FSU-NimbleGen Microarray Facility

Handbook for Users

Handbook of Procedures for Processing Your NimbleGen Samples at Florida State University's NimbleGen Certified Site

Dave Gilbert
Ichiro Hiratani
Dano Fiore
Steve Miller

I. Overview

The FSU NimbleGen Site is part of the FSU Department of Biological Sciences Core Facilities. It is intended to serve FSU and the surrounding area with a place to process their NimbleGen microarray experiments. NimbleGen's superior sensitivity has made it the platform of choice for Sanger Institute, The Epigenome Network of Excellence and the NIH ENCODE project for ChIP-chip and CGH applications. Their recent merge with Roche has led to the development of Sequence Capture, combining NimbleGen array to capture regions of interest with deep sequencing (which can be performed at our neighboring University of Florida's ICBR facility). Of course, the higher sensitivity doesn't hurt for standard transcription array analysis.

The way our facility works is simple. You order the arrays of choice from NimbleGen through a normal purchase order. Your arrays are delivered to FSU. We notify you that we have them. You get the samples to us and we give you back a DVD with your data on it and a little software viewer to browse your data.

II. How do I order my arrays?

If you want to use an existing platform:

- Go to http://www.nimblegen.com/.
- Click on the application you wish to use (see choosing an application, below).
- Each application has a slightly different path to follow, but basically you will either go next to a kingdom and species (gene expression), or you will go to an array type and then see if your species is available in that array type (CGH, ChIP-chip, and DNA methylation).
- Request a quote from Michael Schilling mschilling@nimblegen.com that will indicate "FSU NimbleGen Certified Site" – you must identify yourself as an FSU core user.
- Set up a purchase order
- Order your arrays
- They should be on campus within 2-3 weeks.
- Notify FSU core that you have placed the order.

III. How do I process my sample?

The facility is managed by Dano Fiore (fiore@bio.fsu.edu) and Steve Miller (smiller@bio.fsu.edu) and is located in 2031 KIN (phone: 850-644-8956). You should arrange to have your samples ready close to the time when the arrays are expected to arrive. It will be necessary for you to label your samples with a unique identification for each sample corresponding to each array. You may then submit your samples to Steve. When the arrays arrive on campus, Steve will know which sample it belongs to and will hybridize them accordingly. He will then contact you to get the DVD with your data when ready.

Note that Steve will check your sample for quality control and if there is a problem with your sample, he will contact you. It will be your responsibility to either go ahead with a bad sample, or replace your sample with one that can pass quality control checks. We will do the best we can to advise you of what is wrong with your sample.

IV. How much sample do I need?

- Transcription analysis: 2.5µg of cDNA at >100ng/µl and A260/280 >1.8
- ChIP-chip, DNA methylation or CGH: ~3µg of DNA at >250ng/µl and A260/280 >1.8

V. How much does it cost?

We are in the process of setting up a pricing schedule. Please contact us to obtain current prices. You must purchase your arrays from NimbleGen, at the price they quote for the application you desire (Fall 2008 prices are appx. $500 for the 385K design and appx. $1,000 for the 2.1M design but these prices change frequently, there are various types of discounts, and prices for 4-plex and 12-plex vary-see section VIII). Hence, to be clear, you will receive two charges for this service – one from NimbleGen for the array itself, and one from us for the service.

VI. Can I process my own sample?

Yes, we will give you a discount per hybridization if you learn to process your own samples in house.

VII. After I get my millions of bits of data, what do I do?

There are several options for you. First, you may consult with people working in laboratories that crunch data on a regular basis. If you are working with gene expression, your data will already be reduced to one value per gene, which will fit into an Excel file. If you have someone in your lab who is computer savvy, the best thing to do is to learn to program in "R", and then your people can work with Bioconductor, a freeware package from the NIH. Finally, we have placed on the FSU NimbleGen web site, a link where you can advertise for a computer science or statistics student DIS or OPS. There are many students in Statistics, Computational Biology, Computer Science, Computational Science and Mathematics who will check this site regularly looking for opportunities. The compustat group – a group of interacting faculty and students started by Prof. Gilbert – has started a Wiki on which they intend to set up links for everyone to find information about how to crunch data – information that is directly relevant to FSU people.

VIII. Choosing an application.

The first thing you need to consider is whether NimbleGen has a commercially available array for your purpose, or whether you will have to make a custom design. Another great

advantage of NimbleGen is the ease with which you can make a custom design (the technology is "nimble"). But of course you can get started earlier if there is already a design for you. Available species and applications include:

Transcription: *these arrays contain various exons of the most well annotated genes, as well as some of the less well annotated genes. You hybridize cDNA synthesized from total RNA to the arrays and get back a list of the relative abundance of transcripts for each gene.*

Arabadopsis, C. elegans, Drosophila, Human, Mouse, Rat, budding and fission yeast, Zebrafish

ChIP-chip*: In this application, you ChIP your protein of interest and hybridize it to an array that will have probes 1-2 kb upstream and downstream of the promoter of each annotated gene (Promoter tiling array) or probes that cover the entire genome at one probe per 100 bp (for these arrays, you will need several to cover the entire genome, but you can choose one pre-existing array if you are focusing on one particular chromosome region .*

High Density Promoter Tiling array (2.1 million features): presently only available for human and mouse

Low Density Promoter Tiling array (385,000 features): Arabidopsis, human, mouse and rat

High Density Whole Genome Tiling Arrays (2.1 million features): C. elegans, Drosophila, human, mouse and rat

Low Density Whole Genome Tiling arrays (385,000 features; obviously requires more arrays to cover the genome): Arabidopsis, C. elegans, dog, Drosophila, E. coli, Chicken, Human ENCODE regions, total Human genome, Mouse, Rat, budding yeast.

DNA methylation: *precipitate methylated DNA with either anti-methylated DNA antibodies (MeDip method) or with the methyl DNA binding protein. Hybridize that DNA to the same chip-chip type arrays described above.*

Same arrays as for ChIP-chip.

Comparative Genomic Hybridization (CGH)*: In this method, probes are evenly spaced across the entire genome, at low or high density. This distribution of probes is ideal for evaluating copy number changes in genomes (DNA loss or amplification) or any features of chromosomes that can be evaluated at the resolution afforded with one entire genome on one chip.*

High Density: Human

Low Density: Cow, C. elegans, Dog, Zebrafish, Drosophila, Chicken, Human, Macaque, Mouse, Plasmodium, Rat, Fission and Budding yeast.

Sequence-capture: *This is a very new technology in which you enrich for the sequences of interest by capturing them on an array and then sequence them by deep sequencing methods – this reduces your cost for the sequencing considerably. This application must be discussed individually at this time.*

Custom Design: *So you don't find what you need? For $1,000 you can design anything you want – new species are encouraged! Keep in mind that you can design your array to have the entire array hybridized with one sample, or have your array divided into 4, or 12 separate regions so that your single array can be hybridized with as many different samples. This is useful if you anticipate being able to put 4 or 12 replicates or different experimental conditions on a single chip and save some money, providing a fourth or* 12^{th} *of the chip is enough features for you. Alternatively, you can do different species, or different samples on the same chip. For the design fee to be worthwhile, you should have several experiments in mind. If you have more than 10 experiments in mind, NimbleGen's policy is to waive the design fee. Don't forget to include lots of control features on your array!!*