Genome-wide profiling of epigenetic markers using Next Generation Sequencing

Client Overview

The client is one of the world's premier research and teaching institutions for modern genetics, gene therapy, molecular evolution, statistical genetics and application of model organisms, to problems in biology, medicine, computational and experimental approaches to genome biology.

Problem Statement

The client, working on cancer and functional genomics, was investigating the role of epigenetic factors, especially differentially methylated CpG loci as key indicators in certain types of cancers in humans. Client was interested in studying the epigenomic changes using next generation sequencing, which could help them identify methylated DNA regions across the entire genome more efficiently than the conventional bisulfite sequence analysis.

However, they needed an efficient informatics support that could help them process and analyze the NGS data to project the implications in terms of biological correlations.

The client was facing the following problems when they approached Optra Systems:

- Unlike MSP and the bisulfate methods, the NGS generated complex datasets which could not be interpreted right away
- The raw data required to be processed using sophisticated and custom-made pipelines for quality assessment, read alignment against a reference genome followed by variant identification
- Statistical tools to bring out biologically significant interpretations
- Mapping the results to known genomic features, locations and representing this data in a way that is easy to comprehend was another challenge

Collecting relevant annotation information with reference to the genome (feature co-ordinates, identity, description, references etc.) to map the positive elements on the NGS reads.

Solution

Optra Systems developed a comprehensive bioinformatics pipeline to process the client's NGS methylation data.

The analysis pipeline was decomposed into five distinct steps:

- Quality assessment of the raw data
- Read alignment to a reference genome
- Variant identification using base-by-base methylation calling
- Annotation of the variants
- Data visualization

The services involved:

- Creating a library of NGS platform compatible tools to perform data trimming and filtering tasks, and create output summary graphs and tables to quickly assess the data quality
- Constructing modules for linking the Quality checked data to Reference genome (UCSC) and San Diego Epigenome Center Reference genome to the Alignment program
- Construction of modules to identify true enrichment sites and mapping them on genomic co-ordinates to identify location of methylation enrichment
- Modules to calculate confidence statistics for methylation enrichment sites
- Implementation and integration of the modules
- Mapping, tabulating, validating and reporting final data

Benefits

- Optra System's technology proficiency and domain expertise assisted the client to reduce the time and cost of operations
- The client was provided with a precise report containing allele-specific epigenetic variation with mapping information and accompanying annotation data
- Drawing biologically coherent and meaningful inferences from the arrays was made easy

OptraHealth LLC.

OptraHealth LLC.
530 Lakeside Drive, Ste 250, Sunnyvale, CA 94085
Tel: +1-408-524-5300, Fax: +1-408-524-5302, Email: info@optrahealth.com

© Copyright 2016 OptraHealth LLC.

Analytic Approach

- Reads from multi fasta files were used to assemble and align them with the reference genome,
- Annotation of the sequences was carried out, followed by base-by-base methylation level check, identifying tool specific reporting regions and methylation browser tracks (IGV, UCSC)
- A window of fixed nucleotide width was used to slide over the entire genome and the average log ratio of the array elements in each window position was calculated
- Based on a cutoff value (of the average log ratio), methylation enrichment sites were flagged
- Confidence statistics was calculated for each enrichment site based on the number of positive array elements occurring in the corresponding scan window
- Genomic co-ordinates were mapped for the enrichment sites and supporting annotation data added

Technology Environment

- Assembly alignment pipeline (Velvet/BOWTIE/BWA).
- Methylation calling by Picard
- Handling different NGS data formats (BAM/SAM, VCF, BED/GFF)
- Generation of tracks for viewing in IGV and UCSC genome browser, and summary statistics.
- Bismark, BSMap/RRBSMap, BS Seeker, MethylCoder, RMAP (SE only) were used for methylation analysis and assembly
- Genomic co-ordinates and transcript annotation data were mined and populated from public sequence databases like NCBI, UCSC etc.

About Optra Health

Optra Health is an ISO-certified global organization with deep domain expertise in medical devices, lab automation, life science informatics and healthcare IT solutions. The company provides a fully-scalable, cost-effective OptiShore™ delivery model. This enables customers to choose the optimal balance between on-site, on-shore, and off-shore development that will best

OptraHealth LLC.

© Copyright 2016 OptraHealth LLC.

address their budget and collaboration requirements. With Optra Health, customers are able to shrink their time-to-market by leveraging practical, building-block based solutions. Committed to clear communication and total transparency, the company consistently meets or exceeds its clients' expectations. Offering a full complement of expert engineering and consulting services, Optra Health is aligned to real business needs applied over the entire product development lifecycle. The robust, scalable and efficient IT infrastructure of the company, together with its outstanding project management team, consistently ensures superior results. Optra Health's global delivery model helps its customers cut costs by about 50% without compromising on quality and realize a 200% improved production cycle.

Visit Optra Health today: http://www.optrahealth.com

Contact Optra Health Today

OptraHealth LLC.
530 Lakeside Drive, Ste 250
Sunnyvale, CA 94085
Tel: +1-408-524-5300; Fax: +1-408-524-5302
Email: info@optrahealth.com

OptraHealth LLC.

© Copyright 2016 OptraHealth LLC.