Published in final edited form as:

Nutr Rev. 2008 August ; 66(Suppl 1): S69–S72. doi:10.1111/j.1753-4887.2008.00072.x.

Influence of genetic inheritance on global epigenetic states and cancer risk prediction with DNA methylation signature: challenges in technology and data analysis

Maxwell P Lee,

Laboratory of Population Genetics, Center for Cancer Research, National Cancer Institute, Bethesda, Maryland, USA.

Barbara K Dunn

Basic Prevention Science Research Group, Division of Cancer Prevention, National Cancer Institute, Bethesda, Maryland, USA.

Human health and disease are determined by genetic background and environmental exposures. The interplay between these two in both normal development and disease are mediated by epigenetic regulation of gene expression. Epigenetic regulation causes heritable changes in gene expression that are not associated with DNA sequence changes. Instead, epigenetic changes are mediated by chemical modification of DNA such as DNA methylation or by protein modifications such as histone acetylation and methylation. Epigenetic modifications can be maintained through mitotic cell divisions. This stable transmission of the epigenetic state through mitosis provides a basis for cellular differentiation and the ultimate development of the organism. In contrast to mitosis, few examples exist of inheritance of epigenetic information across generations in model organisms. Specifically, prior to our study no investigation of the global effect of genetic inheritance on chromatin state in humans has been reported. Inheritance of epigenetic modifications is likely to inform our understanding of families that have multiple cancers that remain unexplained by mutations in known cancer-causing genes.

ALLELE-SPECIFIC CHIP-ON-CHIP ANALYSIS OF CHROMATIN MODIFICATIONS

To study the inheritance pattern of epigenetic modifications, we combined two commonly used technologies to develop a unique approach to the analysis of chromatin states in lymphoblastoid cells derived from CEPH (Centre d'Etude de Polymorphisme humain) families: SNP arrays used to trace parental origin together with chromatin immunoprecipitation, allowing “allele-specific chromatin immunoprecipitation (ChIP-on-chip)”, or “AS ChIP-on-chip”. Immunoprecipitating chromatin with antibodies targeting RNA polymerase II and five selected post-translationally modified forms of the histone H3 protein enabled us to distinguish parent- of-origin modifications associated with active

Correspondence: *MP Lee*, Laboratory of Population Genetics, Building 41, Room D702, 41 Library Drive, Bethesda, MD 20892, USA. leemax@mail.nih.gov, Phone: +301-435-8956, Fax: +301-435-8963. .

Declaration of interest. The authors have no relevant interests to declare.

versus inactive chromatin state. The novelty of our approach is seen in several components of our application of ChIP-on-chip technology. Rather than using a regular oligo- array (25–60 nucleotides per oligo, depending on the manufacturer) in the manner of the conventional ChIP-on-chip approach, which attempts to identify DNA segments that bind to immunoprecipitated protein without reference to specific allele, our application of ChIP-on-chip involves the hybridization of immunoprecipitated DNA to the Affymetrix 10K SNP chip (Affymetrix, Santa Clara, CA, USA). In contrast to the regular oligo-array, the 10K SNP chip permits detection of a specific allele (not just the gene region) targeted by the hybridizing DNA. The average heterozygosity of the SNPs represented on the 10K SNP chip is 0.3 in the Caucasian population, which is represented in the CEPH families, indicating that the tested individuals were on average informative for 30% of the assayed SNPs. This allowed us to distinguish selective hybridization of immunoprecipitated DNA to each of the two alleles at a given site for 30% of the SNPs. Since the Affymetrix 10K SNP chip was designed for the purpose of genotyping, we had to modify the protocol in order to use this chip for doing our AS ChIP-on-chip studies, as shown in Figure 1. In sum, we made two key modifications to the laboratory technology: 1) the Affymetrix 10K SNP chip was utilized, rather than the usual oligo array, in combination with ChIP-on-chip; and 2) adaptation of the 10K SNP chip to allele-specific typing, rather than the usual geno-typing. These modifications necessitated accommodating adaptations of the analytic software to carry out the analyses described below.

Ultimately, we applied this logic to analysis of variation within and between families by using multiple cell lines from the CEPH families to evaluate variation of chromatin states across pedigrees. As quality control for the allelic specificity of the ChIP assay, we used a SNP located in the differential methylation region (KvDMR) (define Kv) of the promoter of *LIT1*, a well-known imprinted gene. Our results showed that the paternal allele of *LIT1* was specifically pulled down by antibodies targeting active chromatin (Pol, RNA polymerase II; Ac, histone H3 acetylation; and K4, histone H3 lysine 4 dim- ethylation). Conversely, the maternal allele was preferentially pulled down by antibodies targeting inactive chromatin (K9, histone H3 lysine 9 trimethylation; K27di, histone H3 lysine 27 dimethylation; and K27tri, histone H3 lysine 27 trimethylation).

Having established allele specificity for our ChIP assay using *LIT1* and other imprinted genes as well as X-linked genes, which also exhibit allele-specific epigenetic modifications, we analyzed genome-wide allele- specific chromatin states by combining the ChIP-on-chip method with a SNP chip (AS ChIP-on-chip) as described. We carried out our modified procedure on 12 lym phoblastoid cell lines, six from each of two CEPH families. Each cell line was analyzed with the following six antibodies: Pol II, H3Ac, H3K4, H3K9, H3K27di, and H3K27tri. In addition, for each cell line, we performed two control experiments, one consisting of “input” (non-immunoprecipitated) DNA and the other of genomic DNA using unmodified, standard genotyping protocol. Each SNP had two measurements, one for chromatin binding from the A allele and the other from the B allele. The relative intensity was calculated as the ratio of A allele chromatin binding intensity divided by the total (A+B) binding intensity.

PRINCIPAL COMPONENT ANALYSIS DIFFERENTIATES HISTONE MODIFICATIONS BETWEEN FAMILIES, SUGGESTING A GENETIC INFLUENCE ON CHROMATIN STATE

An understanding of the difference between two families can be effectively achieved by evaluating differences (or variance) between the two families as they relate to differences (or variance) within each family. The final statistical approach utilizes analysis of variance (ANOVA) and weighs the inter-family normalized by the intra-family variance. The ANOVA approach is typically applied to univariate analysis. Thus, high dimensional data generated from the ChIP-on-chip assays described here pose specific challenges. Analysis of such high dimensional data requires reduction in complexity in order to effectively understand the variance structure. We used principal component analysis (PCA) for this purpose.

PCA was carried out via two approaches. In the first, total ChIP intensity of both alleles together (A+B) was calculated. This tells us how much DNA at a particular locus, without regard to allele, is bound to a given modified form of chromatin. The results indicate clustering of samples based on chromatin state, as indicated by antibodies to specific proteins (Figure 2A). For example, the open symbols representing active chromatin modifications cluster in the lower left of the graph. Measurement of total locus intensity (A +B) can also be used to evaluate allele specificity in certain situations, depending on the genotypes involved and the study design. However, for our purposes, given the small size of our study, a second approach, which is more sensitive for elucidating allelic differences in chromatin state, uses the relative intensity ($A/A+B$), a method that is discriminatory only for that 30% of SNP sites that are heterozygous. This method tolerates noise introduced by sample preparation and other technical exposures that are expected to affect both alleles more or less equally.

PCA based on measurement of relative intensity showed that the samples from family 1 and family 2 were separated from each other into two clusters (blue versus red symbols Figure 2B). The differences observed between these two families reflect their genetic differences, based on allelic variation. The separation was largest for the antibodies targeting active chromatin (open symbols). Thus, the global chromatin states as measured by the relative ChIP intensities differ for the individuals in family 1 versus the individuals in family 2. This observation led us to conclude that genetic inheritance can influence histone modifications, the hallmark of the epigenetic phenomena.

Our data offer a novel method by which one can “epigenotype” individuals in a manner analogous to conventional genotyping of genomic DNA for purposes of cancer risk assessment. Thus, our “AS ChIP-on-chip” technique has the theoretical potential for use in risk prediction in clinical settings of familial presentations of cancer. Furthermore, epigenetic changes are more likely than genetic changes to be reversible by appropriate interventions, including nutrients as well as drugs. Hence, our approach to risk assessment in familial settings has the additional value of suggesting epigenetic targets for directed interventions to reverse risk, with an ultimate goal of preventing cancer.

Acknowledgment

We thank our colleagues Drs. Mitsutaka Kadota, Howard H Yang, Nan Hu, Chaoyu Wang, Ying Hu, Philip R Taylor, Kenneth H Buetow for their contribution to this study.

Funding. This research was supported by the Intramural Research Program of the NIH and the National Cancer Institute.

REFERENCE

1. Kadota M, Yang HH, Hu N, Wang C, Hu Y, Taylor PR, Buetow KH, Lee MP. Allele-specific chromatin immunoprecipitation studies show genetic influence on chromatin state in human genome. PLoS Genet. 2007;3:e81. [PubMed: 17511522]

Figure 1. A) The AS ChIP assay shows allele-specific chromatin binding by Pol II and histone H3 protein at the differentially methylated region (KvDMR) in the imprinted *LIT1* promoter.

An allele-specific oligo ligation assay (OLA) was used to detect allele-specific ChIP activities. The two peaks represent allele-specific ChIP activities at the C allele (left) and the T allele (right). GM10858, GM10859, GM10861, GM10870, GM11872, GM11875, and GM11882 are seven CEPH samples from CEPH pedigrees family id 1347 and 1362 that are heterozygous at SNP (rs11023840). **B) Modified 10K SNP chip protocol for the AS ChIP-on-chip experiments.**

Abbreviations: input, DNA from whole cell extract; Pol, RNA polymerase II; Ac, H3Ac; K4, H3K4; K9, H3K9; K27di, H3K27 di-methylation; K27tri, H3K27 tri-methylation.

Figure 2. Clustering of the samples by antibody with total chromatin binding activity (A+B) vs. clustering of the samples by family with the relative allelic chromatin binding activity (A/A+B).

A) Principal component analysis (PCA) was performed with the total ChIP intensity (A+B) for the samples in CEPH family 1362 (red) and 1347 (blue). The clusters formed by experiments with different types of the antibodies are enclosed with ellipses. **B)** PCA was performed with the relative binding intensity (A/A+B) using the same ChIP data as (A).

Abbreviations: Ac, histone H3 lysine 9/14 acetylation; K4, histone H3 lysine 4 di-methylation; Pol, RNA polymerase II; K27di, histone H3 lysine 27 di-methylation; K27tri, histone H3 lysine 27 tri-methylation; K9, histone H3 lysine 9 di-methylation; in, DNA from whole cell extract; gDNA, genomic DNA.