Comprehensive Detection of Germline and Somatic Structural Mutation in Cancer Genomes by Bionano Genomics Optical Mapping

A.W.C. Pang1, J. Lee1, K. Hong1, T. Anantharaman1, E.T. Lam1, Y. Delpr1, S. Bocklandt1, S. Marin1, A. Hastie1, M. Borodkin1
Bionano Genomics, San Diego, California, United States of America

Abstract

Introduction: The ability to identify structural variants (SVs) is crucial in cancer genetics. Karyotype and cytogenetics are manually intensive. Microarrays and sequencing cannot detect calls in segmental duplications and repeats, and miss balanced variants and low-frequency mutations. Materials and Methods: We describe the Bionano Genomics’s Saphyr platform to identify SVs in cancer genomes. DNA >100 kbp is extracted, labelled at specific motifs, and linearized through NanoChannel arrays. Molecule images are digitized and de novo assembled, creating chromosomal-arm scale genome maps. Cancer mutations >500 bp are detected by aligning the molecules or the genome maps to the public reference.

Results: Over the past 12 months, the power of Bionano’s cancer workflow has been demonstrated on nearly 50 various cancers, including leukemia, breast, ovarian, prostate, pancreatic, among others. While the number of SVs varies among samples, we typically observe >3500 calls per genome. Among leukemia samples, we captured the BCR-ABL1 translocation as well as deletions impacting tumor suppressor genes such as PTPN14 and ESRG. We resolved the structure of large duplications (790 kbp) disrupting BRCA1 in early-onset breast cancers, found the amplification of MYC in lung cancers. Conclusions: In conclusion, with one platform, Saphyr can discover a broad range of traditionally refractory but relevant SVs, and improves our understanding of cancer.

Background

Generating high-quality finished genomes replete with accurate identification of structural variation and high completion (minimal gaps) remains challenging using short read sequencing technologies alone. The Saphyr™ system provides direct visualization of long DNA molecules in their native state, bypassing the statistical inference needed to align paired-end reads with an uncertain insert size distribution. These long labeled molecules are de novo assembled into physical maps spanning the entire diploid genome. The resulting provides the ability to correctly position and orient sequence contigs into chromosome-scale scaffolds and detect a large range of homozygous and heterozygous structural variation with very high efficiency.

Methods

1. Extraction of long DNA molecules
2. Label DNA at specific sequence motifs
3. Saphyr Chip linearizes DNA in NanoChannel arrays
4. Saphyr automates imaging of single molecules in NanoChannel arrays
5. Molecules and labels detected in images
6. Bionano Access software assembles optical maps

(1) Long molecules of DNA are labeled with Bionano reagents by (2) incorporation of fluorophores at a specific sequence motif throughout the genome. (3) The labeled genomic DNA is then linearized in the Saphyr Chip using NanoChannel arrays (4) Single molecules are imaged by Saphyr and then digitized. (5) Molecules are uniquely identifiable by distinct distribution of sequence motif labels (6) and then assembled by pairwise alignment into de novo genome maps.

Conclusions

We demonstrate that the Saphyr system can be used to accurately detect genetic mutation hallmarks in samples with cancer. These include large rearrangements ranging from translocations, within chromosome fusions, to copy number alterations. Researchers can now experiment to uncover somatic variation by comparing with Bionano control sample database, or against a matched pair sample. Furthermore, Bionano SV pipelines can detect SVs with complex breakpoint structures that are recalcitrant to detection by other technologies. Our results indicate that the Saphyr system can capture a broad spectrum of variation with functional importance, and can provide easy solutions for cancer studies.

Reference