Whole Genome Sequencing and Epigenome Characterization of Cancer Cells using the PacBio Platform

Jonas Korlach[4], Marie Classon[1], Vasantharajan Janakiraman[2], Eric W. Stawiski[2,3], Matthew Boitano[4], Jason Chin[4], Khai Luong[4], Stephen W. Turner[4], Yuta Suzuki[5], Steffen Durinck[2,3], Shinichi Morishita[5] and Somasekar Seshagiri[2]
[1]Department of Oncology, [2]Molecular Biology, and [3]Bioinformatics, Genentech Inc., South San Francisco, CA 94080, [4]Pacific Biosciences Inc., Menlo Park, CA 94025, [5]University of Tokyo, Tokyo, Japan

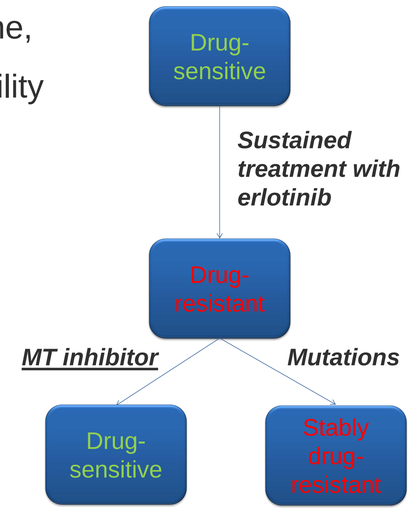
Abstract

The comprehensive characterization of cancer genomes and epigenomes for understanding drug resistance remains an important challenge in the field of oncology. For example, PC-9, a non-small cell lung cancer (NSCL) cell line, contains a deletion mutation in exon 19 (DelE746A750) of *EGRF* that renders it sensitive to erlotinib, an EGFR inhibitor. However, sustained treatment of these cells with erlotinib leads to drug-tolerant cell populations that grow in the presence of erlotinib. However, the resistant cells can be resensitized to erlotinib upon treatment with methyltransferase inhibitors, suggesting a role of epigenetic modification in development of drug resistance.

We have characterized for the first time cancer genomes of both drug-sensitive and drug-resistant PC-9 cells using long-read PacBio® sequencing. The PacBio data allowed us to generate a high-quality, *de novo* assembly of this cancer genome, enabling the detection of forms of genomic variations at all size scales, including SNPs, structural variations, copy number alterations, gene fusions, and translocations. The data simultaneously provide a global view of epigenetic DNA modifications such as methylation. We will present findings on large-scale changes in the methylation status across the cancer genome as a function of drug sensitivity.

Background & Study Design

- PC-9 lung cancer cell line, studying drug susceptibility
- Methyltransferase (MT) inhibitor reverts cells to drug-sensitive



- Utilized PacBio Sequencing for:
  - Generate de *novo* cancer genome assembly
  - Characterize cancer genome structural variation
  - Detect gene fusions
  - Characterize genome-wide methylome

De Novo Assembly & Genome Structure

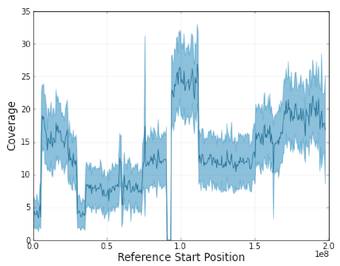
Assembly statistics and comparison to previous short-read cancer genome assembly:

	PacBio	Short-read[1]	Improvement
# of contigs	12,359	424,605	34x
Contig N50	1.044 Mb	0.018 Mb	58x
Max contig length	26.6 Mb	0.28 Mb	95x

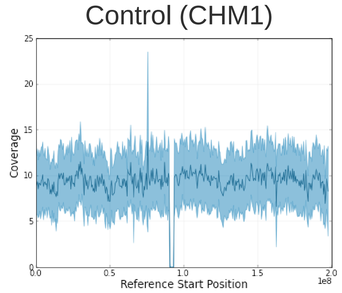
Chromosomal rearrangements

Example: chromosome 3, mapping against hg19:

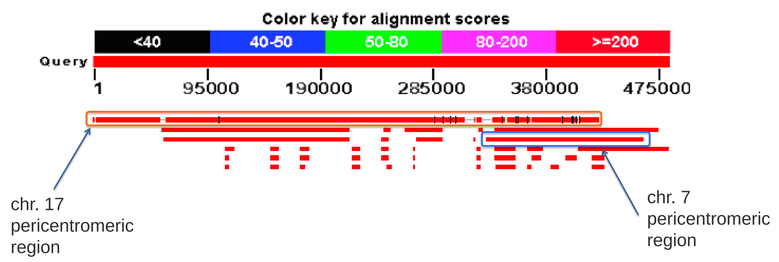
PC-9 cancer cell line



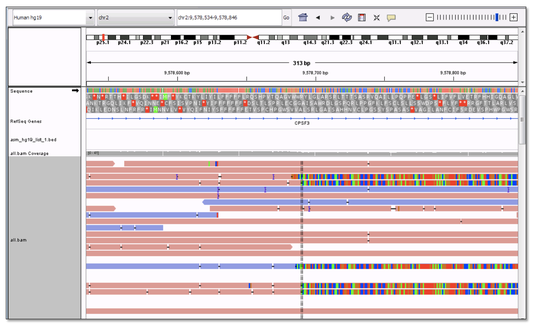
Control (CHM1)



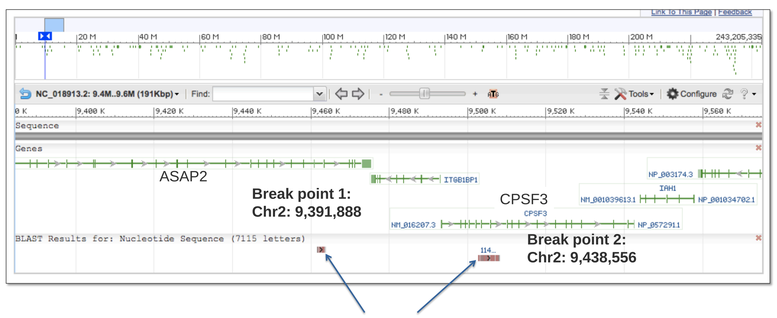
Example: 475 kb assembly contig split between chr. 7 & chr. 17:



Gene fusion detection



- Example: CPSF3-ASAP2 heterozygous gene fusion
- Multi-kilobase anchors on both sides for confident gene fusion detection



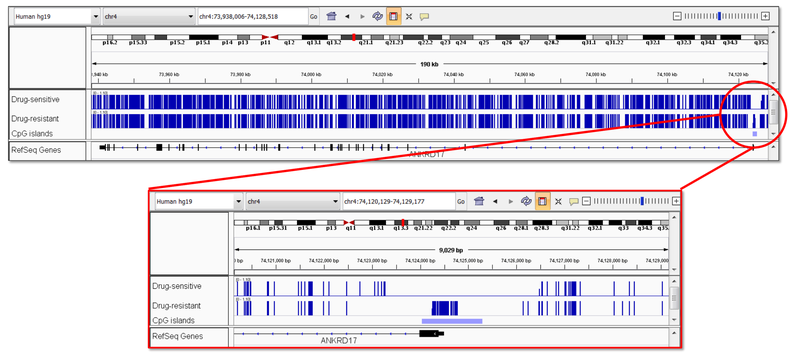
Same read, split-mapped

Cancer Epigenome

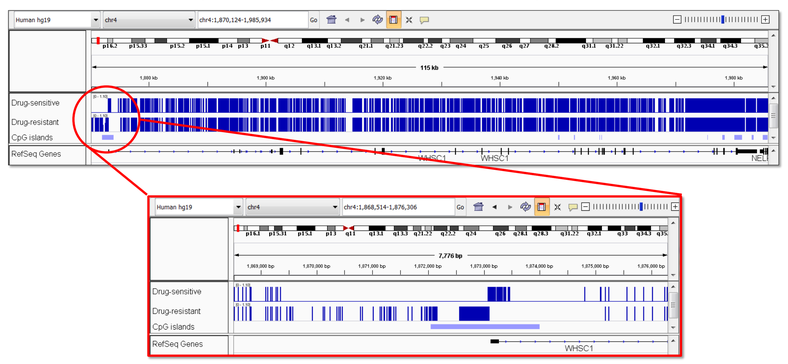
Differential methylation status of CpG islands inferred genome-wide from PacBio Sequencing data, algorithm at https://github.com/hacone/AgIn

Examples:

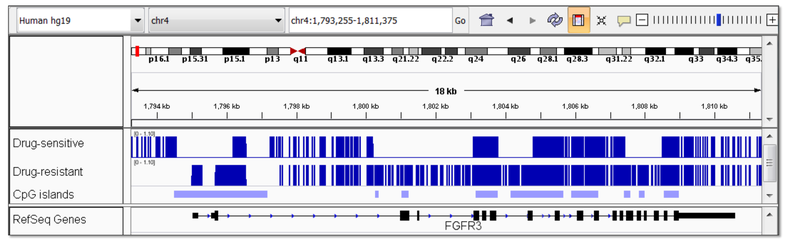
Chr. 4: ANKRD17 (already implicated in breast cancer)



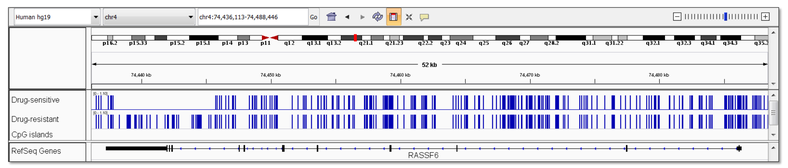
Chr. 4: WHSC1 (already implicated in myelomas)



Chr. 4: FGFR3 (fibroblast growth factor 3)



Chr. 4: RASSF6 (tumor suppressor gene)



References

References

[1]Nagarajan *et al.* (2012) Whole-genome reconstruction and mutational signatures in gastric cancer. *Genome Biology* 13: R115.

Acknowledgements

The authors would like to thank everyone who helped generate data for the poster.