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

Jonas Korlach4, Marie Classon1, Vasantharajan Janakiraman2, Eric W. Stawiski2,3, Matthew Boitano4, Jason Chin4, Khai Luong4, Stephen W. Turner4, Yuta Suzuki5, Steffen Durinck2,3, Shinichi Morishita5 and Somasekar Seshagiri2

1Department of Oncology, 2Molecular Biology, and 3Bioinformatics, Genentech Inc., South San Francisco, CA 94080, 4Pacific Biosciences Inc., Menlo Park, CA 94025, 5University 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.

De Novo Assembly & Genome Structure

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

<table>
<thead>
<tr>
<th></th>
<th>PacBio</th>
<th>Short-read¹</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td># of contigs</td>
<td>12,359</td>
<td>424,605</td>
<td>34x</td>
</tr>
<tr>
<td>Contig N50</td>
<td>1.044 Mb</td>
<td>0.018 Mb</td>
<td>58x</td>
</tr>
<tr>
<td>Max contig length</td>
<td>26.6 Mb</td>
<td>0.28 Mb</td>
<td>95x</td>
</tr>
</tbody>
</table>

Chromosomal rearrangements

Example: chromosome 3, mapping against hg19:

PC-9 cancer cell line  Control (CHM1)

Gene fusion detection

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

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)

Acknowledgements

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

References


For Research Use Only. Not for use in diagnostic procedures.