Comprehensive Profiling of the Osteosarcoma Genome

Jiayi M Sun[1,2], Jianhe Shen[2,3], Alex Yu[2,3], Joseph Luan[2,3], Horatiu Voicu[2,4], Rudy Guerra[3,5], Chris Man[2,3,4], Ching C Lau[1,2,3,4]

[1]Structural and Computational Biology and Molecular Biophysics, Baylor College of Medicine, [2]Texas Children's Cancer Center, [3]Department of Pediatrics, Baylor College of Medicine, [4]Dan L. Duncan Cancer Center, [5]Department of Statistics, Rice University

Background

Osteosarcoma (OS) is the most common malignant bone cancer in children with approximately 400 cases diagnosed per year (~3% of all childhood cancers).[1] The peak incidence is associated with adolescence with the tumor affecting primarily the long bones, such as the distal femur (Fig. 1).

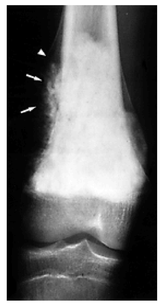


Fig. 1: Typical "sunburst" pattern characteristic of osteosarcoma.[2]

Standard treatment for OS consist of multi-drug chemotherapy followed by tumor resection. Percent tumor necrosis (TN) as a response to neoadjuvant chemotherapy serves as a prognostic marker and guides the choice of post-op chemotherapy. Patients who have a good response (>90% TN) to pre-op chemo have a favorable 5-year overall survival of ~70% and will receive post-op chemotherapy identical to pre-op chemotherapy.[3] Patients with poor response (<90% TN) will receive either dose-intensified therapy or different agents for post-op chemotherapy. Despite numerous efforts to improve the outcome of high risk patients through various clinical trials in the past two decades, there has been no improvements in the 5-year overall survival of < 40% for these patients. Patients with metastasis have even worse overall survival (20%).

Current Knowledge

Several genetic diseases have predisposition for OS:

- **Li-Fraumeni Syndrome**: autosomal dominant hereditary disorder with germline mutations in p53
- **RecQ DNA Helicase**: hereditary genetic disorders with mutations in genes of this family of proteins (Bloom, Werner, Rothmund-Thomson)
- **Retinoblastoma**: mutations in the RB1 gene confer higher disposition for OS

The OS genome is one of the most complex among all cancers (Fig. 2) and consequently, little progress has been made in discovering biomarkers predictive of prognosis. Comprehensive characterization of the genome will aid us in deciphering the relationship between genomic aberrations and tumor biology leading to a better understanding of OS pathogenesis.

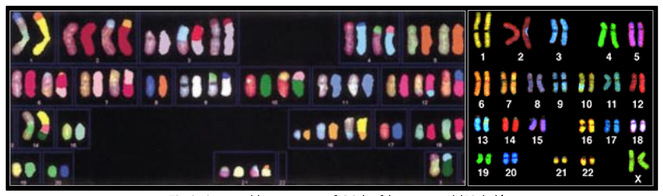


Fig.2: Spectral karyogram of OS (Left) vs normal (Right)[4]

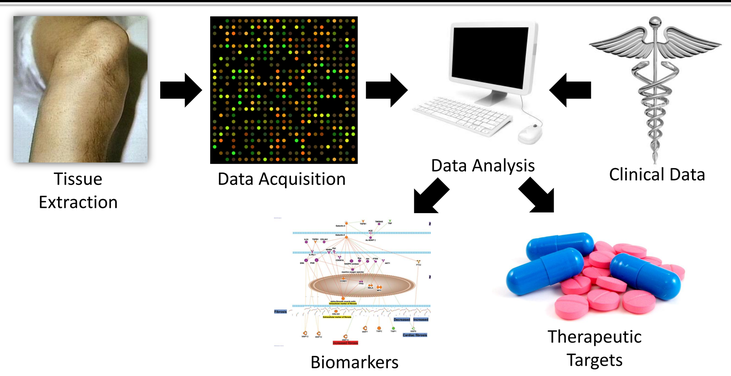
Goals

We hypothesize that the biological and clinical behavior of cancers are determined by the genetic aberrations the cancer cells have acquired. Such genetic alterations could be used as biomarkers to predict clinical outcome or therapeutic targets. We believe that some biomarkers can be used to predict chemoresistance and/or metastatic potential at diagnosis, thus offering the opportunity to customize therapy upfront to counter these high risk features.

We undertook the NCI-funded TARGET (Therapeutically Applicable Research to Generate Effective Therapies) project to accomplish two goals through comprehensive analysis of the OS genome:

1. Identify biomarkers for prediction of high-risk OS (e.g. chemoresistant or metastatic) that could be used at the time of diagnosis for potential stratification of patients to alternative treatments
2. Identify novel therapeutic targets especially for those associated with high risk OS

Analysis Process



Sample Summary

Profiling	Platform	Number of Samples
Copy Number	Affymetrix 6.0 SNP	86
mRNA	Affymetrix Human Exon 1.0 ST	85
microRNA	Life Technologies TLDA card	85
Methylation	Illumina Infinium 450K Array	85
Whole Genome Sequencing (WGS)		48
Whole Exome Sequencing (WES)		86
RNAseq		86

We currently have 85-86 samples for every genomic platform of interest except WGS where we have 48. Samples were collected from the Children's Oncology Group (COG), Texas Children's Cancer Center (TCCC) and other collaborating institutions. 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 (pre-operative chemotherapy, tumor resection, post-operative therapy). All tissue DNA and RNA have gone through full QC/QA evaluation histopathologically and molecularly.

Copy Number

Copy number aberrations were profiled on Affymetrix 6.0 SNP array and data was analyzed using Partek's Genomics Suite. We confirmed copy number changes in genes previously implicated in OS pathogenesis including cyclin-dependent kinase inhibitor 2A (CDKN2A) and cyclin E1 (CCNE1) (Fig. 3).

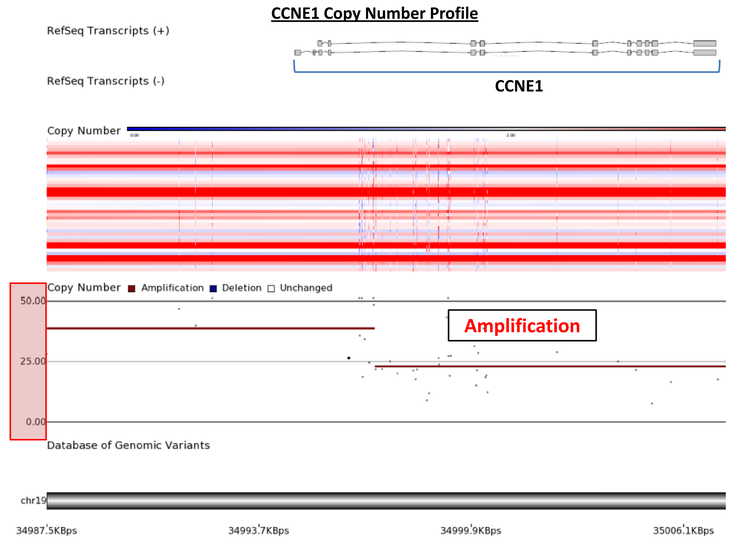


Fig. 3: Copy number profile of CCNE1 a cyclin-dependent kinase across ~40 samples.

Expression

mRNA expression was profiled using Affymetrix Human Exon 1.0 ST array. This platform was used as support when determining how biologically relevant changes in other platforms were. For example, a gene which is aberrantly amplified may seem interesting. However if that gene does not have a corresponding increase in mRNA expression then it would not be as biologically relevant.

Unsupervised clustering revealed two robust clusters (Fig. 4). Further analysis of this data with clinical data may reveal biological subtypes within these membership groups.

Hierarchical Clustering of Expression Data Reveals Two Groups

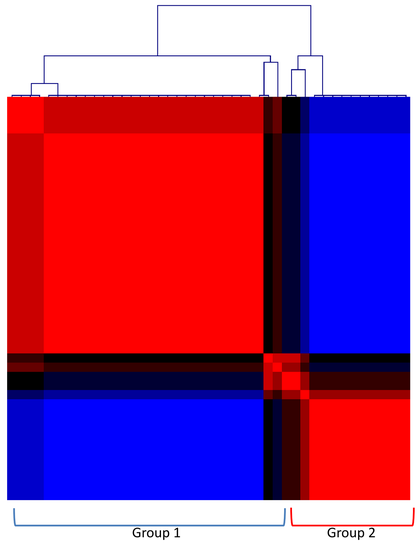


Fig. 4: Unsupervised clustering of human exon data.

MicroRNA

MicroRNAs are short (~22 nt) RNA molecules which bind to mRNA and control transcription. Aberrant microRNA expression has been implicated in many other cancers and therefore we profiled the miRNA expression of these OS patients on TLDA card. Hierarchical clustering revealed two robust clusters of OS based on miRNA profiling data (Fig. 5).

Hierarchical Clustering of miRNA Reveals Two Groups

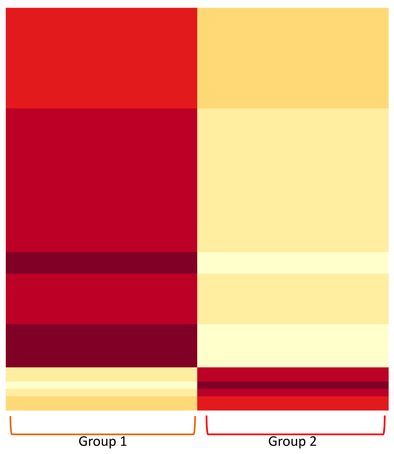


Fig. 5: Unsupervised clustering of miRNA data.

Methylation

Illumina's 450K (HM450k) array was used to profile DNA methylation, a heritable epigenetic mechanism that regulates gene transcription by methylating CpG islands and is known to be involved in controlling cellular differentiation (Fig. 6). CpG methylation can prevent the binding of transcription factors to the gene as well as recruit methyl-binding proteins which further associate with histone modifying proteins that control accessibility to the genome by altering chromatin structure.

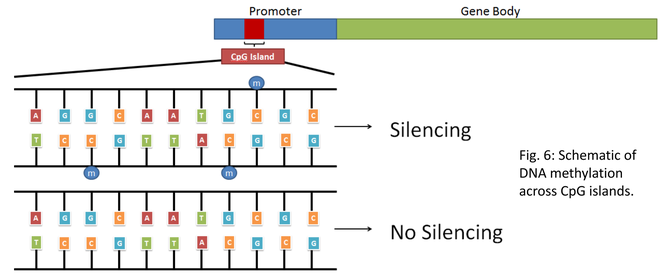


Fig. 6: Schematic of DNA methylation across CpG islands.

The HM450k array has probes throughout the gene including areas within the gene body (Fig. 7). In differentiated cells, the general methylation profile across the promoter region is low to allow for gene transcription while the methylation in the body is high.

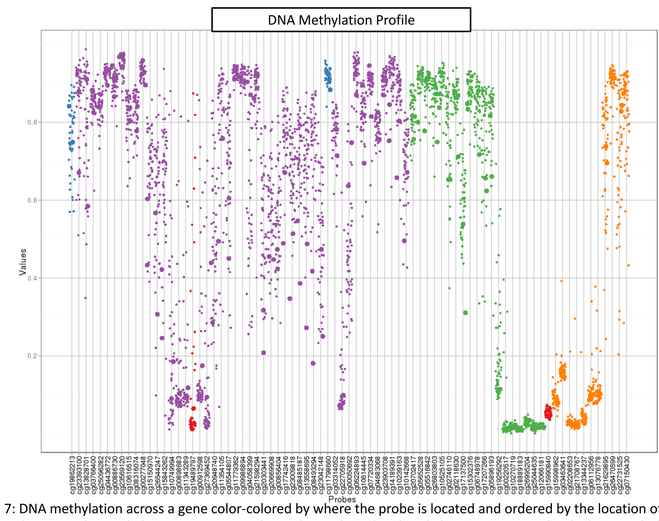


Fig. 7: DNA methylation across a gene color-colored by where the probe is located and ordered by the location of the probe along the gene. The methylation across several samples is plotted with a slight jitter to more clearly distinguish the points

The large degree of variation in methylation within samples across a single probe and across a single gene, in conjunctions with the sheer amount of probes within this array suggests the need for a more sophisticated algorithm for analysis.

QC/QA control, background subtraction and normalization were performed in R. We have currently explored different methods of summarizing the methylation such as:

- taking only methylation values within the promoter region
- specific probes based on annotation (Gene, CpG Island, etc)

Another possible approach is just to pass over summarization and perform single probe comparisons. Ultimately, the approach we decide on will be both computationally efficient and have sound biological reasoning.

Moving forward currently with the single probe comparison, we identified a set of genes with high correlation with expression. Fig. 8 shows an example of a gene with accordant methylation and expression profiles (high methylation-low expression, low methylation-high expression). We took the copy number into account when identifying these genes. Because methylation is impacted by the number of copies, a gene with good correlation between methylation, copy number and expression needs to be filtered out because the effect on expression is attributed more to copy number.

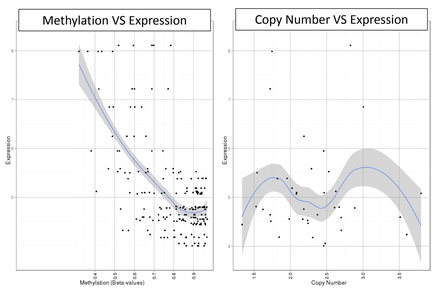


Fig. 8: Methylation VS Expression (Left), Copy Number VS Expression (Right)

Future Plans

- Integrate the different platforms to form a comprehensive view of the genome
- Develop a robust and efficient procedure of analyzing whole-genome methylation data and correlation with expression data
- Correlate genomics data with clinical outcome to identify biomarkers with clinical utility
- Identify potential therapeutic targets
- Validate previous chemoresistance and metastasis signatures or develop new classifiers

References

1. Cancer Facts and Figures 2012, ACS
2. Jerome et al. Chondroblastic osteosarcoma of the lower tibia: a case report. Foot and Ankle Online Journal, 3: 1. 2001.
3. Chou et al. Therapy for osteosarcoma: where do we go from here? Paediatr Drugs, 10:315-27. 2008.
4. Lau et al. Frequent amplification and rearrangement of chromosomal bands 6p12-p21 and 17p11.2 in osteosarcoma. Genes, Chromosomes & Cancer, 39: 11-21. 2004.

This research was funded by a training fellowship from the Keck Center of the Gulf Coast Consortia, on the Training Program in Biomedical Informatics, National Library of Medicine (NLM) T15LM007093, NCI grant (TARGET) NIH-NCI U10 CA98543-07S6, and CPRIT Grant 2532328402