with a biotinylated analogue. Following DREAM analysis, wells of interest were selected for sequencing. The biotinylated PCR product was purified and subjected to pyrosequencing using the PyroMark Q24 System (Qiagen), using forward primers according to the manufacturer's protocol. Sequencing results report the percentage of cytosine versus thymine at each CpG site.

The use of pyrosequencing to verify the methylation densities of DREAMing amplicons represents an atypical sequencing application. In short, prior to amplification, each 'positive' reaction volume (well) ostensibly contains one heterogeneously-methylated epiallele and anywhere from 5 to 500 unmethylated epialleles. Secondly, while the primers are designed to preferentially, but not exclusively, amplify epialleles that are methylated in the primer region, the final ratio of unmethylated to methylated amplicons cannot be precisely controlled. Lastly, while all methylated cytosines from a single epiallele should exhibit the same raw methylation percentage, this rarely occurs in practice, as there is often a considerable amount of noise in pyrosequencing traces that confound analysis. These variables combine to prevent using a single methylation percentage threshold as a means of determining the methylation status of CpG sites for all samples. To deal with this issue, a logic flowchart was developed for evaluating the methylation status of CpG sites on a sample-by-sample basis, the details and an example of which is provided in Supplementary Section 3.1. Wells with primer dimer -dRFU/dT peaks, multiple or very low level heterogeneously-methylated epiallelic peaks were typically unable to be sequenced reliably due to methylation assessment complications and/or lack of pyrosequencing sensitivity and were thus excluded from analyses.

Bulk MS-HRM epiallelic fraction determination

For both p14^{ARF} and *BRCA1*, BST methylated control genomic DNA was serially diluted in NC-BSTDNA to yield an epiallelic fraction ranging from 0.01% to 100%. Following triplicate qPCR (using optimized assay conditions developed above), HRM curves were generated and the absolute value of the negative derivative peaks (-dRFU/dT) at the unmethylated and fully methylated melt temperatures were recorded. The peak heights were averaged and plotted against the epiallelic fraction. Best-fit trendlines were computed for the resulting curves based on power regression analysis.

Statistical framework for quantitative methylation density assessment

Synthetic DNA sequences equivalent to amplified BST-DNA of a single permutation of 0% (unmethylated), 20%, 40%, 60%, 80% and 100% (fully methylated) of the selected *BRCA1* locus were obtained from IDT. DNA for each respective methylation density was diluted to yield 10 000

copies per reaction volume. Following qPCR (using optimized assay conditions developed above), HRM curves were generated and corresponding melt temperatures obtained for each methylation density. Resulting melt temperatures were then compared with *in silico* estimates obtained through uMELT (38) using the Blake & Delcourt thermodynamic library (44). *In silico* and experimental melt temperature values obtained above were then evaluated in order to generate a linear regression.

Melt temperature estimates for all permutations were then calculated via interpolation using the equation for the best-fit trendline of the linear regression. By assuming each permutation of a given methylation density has an equal probability of occurrence, an estimated average melt temperature for each methylation density ($\bar{X}_{permutation, \%methylation}$) was generated, along with the corresponding standard deviation, $s_{permutation, \%methylation}$, which is a composite of the standard deviation of the predicted melt temperatures at each methylation density, $s_{permutation, \%methylation}$, and the standard deviation of $\bar{X}_{permutation, \%methylation}$ from the regression, $s_{regression}$, as:

$$s_{estimation, \%methylation} = \sqrt{s_{regression}^2 + s_{permutation, \%methylation}^2}$$

The measurement standard deviation, $s_{measurement}$, was obtained through 48 replicate assays on NC-BSTDNA and, when combined with the estimation standard deviation was used to generate the overall assay standard deviation, $s_{measurement}$, for each methylation density, as:

$$s_{overall, \%methylation} = \sqrt{s_{measurement}^2 + s_{estimation, \%methylation}^2}$$

Here we assume that $s_{measurement}$ is independent of the sample being measured. Finally, a confidence interval for each methylation density, $\theta_{mean, \%methylation}$, based on the mean of a given number of binned replicates (normally one) from a sample, $n_{replicates}$, was derived from the overall assay standard deviation, as:

$$\theta_{mean, \%methylation} = \bar{X}_{permutation, \%methylation} \pm z^* \frac{s_{overall, \%methylation}}{\sqrt{n_{replicates}}}$$

where z^* is the upper critical value based upon a standard normal distribution.

A step-by-step demonstration of the incorporation of raw melt data into this statistical framework is provided in Supplementary Section 3.2.

DREAMing for detection of *BRCA1* epiallelic variants in MDS/MPN clinical isolates

Peripheral blood samples from patients diagnosed with mixed MDS/MPN were collected at the Johns Hopkins Hospital under research study J0969. All patients provided informed consent according to the Declaration of Helsinki. Mononuclear cells were isolated by density centrifugation with Ficoll-Paque Plus (GE Healthcare). The ZR-DuetTM DNA/RNA MiniPrep kit (Zymo Research) was used to obtain genomic DNA.

BRCA1 promoter methylation was initially analyzed by qMSP (23) using the iTaq SYBR Green mix and 300 nM of each primer. Primers specific for methylated *BRCA1* sequences were: *BRCA1*-methylated sense 5'- GGG TGG TTA ATT TAG AGT TTC GAG AGA CG -3'; *BRCA1*-methylated antisense 5'- AAC GAA CTC ACG CCG CGC AAT CG -3', spanning a 185-bp region (chr17:43,125,361–43,125,545). Sequences for beta-actin recognizing both methylated and unmethylated templates were: actin-sense 5'- TAG GGA GTA TAT AGG TTG GGG AAG TT -3'; actin-antisense 5'- AAC ACA CAA TAA CAA ACA CAA ATT CAC -3', spanning a 104-bp region (chr7:5,532,168–5,532,271). Cycling conditions were 95°C for 5 min, followed by 40 cycles of (95°C for 5 s and 64°C for 60 s). Epiallelic fractions were estimated using cycle threshold values for beta-actin and target gene promoter according to the formula $2^{(\text{ActinCt} - \text{BRCA1Ct})} \times 100\%$.

MDS/MPN patient samples were also preliminarily evaluated in a bulk MS-HRM assay to determine the potential presence of methylated epiallelic variants. The epiallelic fraction from each sample was estimated via the bulk MS-HRM epiallelic fraction method described above according to the number of wells showing anomalous peaks and the respective height of those peaks. Through comparison with methylated:unmethylated control DNA standards, the epiallelic fraction for most samples could be reasonably estimated. For rare epialleles, a reasonable estimate could be obtained by measuring the sample in triplicate and observing the number of replicates exhibiting heterogeneous-methylation peaks. Samples were then diluted according to their respective estimated epiallelic fractions to yield a final working concentration of less than 0.25 methylated epiallelic variants per well. The diluted samples were then mixed with freshly-made master mix, containing a biotinylated analogue of one of the primers (in order to allow subsequent pyrosequencing, as described above), and then pipetted into 84–90 wells of a microtiter plate, along with 6–12 wells containing no-template and/or negative controls. Following qPCR and HRM, resulting HRM curves were then classified according to melt temperature and assigned methylation densities based upon the previously calculated confidence intervals. Epiallelic fractions from the DREAM analysis of each sample were calculated according to the original dilution factor, using beta-actin qPCR to standardize each sample, respectively.

Gaussian mixture modeling and subtraction

For T_m 's that could not be directly identified by CFX Manager software due to low peak-heights relative to a large unmethylated epiallelic peak, the raw melt profile data were exported and analyzed using a MATLAB subroutine based upon a mixed Gaussian model. Unmethylated peaks were automatically fitted to a skewed Gaussian and removed from analysis, leaving only the remaining heterogeneously-methylated epiallelic melt profile and associated peaks. Wells were considered positive if they exhibited a remaining heterogeneous-methylation peak three standard deviations greater than [noise from] the average unmethylated control. Methylation densities were assessed using the previously calculated confidence intervals.

RESULTS

Overview of the DREAMing assay

A schematic illustrating how DREAMing effectively combines quasi-digitization and melt profile analysis to produce an overall assessment of epiallelic/epigenetic heterogeneity directly from a liquid biopsy is shown in Figure 1. Circulating cell-free DNA (cfDNA) is first extracted and bisulfite converted using our previously described MOB technique that results in significantly higher cfDNA recovery and purity when compared with other methods such as phenol chloroform and spin-column-based techniques (40,45). At this point the sample contains a solution of single-stranded BST DNA containing a heterogeneous population of epiallelic variants. Then, using preliminary copy number estimates, the sample is diluted to yield quasi-digital reaction volumes (wells) that contain no more than one partially or fully-methylated epiallelic variant per well. In order to increase signal from heterogeneously-methylated and methylated epialleles, the target loci are amplified with primers that preferentially, but not exclusively, recognize methylated epiallelic variants. Following amplification, HRM is performed in the presence of a DNA-saturating dye such as Evagreen (46) and the resulting negative derivative (-dRFU/dT) curves are compared. Wells containing individual heterogeneously-methylated or fully-methylated epiallelic variants can then be easily identified by the presence of a characteristic secondary melt peak whose respective T_m can be used to identify the methylation density of each original template molecule. Lastly, a 'DREAM analysis' histogram based on the observed frequency of each methylation-density-variant can then be created to allow a clear visual representation of a given sample's epigenetic heterogeneity at the target locus.

Primer selection and assay optimization for DREAMing

DREAMing assays were designed within CpG islands located in the promoter regions of two tumor suppressor genes, *CDKN2A* (p14^{ARF}) and *BRCA1*. These loci were chosen to demonstrate the ability of DREAMing to identify potentially drug-resistant and susceptible epigenetic subpopulations directly from liquid biopsies. For example, epigenetic inactivation of p14^{ARF} results in overexpression of *MDM2* and loss of p53 (47). This has been shown to be a useful prognostic indicator of response of non-small cell lung cancer (NSCLC) to gemcitabine and gefitinib treatment (48–50). Likewise, *BRCA1* is a DNA repair protein whose epigenetic silencing has been shown to correlate to sensitivity to poly(ADP-ribose) polymerase (PARP) inhibition in breast cancer patients (51) and, more recently, patients suffering from myelodysplastic/myeloproliferative neoplasms (MDS/MPN) (52).

Primers were designed by amending previously-published design criteria (36) specifically for the preferential amplification of ultra-rare heterogeneously-methylated species. In particular, one to two CpG sites were included toward the 5' end of each primer to bias amplification against unmethylated sequences (that contain mismatches at the primer CpG sites) resulting in preferential amplification of epialleles that exhibit methylation in the primer regions of the locus.

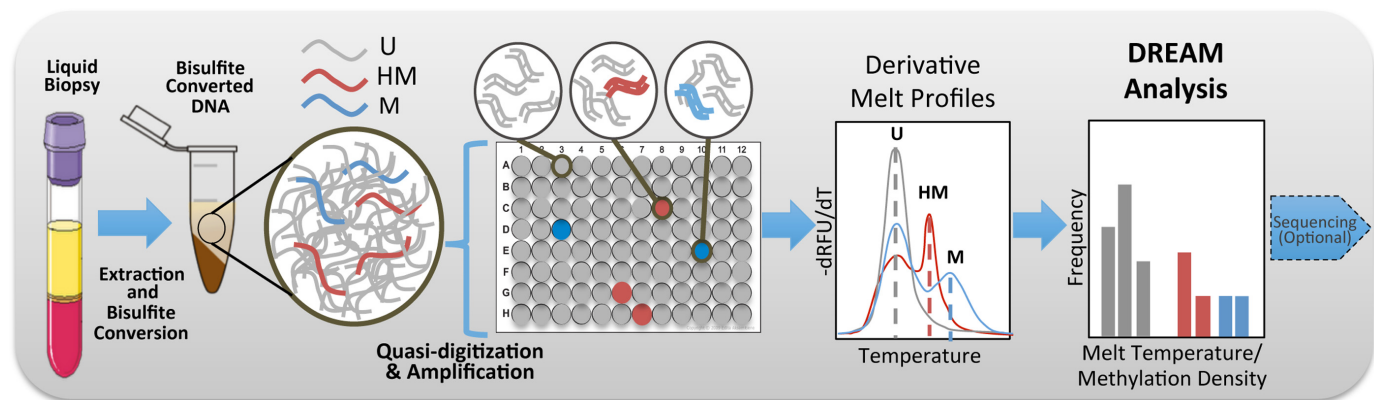


Figure 1. DREAMing: analysis of epigenetic heterogeneity at single-copy sensitivity and single-CpG-site resolution. DNA is extracted from a liquid biopsy and undergoes bisulfite treatment (BST). The sample containing a heterogeneous population of BST epialleles is then diluted such that no more than two epiallelic species are present per reaction volume (quasi-digitized). Following PCR amplification and melt curve generation, wells containing a fully methylated (M; blue DNA/curve) or heterogeneously-methylated (HM; red DNA/curve) epiallele can be easily distinguished from those that are unmethylated (U; gray DNA/curves) by the presence of a secondary melt peak, whose melt temperature (dashed lines) allows discrimination and enumeration of epiallelic variants based on methylation density. Observed melt temperatures can then be plotted as a ‘DREAM analysis’ histogram to easily visualize epigenetic/epiallelic heterogeneity.

Among the critical amendments to previously-described MS-HRM design criteria were the selection of primers/loci that result in the generation of single peak/melt temperature (T_m) negative derivative melt profiles and large differences in melt temperature between unmethylated versus methylated epialleles. Primer pairs were initially evaluated by mixing BST fully-methylated epialleles with negative control BST unmethylated genomic DNA (NC-BSTDNA). Following amplification and melt curve generation, two distinct negative derivative melt peaks are clearly visible, as shown in Figure 2. The first melt peak is derived from the melting of amplified unmethylated (U) epialleles, while the second peak corresponds to the melt profile from the fully methylated (M) epialleles. By carefully adjusting annealing temperatures and magnesium concentrations the effect of the bias against unmethylated sequences could be modulated to achieve the desired level of sensitivity (see Supplementary Figure S1A,B). After fine-tuning this approach to the detection of ultra-rare species, we could readily detect (methylated):(total epialleles) epiallelic fractions for both loci below 0.01% (Supplementary Figure S1C), demonstrating over an order of magnitude greater sensitivity than previously reported (36). The sequences and characteristics of the finally-selected primer pairs and loci are provided in Supplementary Table S1.

We also found that the overall epiallelic fraction of a bulk sample could be readily estimated using the ratio between the heights of the ‘U’ and ‘M’ peaks, as shown in Supplementary Figure S2A,B. By performing a serial dilution with epiallelic fractions ranging from 100% down to 0.01% for the $p14^{ARF}$ and *BRCA1* loci, respective characteristic equations could be defined to approximate the epiallelic fraction based solely on the negative derivative melt curve (Supplementary Figure S2C). We confirmed the validity of this approach by comparing epiallelic fraction estimates of the *BRCA1* locus in three liquid-biopsy-derived mononuclear cell samples from MDS/MPN patients based on the peak-height method against estimation based upon qMSP (Sup-

plementary Figure S2D). The results of these comparisons demonstrated that the peak-height method provides estimates within a factor of two of those obtained by qMSP. However, unlike qMSP-based estimates, the peak-height method is capable of assessing the epiallelic fraction of partially methylated epialleles and does not require comparison to a reference gene such as beta-actin.

Enumeration of individual epiallelic variants at single-copy sensitivity

We next sought to assess the ability of DREAMing to enumerate individual epiallelic variant copies accurately. We first determined the baseline melt characteristics of the $p14^{ARF}$ locus for unmethylated epialleles (Avg. T_m : $78.6 \pm 0.13^\circ\text{C}$) and fully-methylated epialleles (Avg. T_m : $83.8 \pm 0.10^\circ\text{C}$). A mock $p14^{ARF}$ DREAMing assay was then performed using standard samples containing synthetic oligonucleotides corresponding to a BST fully-methylated locus diluted into a large excess of NC-BSTDNA. Figure 3A shows the results of such an assay where a manufacturer-estimated 8 copies of synthetic DNA equivalent to a BST fully-methylated epiallele were diluted into ~ 5000 genomic equivalents of NC-BSTDNA and distributed among 87 microtiter wells. When the resulting melt curves were plotted, 78 wells exhibited a single primary ‘U’ T_m peak centered at 78.6°C (indicating the exclusive presence of NC-BSTDNA), while 8 wells produced a characteristic secondary ‘M’ melt peak at an average of 83.8°C , each indicating the presence of a single copy of a fully-methylated epiallele (one well exhibiting an atypical melt profile was excluded). Following analysis, the observed melt temperatures for each of the wells were plotted into a DREAM analysis histogram as shown in the inset of Figure 3A. Here, melt temperatures derived from wells that contain only unmethylated genomic DNA are shown in gray and those that exhibit an M peak are colored blue.

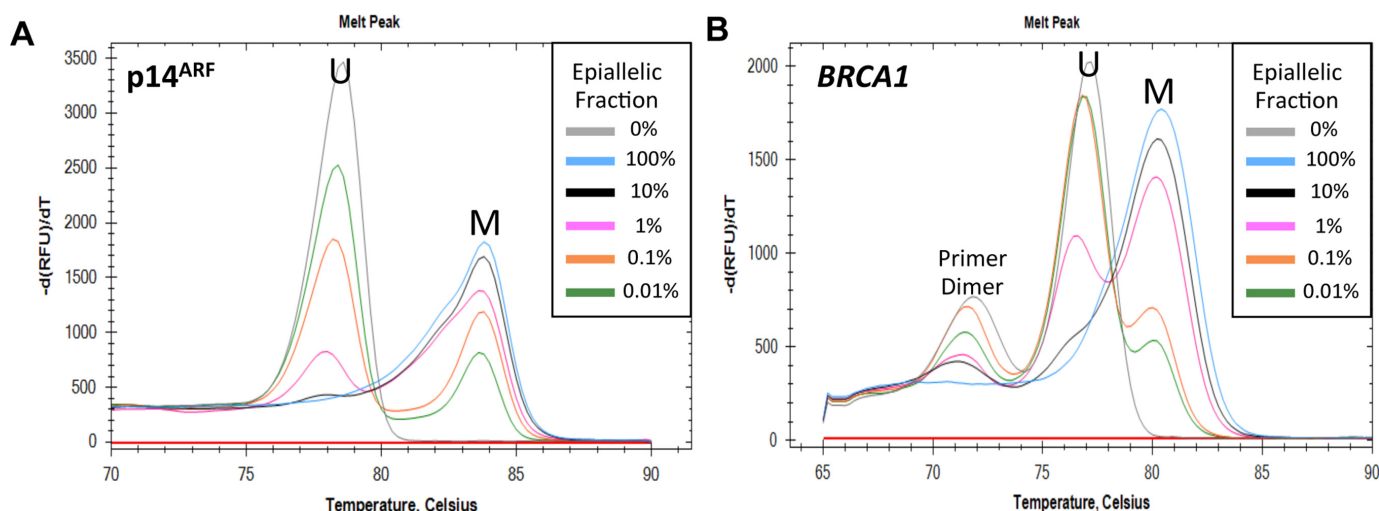


Figure 2. DREAMing primers optimized for high sensitivity. (A) Bulk (-dRFU/dT) melt curves for the p14^{ARF} locus at various genomic DNA methylated:total epiallelic fractions. (B) Bulk (-dRFU/dT) melt curves for the *BRCA1* locus at various genomic DNA methylated:total epiallelic fractions. Both assays exhibit sensitivities that provide detection of epiallelic fractions of 0.01% or lower.

Assessment of p14^{ARF} epigenetic heterogeneity in liquid biopsies from CT-scan-positive NSCLC patients

We next used DREAMing to assess epigenetic heterogeneity of the p14^{ARF} locus in 26 liquid-biopsy-derived cfDNA samples obtained from lung CT-scan-positive patients that had been previously diagnosed for NSCLC according to histopathology. cfDNA from each patient's liquid-biopsy-derived plasma was processed according to the MOB technique and preliminarily assessed for methylation in the *CDKN2A* locus within the p14^{ARF} and p16 promoters by qMSP (23) and MethyLight (24), respectively. In the majority of these samples, methylation levels/copy numbers could not be determined with high confidence by these real-time-PCR-based approaches due to weak or inconsistent amplification (ostensibly due to a combination of intrinsically low ctDNA quantity and the inability of MSP primers to amplify heterogeneously-methylated epialleles efficiently). To perform DREAMing, the BST cfDNA was diluted into microtiter plates such that, based on preliminary qMSP/MethyLight estimates of each sample, anywhere from 5–100 genomic equivalents and no more than one heterogeneously-methylated p14^{ARF} epiallelic variant were expected to be present per reaction volume. Following amplification and melt profile analysis, heterogeneously-methylated epiallelic variants were detected by identifying wells exhibiting secondary melt peaks with a $T_m \geq 79^\circ\text{C}$ (above the NC-BSTDNA T_m 96% confidence interval), thereby indicating the presence of a single heterogeneously-methylated or fully-methylated epiallelic variant. Figure 3B,C shows the results of DREAMing for two NSCLC samples with each HRM profile and DREAM analysis color-coded according to T_m /methylation density.

In order to confirm the presence and methylation density of the epiallelic variants that were detected by DREAMing, we took 137 of the positive wells from the sample analyses and had them verified by pyrosequencing. Based on this analysis, epiallelic variants were positively confirmed in 19 of 20 (95%) of the DREAMing-positive samples and in 133

of 137 (97.1%) positive wells (confirmed patterns are shown in Supplementary Table S2A, though not all wells could be reliably sequenced due to primer dimers and/or multiple secondary peaks). As shown in Figure 3D, epiallelic variants exhibited the predicted linear correlation ($R^2 = 0.97$) between variant T_m 's and methylation density, further validating the use of T_m as a means of easily discriminating between different epiallelic variants to assess epigenetic heterogeneity.

The results of p14^{ARF} DREAM analyses were also compared with *CDKN2A* (p14^{ARF} and p16) qMSP/MethyLight estimates for each of the NSCLC samples. Figure 4 shows a bar graph of the total number of variants detected by DREAMing as compared to qMSP/MethyLight in each sample, grouped according to clinical diagnosis. Variants detected by DREAMing are additionally color-coded according to T_m /methylation density. Overall, DREAMing detected p14^{ARF} methylation in over three times as many samples compared to p14^{ARF} qMSP (22 versus 7 samples, respectively) and up to 28.4-fold more variant copies in samples that were also positive by qMSP. The bars within the graph display the color-coded 'DREAM analyses' that provide a facile means of visualizing and comparing epigenetic heterogeneity between the samples.

Interestingly, previous MSP-based studies showed that while p16 and p14^{ARF} share a common promoter, their methylation status appears to be independent (43). Consistent with these reports, only 7 of the 15 p16-MethyLight-positive samples also exhibited p14^{ARF} methylation by qMSP. However, DREAMing was able to confirm that all 15 of the p16-MethyLight-positive samples also exhibited detectable p14^{ARF} methylation, intimating an association between methylation of the two loci. The disparity between these results is ostensibly due to the higher sensitivity of DREAMing and its ability to detect methylation densities as low as a single CpG site within the entire p14^{ARF} locus.

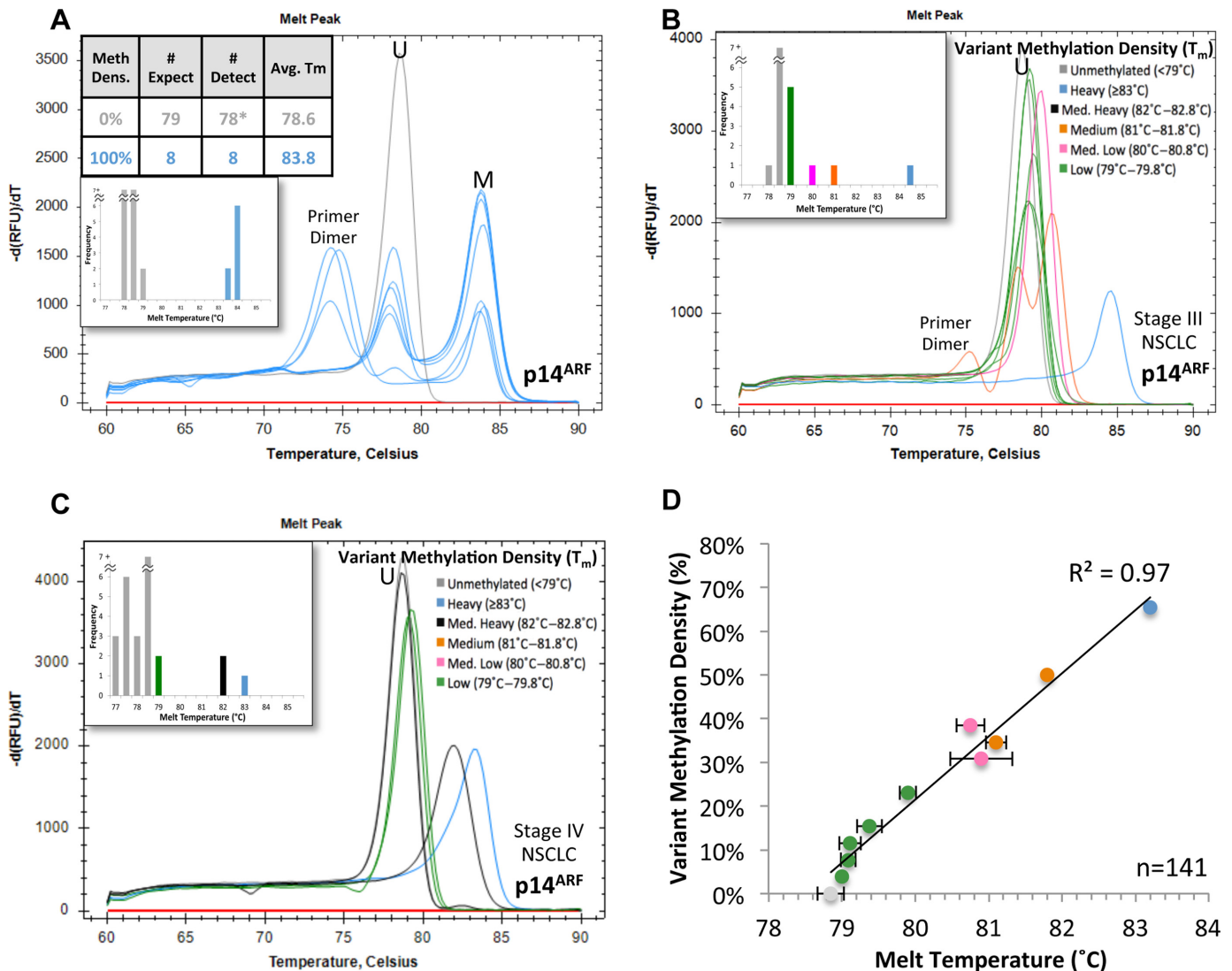


Figure 3. Analysis of p14^{ARF} epigenetic heterogeneity in NSCLC. (A) DREAM analysis showing detection of 8 of 8 copies of synthetic BST fully-methylated p14^{ARF} locus in the presence of ~5000 BST unmethylated genomic equivalents. *One aberrant profile was excluded. (B–C) DREAMing results of p14^{ARF} epiallelic heterogeneity in NSCLC samples. The DREAM analysis of each sample is displayed in the upper left insets, along with the relative T_m/methylation density frequency. (D) Scatterplot color-coded according to methylation density/melt temperature showing correlation between observed of CT-scan positive clinical NSCLC sample p14^{ARF} DREAMing melt temperatures and respective methylation densities as determined by pyrosequencing. Here, CpG sites with an indeterminate methylation status (see Supplementary Section 3.1) were considered half methylated.

Statistical framework for the differentiation of epiallelic variants at single-CpG-site resolution

We next sought to develop a statistical framework for using DREAMing to distinguish epiallelic variants at single-CpG-site resolution based solely on differences in T_m. We utilized the DREAMing assay for the *BRCA1* locus that was designed around a reduced number of CpG dinucleotides to limit variance in T_m due to methylation pattern permutations at each methylation density. We then performed HRM using synthetic oligonucleotide standards equivalent to BST *BRCA1* epialleles with methylation densities of 0% (unmethylated), 20%, 40%, 60%, 80% and 100% (fully methylated) and observed the resulting melt profiles and respective T_m's (Figure 5A). As shown in Figure 5B, linear regression analysis demonstrated a high correlation

between experimental T_m's and *in silico* estimates, thereby allowing interpolation of the melt temperatures of all possible 2ⁿ epiallelic permutations, where n is the number of CpG sites (here, n = 5) located between the primers within the locus of interest. By accounting for variances due to experimental deviation from the regression curve and differences in T_m due to permutations at each methylation density, as well as measurement imprecision, we were able to calculate the probability of a given variant methylation density as a function of observed T_m, as shown in Figure 5C (see methods for more details). Based on these probabilities, non-overlapping confidence intervals were established for distinguishing epiallelic variants at single-CpG-site resolution based solely on T_m, as shown in Figure 5D. An example of the process of incorporating raw *BRCA1* melt data

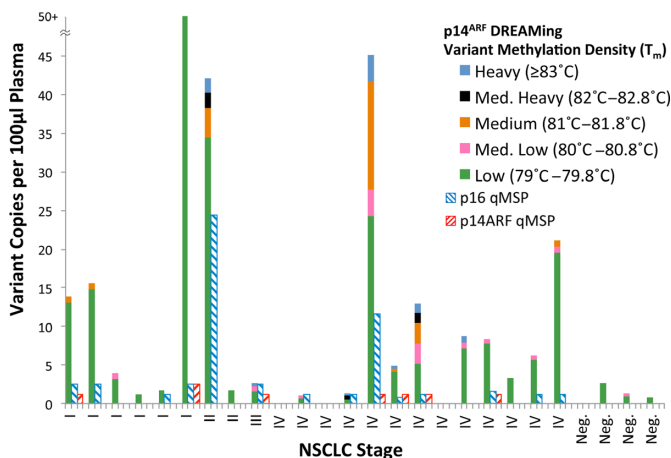


Figure 4. p14^{ARF} methylation in NSCLC via DREAMing and qMSP. Plot comparing the number of detected copies of methylated p14^{ARF} epialleles via DREAMing versus p16 qMSP (estimated) and p14^{ARF} qMSP (estimated) per 100 μ l plasma from 26 CT-scan-positive patients, plotted according to histopathological diagnosis. p14^{ARF} DREAMing detected heterogeneously-methylated p14^{ARF} epialleles in 3 times as many samples as compared to p14^{ARF} qMSP. All samples positive by p16 qMSP were also positive by p14^{ARF} DREAMing, while less than half (47%) were positive by p14^{ARF} qMSP.

into this framework is provided in Supplementary Section 3.2.

We next performed a mock assay to demonstrate the ability of DREAMing to identify and enumerate individual *BRCA1* epiallelic variants at single-CpG-site resolution. Here, a manufacturer-estimated 4.2 copies each of synthetic control DNA equivalent to BST *BRCA1* epiallelic variants at single-CpG-site (20%) and 100% methylation densities were mixed with \sim 5000 genomic equivalents of NC-BSTDNA and pipetted into 84 wells. Following amplification, three wells exhibited secondary melt peaks at 77.8–78.2°C, indicating methylation of a single CpG site (20% density), and five wells showed secondary melt peaks at 80.4–80.6°C, indicating 100% methylation density, as shown in Figure 6A. These results are well anticipated by the Poissonian probability ($\lambda = 4.2$). The right inset of Figure 6A shows the overall DREAM analysis of the mock experiment, color-coded according to the methylation density of the epiallelic variants predicted by the T_m confidence intervals.

In order to assess the fraction of false-positive epiallelic variants due to incomplete bisulfite conversion, we also performed *BRCA1* DREAMing on Epitect NC-BSTDNA and compared it with the synthetic equivalent of fully-converted unmethylated DNA. DREAMing detected an overall 0.005% heterogeneously-methylated epiallelic fraction in the NC-BSTDNA, as compared to 0.000% for the synthetic control DNA. This difference reflects the notable ability of DREAMing to easily assess the false-positive rate stemming from inherent inefficiencies in bisulfite conversion (53,54).

Assessment of *BRCA1* epigenetic heterogeneity in MDS/MPN patient liquid biopsies

Lastly, we sought to demonstrate the ability of DREAMing to directly assess the methylation density/clonal heterogeneity of the *BRCA1* locus in liquid-biopsy-derived mononuclear cells from 10 patients diagnosed with MDS/MPN. These samples were initially evaluated for *BRCA1* methylation by both qMSP as well as the bulk peak-height ratio method described above to inform the level of dilution appropriate for the DREAMing assay. Figure 6B–D shows the results of DREAM analyses of *BRCA1* heterogeneity in three of the ten clinical samples. Through simple observation of the resulting secondary-peak T_m 's, the presence of heterogeneously-methylated epiallelic variants with 20%, 60%, 80% and 100% methylation densities could be readily discriminated based on T_m confidence intervals. Secondary peaks with a difficult to discern T_m were evaluated by employing a Gaussian mixture model to remove interference from the large unmethylated peak (see Supplementary Figure S3). By counting the total number of wells that exhibited secondary heterogeneous-methylation peaks, direct enumeration of the epiallelic variants could be easily obtained, as shown in Table 1. In addition to providing assessment of epiallelic subclonal heterogeneity, DREAMing also detected up to 30-fold more epiallelic variants than qMSP per sample. Subsequent pyrosequencing analysis confirmed the predicted epiallelic variant methylation densities at single-CpG-site resolution in 33 of 34 wells chosen for sequencing analysis (see Supplementary Table S2B for sequencing results). The upper right insets in Figure 6B–D show the methylation patterns that were confirmed by pyrosequencing.

DISCUSSION

Intratumor heterogeneity has emerged as a critical component in the understanding, assessment and treatment of cancer (55). At present, the exact mechanism that accounts for this heterogeneity remains an issue of hot debate (56). Overall, two paradigms have dominated this discussion: the cancer stem cell (CSC) hypothesis (57,58) and the clonal evolution model (59). Both of these can be used to account for the commonly-observed phenomenon that only a small subset of cancer cells can transfer disease in immunocompromised mice (60,61). Yet while the two competing theories differ on how this small subset of cells arises during carcinogenesis, both paradigms posit that it is epigenetic differences that predominantly determine tumorigenic potential between cancer cells due the implausibility that only the rare tumorigenic cells contain genotypes that allow high rates of proliferation. Likewise, epigenetic heterogeneity arising from stochastically disordered methylation has been shown to directly correspond to adverse clinical outcome (14,15).

Ultra-high-depth genome-wide analyses of cancer cell populations as well as single-cell studies will be instrumental in shedding light on carcinogenesis and elucidating the epigenetic differences and mechanisms that determine cancer cell tumorigenicity and drug resistance. While such approaches have clear utility in solving these fundamental questions and for the discovery of epigenetic biomarkers, they demand considerable time, effort and resources that

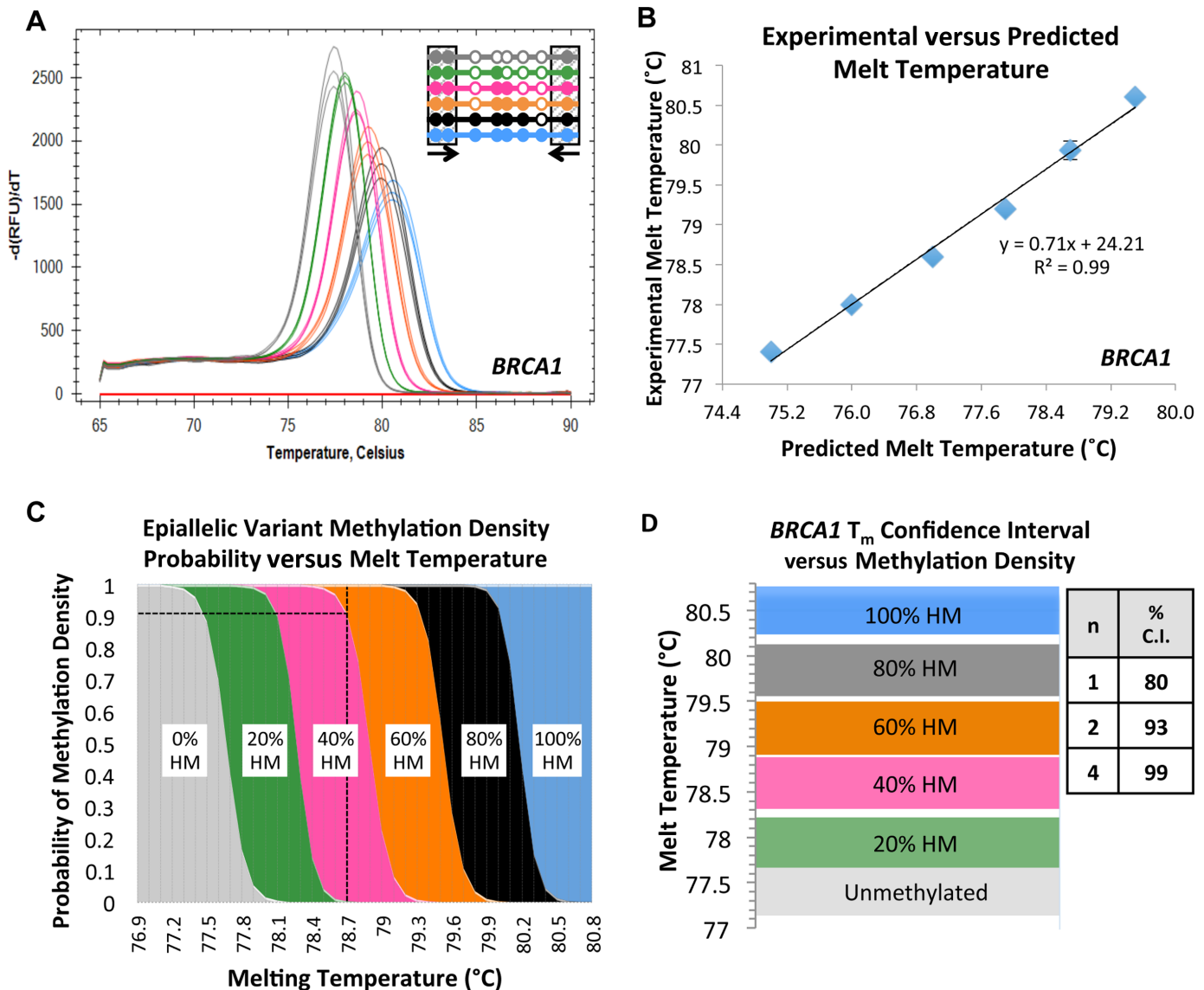


Figure 5. Statistical basis for DREAMing at single-CpG-site resolution. (A) Experimental melt profiles of synthetic DNA equivalent to the BST *BRCA1* locus at various methylation densities. Right inset indicates template methylation patterns; solid circles and empty circles indicate methylated and unmethylated CpG sites, respectively. (B) Linear regression between experimental melt temperatures and those predicted by uMELT using the 'Blake & Delcourt' thermodynamic library. (C) Probability of true *BRCA1* epiallelic variant methylation density as a function of observed melt temperature. For example, a T_m of 78.7°C (dashed line) has a ~91% probability of being derived from a variant with a 40% density and ~9% from a variant with a 60% density. (D) Confidence intervals for the melt temperature mean of 1, 2 and 4 wells grouped by methylation density.

are currently impractical for many applications in basic science and the clinic. There thus remains a critical need for methods with the sensitivity and resolution required for reliably detecting the presence and heterogeneity of epigenetic biomarkers that are sufficiently cost-effective to allow for routine use. This need is greatly exacerbated in studies involving the use of samples such as liquid biopsies that contain exceedingly low copy numbers of tumor-specific DNA, making sequencing approaches particularly cost prohibitive.

In the present work, we sought to describe a simple and inexpensive method, termed DREAMing, which allows evaluation of epigenetic heterogeneity at extremely high sensitivity and resolution. While previous reports have

shown that higher methylation densities result in higher T_m 's due to the greater GC content of methylated sequences following bisulfite treatment (28), these bulk assay techniques are limited due to issues arising from the generation of heteroduplex amplicons whose melt curve contributions obfuscate quantitative analysis (28). In contrast, DREAMing incorporates a semi-limiting (quasi-digital) dilution approach that provides that no more than two epiallelic species (typically one heterogeneously-methylated epiallelic variant amongst 5–500 unmethylated epialleles/genomic copies) be present per reaction volume. In this case, any heteroduplexes that form during amplification will melt at a lower temperature than any epiallelic homoduplexes within the same reaction volume, and can be readily excluded from

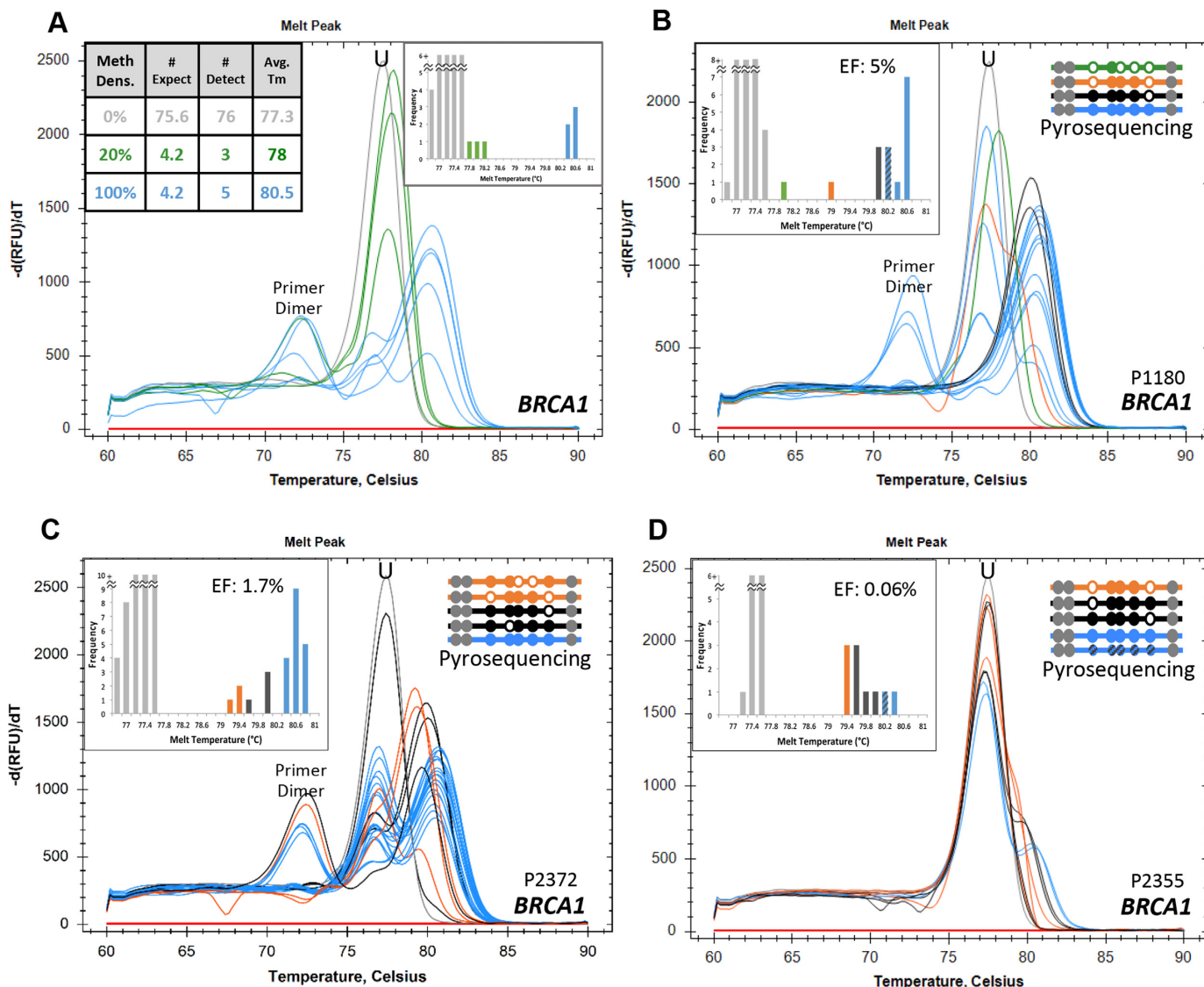


Figure 6. DREAM analysis of *BRCA1* epigenetic heterogeneity in MDS/MPN at single-copy sensitivity and single-CpG-site sensitivity. (A) DREAM analysis showing detection and methylation density assessment of synthetic epiallelic variants equivalent to a BST 20% (green) and fully-methylated (blue) *BRCA1* locus. (B–D) DREAM analysis of *BRCA1* epigenetic heterogeneity in liquid-biopsy-derived mononuclear cells from MDS/MPN patients. Raw melt profiles color-coded according to methylation density as predicted by the confidence intervals shown in Figure 5D. DREAM analysis, along with the calculated overall epiallelic fraction of each sample, are displayed in the upper left insets, while the methylation patterns, as determined by pyrosequencing, are shown in the upper right insets. Hatched coloring indicates an indeterminate melt temperature (in between the previously-determined confidence intervals).

melt curve analysis. The resulting derivative ($-dRFU/dT$) melt curve thus consists of either one or two primary T_m peaks (excluding primer dimers): one from amplified BST unmethylated (U) epialleles and a secondary peak from each well containing an amplified single copy of a BST heterogeneously-methylated or fully-methylated (M) epiallelic variant. Based on this paradigm, here we have demonstrated that DREAMing can be used to enumerate and discriminate individual epiallelic variants in fractions as low as 0.005% at single-CpG-site resolution using only a single 96-well microtiter plate. This makes it far more efficient than fully digital approaches (62), which would nominally require HRM analysis of $\sim 50\,000$ reaction-wells to achieve a similar sensitivity.

There are a number of issues that are important to consider for the successful design and implementation of DREAMing assays. First, loci should be specifically selected based upon the desired application. For example, the number of CpG sites located between the primers can greatly influence the number of copies of heterogeneously-methylated DNA that will be detected in a given assay. Choosing to include more CpGs will ostensibly increase the sensitivity of the assay to lower methylation densities and, potentially, to disease or phenotype-specific methylation. On the other hand, the inclusion of large numbers of CpG's (e.g. >10) typically reduces assay resolution, thereby limiting the ability to differentiate between epialleles at similar methylation densities, potentially compromising over-

Table 1. Results of *BRCA1* DREAM analysis of MDS/MPN mononuclear cells and pyrosequencing confirmation

Sample	EF qMSP	20%HM EF (Conf.)	40%HM EF (Conf.)	60%HM EF (Conf.)	80%HM EF (Conf.)	100%HM EF (Conf.)	BorderlineTm EF (Conf.)	Total EF DREAMing (Conf.)	Variants Detected versus qMSP
P1143	6%	0.3% (1/1)	-	-	2.4% (5/6)*	7.4% (2/2)	0.5% (2/2)	11% (10/11)	183%
P1180	5%	0.3% (1/1)	-	0.3% (1/1)	0.9% (1/1)	2.5% (3/3)	0.9%*	5% (6/6)	100%
P2338	0.02%	-	-	0.01%*	0.01%*	0.01%*	-	0.03%	150%
P2355	Below LOD (Est. 0.002%)	-	-	0.02% (1/1)	0.03% (2/2)	0.01% (2/2)	-	0.06% (5/5)	~3000%
P2361	0.00%	-	-	-	-	-	-	0.00%	-
P2366	0.02%	-	-	-	0.005%*	0.02% (3/3)	0.005%*	0.03% (3/3)	150%
P2368	0.00%	-	-	-	-	-	-	0.00%	-
P2372	0.8%	-	-	0.2% (2/2)	0.3% (2/2)	1.2% (5/5)	-	1.7% (9/9)	213%
P2434	0.01%	-	-	-	0.01%*	-	-	0.01%	100%
P2437	0.02%	-	-	0.03%*	0.03%*	-	0.009%*	0.06%	300%
NC-	0.00%	0.005%	-	-	-	-	-	0.005%	∞
BSTDNA									
NC-	0.00%	-	-	-	-	-	-	0.000%	-
Synthetic									

Ten MDS/MPN samples were analyzed and methylated:total epiallelic fractions (EF) compared with independent qMSP analysis. DREAMing methylation density assessments were confirmed by pyrosequencing in select positive wells, the results of which are shown in parentheses. Of the 34 wells that were sequenced, 33 confirmed the predicted level of methylation. Key: *Some sequencing not possible due to presence of primer dimers.

all analyses of epigenetic heterogeneity. While not comprehensively evaluated here, longer loci may also be subject to higher backgrounds and/or false positives due to the inclusion of an increased number of [potentially unconverted] cytosines within the template strands.

Another issue that deserves particular clarification is the generation and effect of nonspecific hybridizations (e.g. primer dimers and heteroduplexes) of amplified products on the DREAMing assay. This issue is particularly notable in that the primer design and assay optimization strategy results in a heavy bias against unmethylated sequences leading to relatively less competition for hybridization and an increase in potential amplifications due to nonspecific interactions. Additionally, the semi-limiting dilution technique reduces the number of target strands further exacerbating these phenomena and making nonspecific hybridizations not uncommon in DREAMing assays. Evidence of these nonspecific hybridizations is readily observable in the resulting melt profiles as distinct [lower melt temperature] peaks or reductions in the T_m of the 'U' peak. However, unlike real time assays, such as qMSP, these nonspecific interactions very rarely affect the DREAMing assay itself due to fact that they [all but always] result in melt temperature peaks below unmethylated controls and thus do not interfere with the evaluation of peaks or the quantification of amplified epialleles.

A third issue that arises in the implementation of DREAMing for clinical samples is the determination of the appropriate dilution to achieve quasi-digitization of the sample epialleles. This is primarily an issue stemming from the limited number of reaction volumes available in a microtiter plate. Given that the number of wells containing a methylated epiallele is expected to follow a normal Poissonian distribution, samples should be diluted so as to yield a quasi-digital state (0 or 1 methylated epiallele) for the vast majority of wells. It is thus suggested that samples be diluted such that a methylated epiallele is expected to be present in no more than 40% of the total assay wells ($\lambda = 0.4$ per plate) in order to maintain a quasi-digital state for >95% of the wells. Due to this constraint, it is generally advisable that all samples be preliminarily assessed for overall DNA copy numbers by performing methylation-independent qMSP on

a housekeeping gene such as beta-actin. For samples whose epiallelic fraction cannot be *a priori* estimated within a factor of ~40, such as the MDS/MPN mononuclear cell isolates tested here, a preliminary bulk technique such as qMSP or the previously described peak height ratio method can also be used to approximate the epiallelic fraction and determine the minimum level of dilution required to produce a quasi-digital distribution ($\lambda < 0.4$ per plate). As long as this condition is met, the sensitivity of the assay is then determined by the bulk sensitivity of the optimized assay divided by the number of wells or by the bisulfite conversion efficiency of the target locus. In the case of cfDNA the situation is typically simpler as most clinical samples can be easily evaluated by distributing the sample into the reaction wells such that there are no more than 100 genomic copies per well. This level of dilution (100 copies per well) in a 96-well plate provides a sensitivity of approximately 0.01% and maintains reasonable quasi-digitization for samples with epiallelic fractions as high as ~0.4%, which, in our experience, is suitable for the vast majority of plasma samples.

The chosen level of dilution also has a direct relationship to the statistical significance of a given DREAMing assay. Statistically speaking, each DREAMing assay can be considered as a binomial observation of the fraction of wells (or number of genomic copies) containing heterogeneously-methylated epialleles that meet a given set of criteria. For example, in a certain assay a researcher may be interested in the overall epiallelic fraction, the fraction at a given methylation density or, if supplemented by sequencing, the fraction of unique methylation patterns. The statistical significance of the population meeting the chosen criteria would then be calculated using standard binomial statistical tests (63) that can be simplified by assuming a normal distribution, and where the number of events, n , is equal to the number of assay wells and the proportion of successes, \hat{p} would be the number of wells meeting the desired criteria divided by n .

As alluded to previously, the high sensitivity and resolution of DREAMing also make it particularly dependent upon the bisulfite conversion efficiency of the target locus. For this reason, it is recommended that proper negative con-

Table 2. Comparison of DREAMing against commonly employed methylation analysis platforms

Technology	Average Prep Time	Average Assay Time	Cost per Sample	Native Sensitivity	Heterogeneous Methylation Detection	Multi-Epiallelic Discrimination	Directly Quantitative	Main application
MSP/MethyLight	<30 min	2–3 h	\$10–\$100	0.01–0.1%	No	No	No	Fully methylated detection
MS-HRM (biased primers)	<30 min	2–3 h	\$10–\$100	0.1%	Yes	No	No	Rough HM Epiallele detection
Pyrosequencing	~1 h	~6 h	\$50–\$100	5%	Yes	No	No	Abundant Methylation Detection
Bisulfite sequencing	days	0.5–9 days	\$1K–\$10K	0.1%–1%	Yes	Yes	No	Methylome Analysis
DREAMing	<30 min	2–3 h	\$50–\$100	0.005%	Yes	Yes	Yes	Epiallelic Heterogeneity analysis

HM: heterogeneously-methylated.

trols, such as healthy human DNA and Epitect unmethylated control DNA, be incorporated and rigorously evaluated prior to evaluating any samples of interest in order to determine the rate of false positives due to incomplete bisulfite conversion. For all research studies, healthy or normal controls should also be extensively evaluated in order to distinguish between phenotypic or disease-specific methylation versus low-level methylation that is within the normal range of control populations.

Lastly, a number of fundamental limitations may render DREAMing unsuitable for some applications. For example, due to the demand for quasi-digitization, the sensitivity of the assay is ultimately determined by the dominant [non-zero] epiallelic methylation density within the sample as well as the number of reaction volumes tested. In the case where one species is particularly over represented and the assay is performed in a 96-well microtiter plate, epiallelic species present at an order of magnitude or lower than the dominant fraction will be not be accurately represented or may be missed altogether. While this limitation can, in principle, be overcome through the use of higher density arrays, doing so may undermine the simplicity and cost-effectiveness of the assay.

Another fundamental limitation of DREAMing is the relatively low throughput of the assay. This all but prevents its use for biomarker discovery, as screening multiple loci would be costly, time-consuming and necessitate large sample volumes. Likewise, its low throughput makes DREAMing far more suitable for individual biomarkers than biomarker panels. These problems could be somewhat alleviated by limiting the number of wells per target to allow for more targets to be run in the same plate, but would concomitantly reduce the sensitivity and statistical significance for each target-assay.

A third limitation is that DREAMing does not directly provide sequence information. This limitation ultimately prevents full characterization of epigenetic heterogeneity within a given sample. Nonetheless, DREAMing can be easily augmented with sequencing, as demonstrated here, through incorporation of modified primers. It should be noted, however, that while [supplemental] sequencing can be used to identify epiallelic clones with different patterns of methylation amongst the detected epiallelic variants of a given methylation density, it is currently unclear whether there is significant biological consequence to the methylation of *particular* CpG sites. There is, however, strong evidence that locus methylation density has a direct effect on

epigenetic silencing and the generation of distinct epigenetic phenotypes (21,64–65).

Overall, DREAMing represents a simple and practical means of assessing epigenetic heterogeneity in difficult samples such as liquid biopsies by providing direct quantification and discrimination of epiallelic variants at single-CpG-site resolution. The technique offers a number of advantages over other methods (see Table 2) and can be performed at minimal time and cost in a microtiter plate using a standard qPCR machine capable of HRM analysis. DREAMing is suitable for ultra-sensitive assessment of epigenetic heterogeneity in any specimen of interest, particularly those containing low abundance epialleles, and has potential utility in the evaluation of DNA methylation dynamics in cell populations (17), prenatal diagnostics (66), as well as early cancer diagnostic, companion diagnostic and predictive applications (9).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

The authors would also like to thank Michael A. McDewitt for providing the MDS/MPN blood samples, as well as Stephanie Fraley and Alejandro Stark for their excellent advice in HRM and bisulfite conversion, respectively.

FUNDING

National Institutes of Health [R01CA155305, R21CA186809 and U54CA151838]. Funding for open access charge: National Institutes of Health. Pornpat Athamanolap was supported by the Royal Thai Government fellowship. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest statement. None declared.

REFERENCES

- Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A. and Kinzler, K.W. (2013) Cancer Genome Landscapes. *Science*, **339**, 1546–1558.
- Hanahan, D. and Weinberg, R.A. Hallmarks of Cancer: The Next Generation. *Cell*, **144**, 646–674.
- Herman, J.G. and Baylin, S.B. (2003) Gene Silencing in Cancer in Association with Promoter Hypermethylation. *N. Engl. J. Med.*, **349**, 2042–2054.

4. Nephew, K.P. and Huang, T.H.-M. (2003) Epigenetic gene silencing in cancer initiation and progression. *Cancer Letters*, **190**, 125–133.
5. Heyn, H. and Esteller, M. (2012) DNA methylation profiling in the clinic: applications and challenges. *Nat. Rev. Genet.*, **13**, 679–692.
6. Landan, G., Cohen, N.M., Mukamel, Z., Bar, A., Molchadsky, A., Brosh, R., Horn-Saban, S., Zalcenstein, D.A., Goldfinger, N., Zundelovich, A. *et al.* (2012) Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. *Nat. Genet.*, **44**, 1207–1214.
7. Feinberg, A.P., Ohlsson, R. and Henikoff, S. (2006) The epigenetic progenitor origin of human cancer. *Nat. Rev. Genet.*, **7**, 21–33.
8. Easwaran, H., Tsai, H.-C. and Baylin, S.B. Cancer Epigenetics: Tumor Heterogeneity, Plasticity of Stem-like States, and Drug Resistance. *Mol. Cell*, **54**, 716–727.
9. Brown, R., Curry, E., Magnani, L., Wilhelm-Benartzi, C.S. and Borley, J. (2014) Poised epigenetic states and acquired drug resistance in cancer. *Nat. Rev. Cancer*, **14**, 747–753.
10. Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A. *et al.* (2010) A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. *Cell*, **141**, 69–80.
11. Meacham, C.E. and Morrison, S.J. (2013) Tumour heterogeneity and cancer cell plasticity. *Nature*, **501**, 328–337.
12. Bedard, P.L., Hansen, A.R., Ratain, M.J. and Siu, L.L. (2013) Tumour heterogeneity in the clinic. *Nature*, **501**, 355–364.
13. Turner, N.C. and Reis-Filho, J.S. (2012) Genetic heterogeneity and cancer drug resistance. *Lancet Oncol.*, **13**, e178–e185.
14. Landau, D.A., Clement, K., Ziller, M.J., Boyle, P., Fan, J., Gu, H., Stevenson, K., Sougnez, C., Wang, L., Li, S. *et al.* Locally Disordered Methylation Forms the Basis of Intratumor Methylation Variation in Chronic Lymphocytic Leukemia. *Cancer Cell*, **26**, 813–825.
15. Oakes, C.C., Claus, R., Gu, L., Assenov, Y., Hüllelin, J., Zucknick, M., Bieg, M., Brocks, D., Bogatyrova, O., Schmidt, C.R. *et al.* (2014) Evolution of DNA Methylation Is Linked to Genetic Aberrations in Chronic Lymphocytic Leukemia. *Cancer Discov.*, **4**, 348–361.
16. Schwarzenbach, H., Hoon, D.S.B. and Pantel, K. (2011) Cell-free nucleic acids as biomarkers in cancer patients. *Nat. Rev. Cancer*, **11**, 426–437.
17. Shipony, Z., Mukamel, Z., Cohen, N.M., Landan, G., Chomsky, E., Zelig, S.R., Fried, Y.C., Ainbinder, E., Friedman, N. and Tanay, A. (2014) Dynamic and static maintenance of epigenetic memory in pluripotent and somatic cells. *Nature*, **513**, 115–119.
18. Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E.S. and Getz, G. (2013) Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotech.*, **31**, 213–219.
19. Diaz, L.A. and Bardelli, A. (2014) Liquid Biopsies: Genotyping Circulating Tumor DNA. *J. Clin. Oncol.*, **32**, 579–586.
20. Ilie, M., Hofman, V., Long, E., Bordone, O., Selva, E., Washetine, K., Marquette, C.H. and Hofman, P. (2014) Current challenges for detection of circulating tumor cells and cell-free circulating nucleic acids, and their characterization in non-small cell lung carcinoma patients. What is the best blood substrate for personalized medicine? *Ann. Transl. Med.*, **2**, 107.
21. Curradi, M., Izzo, A., Badaracco, G. and Landsberger, N. (2002) Molecular Mechanisms of Gene Silencing Mediated by DNA Methylation. *Mol. Cell. Biol.*, **22**, 3157–3173.
22. Herman, J.G., Graff, J.R., Myöhänen, S., Nelkin, B.D. and Baylin, S.B. (1996) Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Nat. Acad. Sci.*, **93**, 9821–9826.
23. Lo, Y.M.D., Wong, I.H.N., Zhang, J., Tein, M.S.C., Ng, M.H.L. and Hjelm, N.M. (1999) Quantitative Analysis of Aberrant p16 Methylation Using Real-Time Quantitative Methylation-specific Polymerase Chain Reaction. *Cancer Res.*, **59**, 3899–3903.
24. Eads, C.A., Danenberg, K.D., Kawakami, K., Saltz, L.B., Blake, C., Shibata, D., Danenberg, P.V. and Laird, P.W. (2000) MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res.*, **28**, e32.
25. Bibikova, M., Lin, Z., Zhou, L., Chudin, E., Garcia, E.W., Wu, B., Doucet, D., Thomas, N.J., Wang, Y., Vollmer, E. *et al.* (2006) High-throughput DNA methylation profiling using universal bead arrays. *Genome Res.*, **16**, 383–393.
26. Taylor, K.H., Kramer, R.S., Davis, J.W., Guo, J., Duff, D.J., Xu, D., Caldwell, C.W. and Shi, H. (2007) Ultradeep Bisulfite Sequencing Analysis of DNA Methylation Patterns in Multiple Gene Promoters by 454 Sequencing. *Cancer Res.*, **67**, 8511–8518.
27. Meissner, A., Gnirke, A., Bell, G.W., Ramsahoye, B., Lander, E.S. and Jaenisch, R. (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.*, **33**, 5868–5877.
28. Mikeska, T., Candiloro, I.L.M. and Dobrovic, A. (2010) The implications of heterogeneous DNA methylation for the accurate quantification of methylation. *Epigenomics*, **2**, 561–573.
29. Laird, P.W. (2010) Principles and challenges of genome-wide DNA methylation analysis. *Nat. Rev. Genet.*, **11**, 191–203.
30. Newman, A.M., Bratman, S.V., To, J., Wynne, J.F., Eclov, N.C.W., Modlin, L.A., Liu, C.L., Neal, J.W., Wakelee, H.A., Merritt, R.E. *et al.* (2014) An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.*, **20**, 552–558.
31. Varley, K.E., Mutch, D.G., Edmonston, T.B., Goodfellow, P.J. and Mitra, R.D. (2009) Intra-tumor heterogeneity of MLH1 promoter methylation revealed by deep single molecule bisulfite sequencing. *Nucleic Acids Res.*, **37**, 4603–4612.
32. Wang, Q., Jia, P., Cheng, F. and Zhao, Z. (2015) Heterogeneous DNA methylation contributes to tumorigenesis through inducing the loss of coexpression connectivity in colorectal cancer. *Genes Chromosomes Cancer*, **54**, 110–121.
33. Graff, J.R., Gabrielson, E., Fujii, H., Baylin, S.B. and Herman, J.G. (2000) Methylation patterns of the E-cadherin 5' CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression. *J. Biol. Chem.*, **275**, 2727–2732.
34. Moelans, C.B., de Groot, J.S., Pan, X., van der Wall, E. and van Diest, P.J. (2014) Clonal intratumor heterogeneity of promoter hypermethylation in breast cancer by MS-MLPA. *Mod. Pathol.*, **27**, 869–874.
35. Nabili, N.H., Deleyrolle, L.P., Darst, R.P., Riva, A., Reynolds, B.A. and Kladde, M.P. (2014) Multiplex mapping of chromatin accessibility and DNA methylation within targeted single molecules identifies epigenetic heterogeneity in neural stem cells and glioblastoma. *Genome Res.*, **24**, 329–339.
36. Wojdacz, T.K., Dobrovic, A. and Hansen, L.L. (2008) Methylation-sensitive high-resolution melting. *Nat. Protoc.*, **3**, 1903–1908.
37. Tusnady, G.E., Simon, I., Váradi, A. and Arányi, T. (2005) BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes. *Nucleic Acids Res.*, **33**, e9.
38. Dwight, Z., Palais, R. and Wittwer, C.T. (2011) uMELT: prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application. *Bioinformatics*, **27**, 1019–1020.
39. Rhee, I., Bachman, K.E., Park, B.H., Jair, K.-W., Yen, R.-W.C., Schuebel, K.E., Cui, H., Feinberg, A.P., Lengauer, C., Kinzler, K.W. *et al.* (2002) DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature*, **416**, 552–556.
40. Keeley, B., Stark, A., Pisanic Ii, T.R., Kwak, R., Zhang, Y., Wrangle, J., Baylin, S., Herman, J., Ahuja, N., Brock, M.V. *et al.* (2013) Extraction and processing of circulating DNA from large sample volumes using methylation on beads for the detection of rare epigenetic events. *Clin. Chim. Acta*, **425**, 169–175.
41. Genereux, D.P., Johnson, W.C., Burden, A.F., Stöger, R. and Laird, C.D. (2008) Errors in the bisulfite conversion of DNA: modulating inappropriate- and failed-conversion frequencies. *Nucleic Acids Res.*, **36**, e150.
42. Korbie, D.J. and Mattick, J.S. (2008) Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat. Protoc.*, **3**, 1452–1456.
43. Esteller, M., Tortola, S., Toyota, M., Capella, G., Peinado, M.A., Baylin, S.B. and Herman, J.G. (2000) Hypermethylation-associated inactivation of p14ARF Is Independent of p16INK4a Methylation and p53 Mutational Status. *Cancer Res.*, **60**, 129–133.
44. Blake, R.D. and Delcourt, S.G. (1998) Thermal stability of DNA. *Nucleic Acids Res.*, **26**, 3323–3332.
45. Bailey, V.J., Zhang, Y., Keeley, B.P., Yin, C., Pelosky, K.L., Brock, M., Baylin, S.B., Herman, J.G. and Wang, T.-H. (2010) Single-Tube Analysis of DNA Methylation with Silica Superparamagnetic Beads. *Clin. Chem.*, **56**, 1022–1025.

46. Mao, F., Leung, W.-Y. and Xin, X. (2007) Characterization of EvaGreen and the implication of its physicochemical properties for qPCR applications. *BMC Biotechnol.*, **7**, 76.
47. Stott, F.J., Bates, S., James, M.C., McConnell, B.B., Starborg, M., Brookes, S., Palmero, I., Ryan, K., Hara, E., Vousden, K.H. *et al.* (1998) The alternative product from the human CDKN2A locus, p14(ARF), participates in a regulatory feedback loop with p53 and MDM2. *EMBO J.*, **17**, 5001–5014.
48. Hsu, H.-S., Wang, Y.-C., Tseng, R.-C., Chang, J.-W., Chen, J.-T., Shih, C.-M., Chen, C.-Y. and Wang, Y.-C. (2004) CpG Island Methylation Is Responsible for p14ARF Inactivation and Inversely Correlates with p53 Overexpression in Resected Non-Small Cell Lung Cancer. *Clin. Cancer Res.*, **10**, 4734–4741.
49. Fischer, J.R., Ohnmacht, U., Rieger, N., Zemaitis, M., Stoffregen, C., Manegold, C. and Lahm, H. (2007) Prognostic significance of RASSF1A promoter methylation on survival of non-small cell lung cancer patients treated with gemcitabine. *Lung Cancer*, **56**, 115–123.
50. Saito, K., Takigawa, N., Ohtani, N., Iioka, H., Tomita, Y., Ueda, R., Fukuoka, J., Kuwahara, K., Ichihara, E., Kiura, K. *et al.* (2013) Antitumor Impact of p14ARF on Gefitinib-Resistant Non-Small Cell Lung Cancers. *Mol. Cancer Ther.*, **12**, 1616–1628.
51. Veeck, J., Roper, S., Setien, F., Gonzalez-Suarez, E., Osorio, A., Benitez, J., Herman, J.G. and Esteller, M. (2010) BRCA1 CpG Island Hypermethylation Predicts Sensitivity to Poly(Adenosine Diphosphate)-Ribose Polymerase Inhibitors. *J. Clin. Oncol.*, **28**, e563–e564.
52. Poh, W., Moliterno, A.R., Pratz, K.W., Gojo, I., McDevitt, M.A. and Herman, J.G. (2014) Epigenetic Silencing of BRCA1 Is Linked to Homologous Recombination Repair Defects and Elevated Mir-155 Expression in Myeloid Neoplasms. *Blood*, **124**, 3525–3525.
53. Holmes, E.E., Jung, M., Meller, S., Lisse, A., Sailer, V., Zech, J., Mengdehl, M., Garbe, L.-A., Uhl, B., Kristiansen, G. *et al.* (2014) Performance Evaluation of Kits for Bisulfite-Conversion of DNA from Tissues, Cell Lines, FFPE Tissues, Aspirates, Lavages, Effusions, Plasma, Serum, and Urine. *PLoS ONE*, **9**, e93933.
54. Warnecke, P.M., Stirzaker, C., Song, J., Grunau, C., Melki, J.R. and Clark, S.J. (2002) Identification and resolution of artifacts in bisulfite sequencing. *Methods*, **27**, 101–107.
55. McGranahan, N. and Swanton, C. (2015) Biological and Therapeutic Impact of Intratumor Heterogeneity in Cancer Evolution. *Cancer Cell*, **27**, 15–26.
56. Shackleton, M., Quintana, E., Fearon, E.R. and Morrison, S.J. (2009) Heterogeneity in Cancer: Cancer Stem Cells versus Clonal Evolution. *Cell*, **138**, 822–829.
57. Reya, T., Morrison, S.J., Clarke, M.F. and Weissman, I.L. (2001) Stem cells, cancer, and cancer stem cells. *Nature*, **414**, 105–111.
58. Magee, J.A., Piskounova, E. and Morrison, S.J. (2012) Cancer Stem Cells: Impact, Heterogeneity, and Uncertainty. *Cancer Cell*, **21**, 283–296.
59. Nowell, P.C. (1976) The Clonal Evolution of Tumor Cell Populations. *Science*, **194**, 23–28.
60. Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J. and Clarke, M.F. (2003) Prospective identification of tumorigenic breast cancer cells. *Proc. Nat. Acad. Sci.*, **100**, 3983–3988.
61. Ricci-Vitiani, L., Lombardi, D.G., Pilozzi, E., Biffoni, M., Todaro, M., Peschle, C. and De Maria, R. (2007) Identification and expansion of human colon-cancer-initiating cells. *Nature*, **445**, 111–115.
62. Candiloro, I., Mikeska, T., Hokland, P. and Dobrovic, A. (2008) Rapid analysis of heterogeneously methylated DNA using digital methylation-sensitive high resolution melting: application to the CDKN2B (p15) gene. *Epigenetics & Chromatin*, **1**, 7.
63. Newcombe, R.G. (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat. Med.*, **17**, 857–872.
64. Lorincz, M.C., Schübeler, D., Goeke, S.C., Walters, M., Groudine, M. and Martin, D.I.K. (2000) Dynamic Analysis of Proviral Induction and De Novo Methylation: Implications for a Histone Deacetylase-Independent, Methylation Density-Dependent Mechanism of Transcriptional Repression. *Mol. Cell. Biol.*, **20**, 842–850.
65. Cameron, E.E., Baylin, S.B. and Herman, J.G. (1999) p15INK4B CpG Island Methylation in Primary Acute Leukemia Is Heterogeneous and Suggests Density as a Critical Factor for Transcriptional Silencing. *Blood*, **94**, 2445–2451.
66. Lo, Y.M.D. and Chiu, R.W.K. (2007) Prenatal diagnosis: progress through plasma nucleic acids. *Nat. Rev. Genet.*, **8**, 71–77.