Characterizing the DNA Methylome of Osteosarcoma

Jiayi M. Sun, Jianhe Shen, Howard Lin, Aaron Taylor, Horatiu Voicu, Alex Yu, Joseph Luan, Rudy Guerra, Chris Man, Ching C. Lau

Structural and Computational Biology and Molecular Biophysics, Baylor College of Medicine, Texas Children's Cancer and Hematology Center, Rice University, Dan L. Duncan Cancer Center

Introduction

Osteosarcoma is the most common malignant bone cancer in children and adolescents with approximately 400 cases diagnosed per year. It occurs primarily in the distal femur and proximal tibia but can also occur in the upper limb long bones and pelvis. Standard treatment for osteosarcoma begins with neoadjuvant therapy followed by tumor resection. Percent tumor necrosis as a response to chemotherapy serves as a prognostic marker and guides the choice of post operative chemotherapy. Patients who respond positively to pre-op chemo (>90%) have a favorable 5-year overall survival rate of approximately 70%. Patients with poor response (<90%) will receive either dose-intensified therapy or different agents for post-op chemo. Despite efforts to improve the outcome of these high-risk patients through various clinical trials, there's been very little improvement in patient survival over the past two decades. These patients typically have a 40% 5-year survival rate and patients with metastasis have an even worse overall survival of approximately 20-25%.

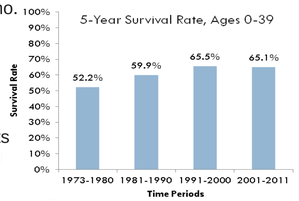


Figure 1: Osteosarcoma 5-year survival rate in individuals Age 0-39 (SEER Incidence and Survival 1975-2011)

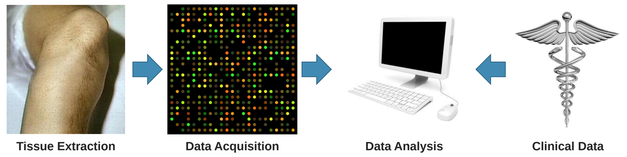
There are two main reasons for this:

- Lack of reliable biomarkers for predicting prognosis and identifying high-risk patients at the time of diagnosis
- Osteosarcoma has a high degree of genomic instability making identification of robust biomarkers and therapeutic targets difficult

Biological and clinical behavior of cancers are determined by the aberrations the cancer cells have acquired. Genetic aberrations can be used as biomarkers to predict clinical outcome or identify therapeutic targets. Some of these markers can be used as predictors of chemoresistance or metastasis at diagnosis, offering the opportunity to customize therapy upfront to counter these high risk features at an earlier time. We undertook the NCI-funded TARGET (Therapeutically Applicable Research to Generate Effective Treatments) initiative to accomplish two goals through comprehensively characterizing the osteosarcoma genome:

- Identify biomarkers for prediction of high-risk osteosarcoma (e.g. chemoresistance or metastasis) that can be used to stratify patients at the time of diagnosis to alternative treatments
- Identify novel therapeutic targets, especially for high-risk osteosarcoma

Methods



Samples in our completed discovery set were collected from the Children's Oncology Group (COG), Texas Children's Cancer and Hematology Center and other collaborating institutions around the world. All samples have matched clinical outcome data from COG's Statistics and Data Center. All patients in this study have been treated with the same protocol (neoadjuvant chemotherapy, tumor resection, post-operative chemotherapy). All tissue DNA and RNA have gone through full QC/QA evaluation histopathologically and molecularly.

References and Funding

1. Günesdogan Ufuk, Jäckle H, Herzig A. Histone supply regulates S phase timing and cell cycle progression. eLife 2014; 3: e02443.
2. Zhao Y, et al. Abnormal methylation of seven genes and their associations with clinical characteristics in early stage non-small cell lung cancer. Oncol Lett 2013; 4: 1211-1218.
3. Litovkin K, et al. Methylation of PITX2, HOXD3, RASSF1 and TDRD1 predicts biochemical recurrence in high-risk prostate cancer. J Cancer Res Clin Oncol 2014; 11: 1849-61.
4. Irizarry RA, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 2009; 2: 178-86.

Funding: National Cancer Institute's TARGET Initiative, CPRIT Grant, Keck Center, Gulf Coast Consortia National Library of Medicine (NLM), Gillson Longenbaugh Foundation, Anderson Foundation, SCBMB, Baylor College of Medicine

DNA Methylation Profiling Overview

DNA methylation is a heritable epigenetic modification that regulates gene expression through adding methyl groups to cytosines. A large concentration of these cytosines are located within CpG islands which are regions of the genome rich in 'CG' dinucleotide repeats. CpG islands occur commonly in the promoter region of genes and DNA methylation within these regions typically leads to gene silencing through the recruitment of methyl-binding proteins and compaction of the surrounding chromatin structure. We sued Illumina's Inifinium HumanMethylation450 (HM450) BeadChip for DNA methylation profiling. This array assesses over 485,000 unique CpG sites across the genome.

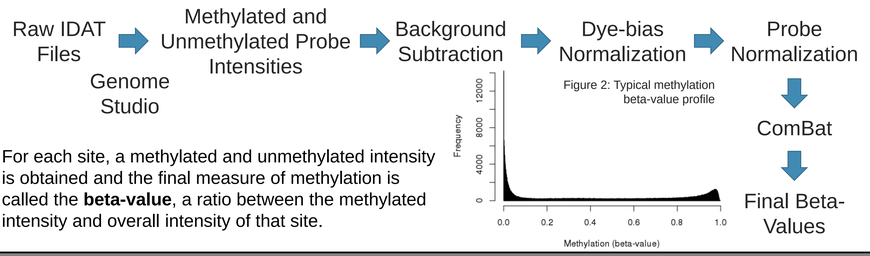


Figure 2: Typical methylation beta-value profile

For each site, a methylated and unmethylated intensity is obtained and the final measure of methylation is called the **beta-value**, a ratio between the methylated intensity and overall intensity of that site.

Unsupervised Clustering

DNA methylation profiling has been used for subtype classification before, notably in the CpG island methylator phenotype (CIMP) first studied in colorectal cancer. We used non-negative matrix factorization (NMF), an unsupervised clustering method, on the top 1500 CpG probes with the highest standard deviation. Results from NMF were quality assessed through looking at the measures such as the cophenetic coefficient, dispersion and silhouette analysis as well as comparison to a randomly permuted set of input data. Through this we determined that three clusters was the best fit. The probes which were significant between the groups (ANOVA multiple testing corrected p-value < 0.05) were plotted in the heatmap seen in Fig. 3.

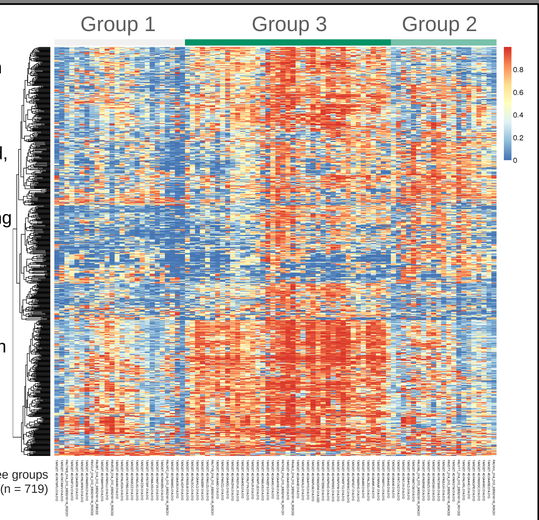


Figure 3: ANOVA significant probes between three groups from NMF (n = 719)

Methylation-based Subgroups Are Clinically Significant

We further identified that the NMF groups have a clinically significant survival difference. Groups 1 and 2,apart from one sample generally have better prognosis while the samples in Group 3 which is characterized by noticeably increased methylation (hypermethylation) have a much worse overall and event-free survival rate. Group 2 has several early cases of death and relapse similar to Group 3, however after a certain time point they have much better prognosis and end up more similar to Group 1 in terms of survival.

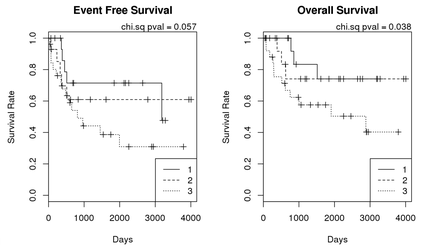


Figure 4: Kaplan-meier survival curves of NMF 3 group membership

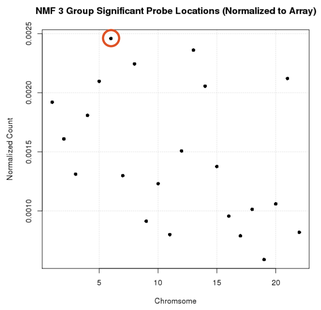


Figure 5: Proportion of significant probes per chromosome, chromosome 6 is indicated with the red circle

Concentration of Probes in Chromosome 6

When analyzing the ANOVA significant probes, we noticed a disproportionate number of probes in chromosome 6. Some of these probes were located very close to each other suggesting the presence of an entire region being differentially methylated. Chr6 aberrations have been reported previously in osteosarcoma, mostly related to the region around RUNX2, a transcription factor involved in osteoblast differentiation. The region we discovered however is located upstream of this within a cluster of histone genes and is significantly hypermethylated in Group 2 compared to Group 1 and 3. Our hypothesis is that upregulation of histone mRNA plays a positive role in the resulting better prognosis of these patients. Histone mRNA is strictly regulated during cell division and this observation may serve as an indicator of an alternative phenotype[1]. Fig. 6 shows the differentially methylated regions in chromosome 6.

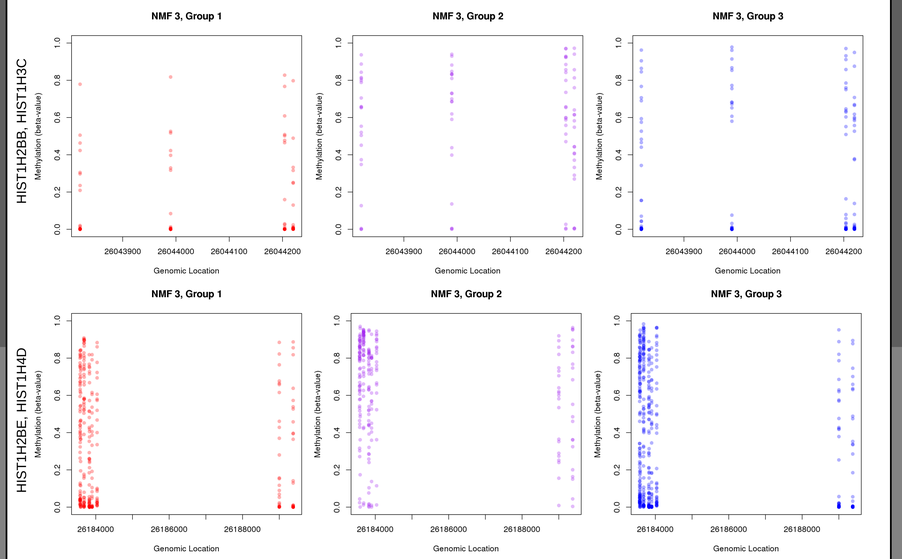


Figure 6: Differential methylation within two regions on chromosome 6 overlapping a cluster of histone genes.

Potential Biomarkers for Poor Prognosis

To identify probes which distinguish samples with worse prognosis, such as those in Group 3, we used the *glmnet* package in R which utilizes a penalized lasso regression to fit generalized linear models. This method tries to reduce the number of covariates which predict the given model leading to dimensionality reduction. In our case, we modeled the overall and event-free survival using a cox proportional hazards model.

Through this we identified two genes, NPAS3 (neuronal PAS domain protein) and PITX2 (paired-like homeodomain transcription factor 2). PITX2 was differentially methylated between Group 3 (the worst prognosis group) and the other groups by the previous ANOVA test as well. The literature suggests that hypermethylation of PITX2 corresponds to an increased risk for metastasis in cancers such as prostate and lung cancer[2-3]. NPAS3's role in cancer is not well known however some evidence in astrocytomas suggest a tumor suppressive role.

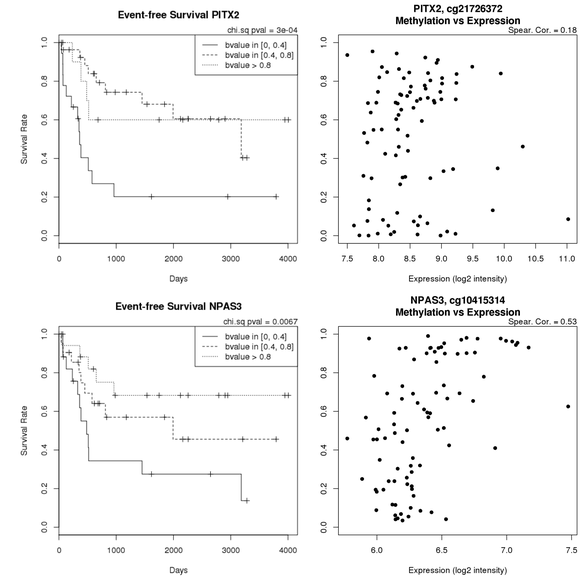


Figure 7: Kaplan-meier survival curves stratifying PITX2 and NPAS3 methylation beta value (Left) Methylation and expression correlation (Right)

Our results corroborate previous studies of PITX2 with hypermethylation corresponding to better prognosis however our results suggest that hypermethylation of NPAS3 is beneficial, contrasting its tumor suppressive role in the literature (Fig. 7). The shape of the Kaplan-meier curve for PITX2 indicates an on-off trigger while the methylation of NPAS3 suggests more of an ordinal increase. Both NPAS3 and PITX2 have a positive relationship to expression. The significant probe for PITX2 is located within a CpG island outside of the gene promoter and the probe for NPAS3 also located in the gene body. These PITX2 and NPAS3 probes seem to be potential biomarkers for osteosarcoma prognosis given the relationship between methylation beta-value and survival outcome however their positive methylation to expression correlation suggests a more complicated interaction[4].