Weill Cornell Medicine

Core Laboratories Center

Genomics and Epigenomics Core Facility

Genomics Services

ABSTRACT

The Genomics and Epigenomics Core Facility of the WCM Core Laboratories Center (CLC) provides state-of-the-art instruments and services, and expertise in their applications, to the Weill Cornell Medicine (WCM) community and to outside investigators. Resources and services include next generation sequencing, epigenomic assays, single cell genomics, nucleic acid mass spectrometry, and real-time PCR. The facility provides consultation on project design and data analysis, and offers seminars, educational workshops and training.

OVERVIEW

History: The WCM CLC Genomics and Epigenomics Core Facility is the product of an ongoing process of integration of two previously separate core facilities, the Genomics Resources Core Facility (established in 2000) and the Epigenomics Core Facility (established in 2008).

Location: The core is located on the 1st and 4th floors of the "A" building and the 5th floor of the "C' building at 1300 York Avenue, New York, NY.

WCM Core Laboratories Center (CLC): The WCM CLC was established in 2015. In addition to genomics and epigenomics, the CLC includes core facilities that offer resources and services in proteomics and metabolomics, synthetic and analytical chemistry, NMR, flow cytometry, imaging (optical and electron microscopy, high content screening), CBIC (MRI, PET/CT, and ultrasound), biorepository, bioinformatics, and advanced technology assessment.

Next Generation Sequencing: The core facility has 5 Illumina instruments, including HiSeq 4000, 2 HiSeq 2500/1T, HiSeq 2500, NextSeq 500, & MiSeq.

Genomics: NGS genomics applications include exome capture, whole genome and targeted region *de novo* sequencing and resequencing, targeted region and whole transcriptome profiling, small RNA analysis, and metagenomics projects.

Epigenomics: NGS epigenomics applications include (1) DNA methylation profiling, including whole genome bisulfite sequencing (WGBS), targeted methylome sequencing, enhanced reduced representation bisulfite sequencing (ERRBS), genome wide detection of 5-hydroxymethylcytosine modification (5hmC-BIC-seq); (2) assessment of the functional state of chromatin by assaying for transposase activity (ATAC-seq); and (3) genome wide protein-nucleic acid association mapping by chromatin immunoprecipitation sequencing (ChIP-seq).

Single Cell Genomics and Epigenomics: The core offers strong support for single cell genomics and epigenomics projects with a 10X Genomics Chromium system, a Drop-seq instrument, access to a Fluidigm C1 single-cell auto prep platform, and close collaboration with the CLC Flow Cytometry Core Facility

Nucleic Acid Mass Spectrometry: The core offers the Agena Bioscience (Sequenom) Compact MassArray system, a platform that uses MALDI-TOF mass spectrometry to measure PCR-derived amplicons. Currently used for methylation analysis; other applications include genotyping and mutation detection, gene expression analysis, and pathogen detection.

Real-time PCR: An ABI 7900 HT Sequence Detection System provides real-time PCR for quantitative gene expression and SNP genotyping.

Sanger Sequencing & Fragment Analysis: The core offers a sample "drop and ship" location to facilitate WCM investigators' access to Sanger sequencing and fragment analysis services done on two ABI 3730xl DNA Analyzers in the BRC Genomics Facility on the Cornell University campus in Ithaca, NY.

Administration: The Genomics and Epigenomics Core Facility is part of the Weill Cornell Medicine (WCM) Core Laboratories Center (CLC).

Open to all: The resources and services of the core facility are open to all investigators at Weill Cornell Medicine, Cornell University and Cornell-affiliated institutions. The facility also provides services to external investigators at both academic institutions and commercial enterprises.

RESOURCES

Next Generation Sequencing: 5 Illumina NGS instruments, including HiSeq 4000, two HiSeq 2500/1T, HiSeq 2500, NextSeq 500, and MiSeq.

Single Cell Genomics and Epigenomics: 10X Genomics Chromium, custom built Drop-seq instrument, access to Fluidigm C1.

Nucleic Acid Mass Spectrometry: Agena Bioscience Compact MassArray.

Real-time PCR: ABI 7900 HT Sequence Detection System.

General: 2 Covaris S2, 2 Agilent 2100 Bioanalyzers, Agilent tape station, 2 Nanodrop1000 spectrophotometers, 2 Qubit fluorometers, Glomax plate reader, Gel Logic 200 Imaging System, 3 liquid handling robots (Eppendorf epMotion 5075m, Perkin Elmer Zephyr, Hamilton Starlet).

Next Generation Sequencing

Illumina HiSeq 4000

Applications include:

- whole genome sequencing
- targeted region sequencing
- metagenomics
- whole transcriptome
- targeted gene expression
- small RNA profiling
- epigenomics

Illumina HiSeq 4000 can run two 8-channel flowcells in high output mode or two 2-channel flowcells in rapid run mode. Currently, maximum read lengths are paired-end 2 x 150 bp. Depending on the application, run times in high output mode range from 1 to 3.5 days, with a maximum of about 2.5 billion reads per flowcell (5 billion per run). Multiple samples can be pooled into an individual channel.

Illumina NextSeq 500

Illumina NextSeq 500 can provide up 39 Gb per run with maximum paired-end read lengths of 2 x 150 bp. Depending on the application, run time is up to about 30 hours, with a maximum of about 260 million reads per flowcell.

Illumina MiSeq

Provides longer reads that are ideal for:

- *de novo* genomic sequencing
- *de novo* transcriptome sequencing
- structural variation profiling
- metagenomics

The Illumina MiSeq can rapidly run a single sample or pool of samples. Currently, maximum read lengths are 600 bp per read. Depending on the application, run times range from 1 to 2 days, with a maximum of about 50 M reads and 15 GB per flowcell.

Sample Preparation

Sample preparation (sequencing library construction) services are available for next generation sequencing for both DNA and RNA samples. Available instruments include the Covaris S2 for DNA fragmentation and the Sage Sciences Blue Pippen for DNA molecule size selection.

Sample quality assessment/quality control (QA/QC) is available for all services, including checking RNA and DNA sample purity and integrity on the Advanced Analytical Fragment Analyzer, and quantification using the NanoDrop 1000 spectrophotometer or Invitrogen Qubit spectrofluorimeter. The core facility provides recommendations for both DNA and RNA sample QA/QC.

High throughput liquid handling robots are available to support large scale projects that require sample and/or reagent arraying, rearraying, cherry picking, and concentration adjustment. These robots include the Eppendorf epMotion 5075m, Perkin Elmer Zephyr, and Hamilton Starlet.

Long Reads & Single Cell Genomics

10X Genomics Chromium

Applications include:

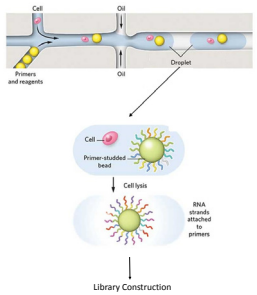
- performs deep profiling of complex cell populations with high-throughput digital gene expression on a cell-by-cell basis.
- provides long-range information on a genome-wide scale, including variant calling, phasing and extensive characterization of genomic structure
- provides long-range information for phasing, structural variant detection and copy number determination
- Provides sample-specific sequence and produces diploid genome structure.

The 10X Genomics Chromium system segregates cells or long DNA molecules (including >100 kb) into individual partitions (i.e., microwells), where DNA in each well can be labeled with a unique barcode sequence, and prepares sequencing libraries in parallel such that all fragments produced within a partition share a common barcode. It combines large partition numbers with a massively diverse barcode library to generate >100,000 barcode containing partitions in a matter of minutes. The short-read data is mapped to the original long molecules using the barcodes. This long range information enables applications impossible with standard short read sequencing, including haplotype phasing and vastly improved structural variant detection. Profiling 1,000s to 10,000s of cells per experiment increases sensitivity and accuracy for the detection of rare cell types.

Single Cell Genomics

Drop-seq

- Massively parallel single-cell RNA sequencing technique
- Performs deep profiling of complex cell populations with high-throughput digital gene expression on a cell-by-cell basis.



The core facility offers a custom built Drop-seq instrument, a droplet-based platform for profiling the transcriptomes of thousands of individual cells at low cost and high speed. The platform includes an inverted microscope, syringe pumps, a magnetic mixing system, and a custom-designed microfluidic droplet generator. Drop-seq quickly profiles thousands of individual cells by separating them into nanoliter-sized aqueous droplets, associating a different barcode with each cell's RNAs, and sequencing them all together. Drop-seq analyzes mRNA transcripts from thousands of individual cells simultaneously while retaining information on the transcripts' cell of origin. The main advantage of Drop-Seq over existing methods is its low cost and speed -- a single investigator can prepare 10,000 single-cell libraries per day.

Nucleic Acid Mass Spectrometry

MassArray

Applications include:

- methylation analysis
- genotyping and mutation detection
- gene expression analysis
- pathogen detection

The Agena Bioscience (formerly known as Sequenom) Compact MassArray is a non-fluorescent detection platform utilizing MALDI-TOF mass spectrometry to measure PCR-derived amplicons.

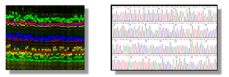
Real-time PCR

Applications include:

- gene expression analysis
- SNP genotyping analysis
- plate format
- real-time PCR mode
- endpoint analysis mode
- melt curve analysis of PCR products

The ABI 7900 HT Sequence Detection System for real-time PCR (rtPCR) allows quantitation of starting material (genomic DNA, plasmid, RNA), along with heterozygote detection and SNP genotyping, using either fluorescent probes or the SYBR green reagent for detection. Core facility users reserve time on the instrument, prepare plates of reactions, and set up the analysis runs. The core facility offers training on the use of this instrument.

Sanger Sequencing and Fragment Analysis



The core offers a sample "drop and ship" location to facilitate WCM investigators' access to Sanger sequencing and fragment analysis services done on two ABI 3730xl DNA Analyzers located in the BRC Genomics Facility on the Cornell University campus in Ithaca, NY. Average data return time is ≤ 30 hours.

Consultation, Workshops and Training

Consultation on project design and data analysis available upon request.

Software analysis tools and training available through the CLC Genomics and Epigenomics Core Facility.

Educational workshops and hands-on training on data analysis available through the CLC Applied Bioinformatics Core (ABC).

Seminars on software tools organized by the CLC Applied Bioinformatics Core.

Seminars on emerging genomics technologies organized by the CLC Advanced Technology Assessment Facility.

Coordinated project design consultation and data analysis support available with the CLC proteomics and metabolomics, synthetic and analytical chemistry, NMR, flow cytometry, imaging (e.g., optical and multiphoton microscopy, high content screening, MRI, PET/CT, and high resolution ultrasound) biorepository, bioinformatics, and advanced technology assessment core facilities.

Contact Information

Genomics and Epigenomics Core Facility

Jenny Xiang, M.D., Director of Genomics Services
jzx2002@med.cornell.edu

Alicia Alonso, Ph.D., Director of Epigenomics Services
epigenomicscore@med.cornell.edu

Doron Betel, Ph.D., Director of Informatics Services
epigenomicscore@med.cornell.edu

http://corefacilities.weill.cornell.edu/genomics
http://epicore.med.cornell.edu

For questions about the WCM Core Laboratories Center, please contact George Grills at ggg2016@med.cornell.edu