

# NCGR Sequencing Center

## FAQ

### Awards

NCGR awarded 2009 Bio-IT World Best Practices in Basic Research: Schizophrenia studies  
Analysis Software *Alpheus* awarded 2009 Laureate by Computerworld

### Certified Service Provided

- First Illumina Certified Service Provider (CS-Pro™) for Sequencing in North America (1/9/09)  
<http://investor.illumina.com/phoenix.zhtml?c=121127&p=irol-newsArticle&ID=1242467&highlight=>

### Premiere Organization

- First in the world to break 20 billion base pair run using single Illumina flowcell (Jan 2009)
- #1 Service Provider in completed runs (over 450)
- Top 10 world-wide based on number of Illumina Genome Analyzers
- Top 20 world-wide based on number of all available next-generation sequencing platforms (454-FLX, ABI, and Illumina).
- Large satisfied customer base both commercial and academic

### Our Publications/Articles/Webinars

*A highly annotated whole-genome sequence of a Korean individual* Nature 2009

<http://www.nature.com/nature/journal/vaop/ncurrent/full/nature08211.html>

*Alternative isoform regulation in human tissue transcriptomes* Nature 2008 [PMID: 18978772](#)

*Management of High-Throughput DNA Sequencing Projects: Alpheus* J Comp Sci Syst Biol 2008 [\[Full Text\]](#)

*Genomic Convergence Analysis of Schizophrenia: mRNA Sequencing Reveals Altered Synaptic Vesicular Transport in Post-Mortem Cerebellum* PLoS One 2008 [PMID: 18985160](#) [\[Abstract\]](#)

*Transcriptome sequencing of malignant pleural mesothelioma tumors.* Proc Natl Acad Sci U S A. 2008 Mar 4;105(9):3521-6. Epub 2008 Feb 26. [PMID: 18303113](#) [\[Abstract\]](#)

BioInform: [http://www.bioinform.com/issues/12\\_45/features/150731-1.html?CMP=OTC-RSS](http://www.bioinform.com/issues/12_45/features/150731-1.html?CMP=OTC-RSS)

MarketWatch: <http://www.marketwatch.com/news/story/JMPR-Genomics-NCGR-Partnership-Foster/story.aspx?guid={7AC9DE36-B6AA-4EDE-9CD5-633B29FE6154}>

Webinar: "Using mRNA-Seq to Explore Agricultural Transcriptomics" by Dr. Greg May. See <http://www.illumina.com/webinars/archives/index.ilmn>

### Sequencing Platform/Services

Eight (8) Illumina Genome Analyzer II's. Six (6) with Paired End Modules

- Over 400 runs / Over 50 satisfied customers
- On-site full-time Illumina Field Service Engineer
- IT Bottleneck solved w/millions \$ worth of compute Infrastructure + IT and Bioinformatics Leadership/Expertise
- LIMS delimited process - Grindstone
- mRNA, Whole Genome Shotgun & ChIP Sequencing
- DGE tagged and small RNA sequencing w/customer provided libraries
- Single and Paired End

# NCGR Sequencing Center

## FAQ

Alpheus® variant and expression detection pipeline for Sequence Analysis

- Over 50 instances/customers
- Web-based for worldwide analysis
- Alignment to any species + transcriptome/genome alignments
- JMP-Genomics interface for statistics
- Kognitio WX2 database
- GSNAP alignments for long reads SE/PE

### Results

Standard Output (base calling and quality scores):

FASTQ files of passing reads (with script to convert to FASTA and qual files) are delivered via a secure FTP site (Gigabyte scale).

Non-standard Output (image files):

If desired, raw image files can be mailed on a hard-disk for an additional fee (Terabyte scale).

### Archiving Policy

Once data is delivered, raw images are discarded and standard output is retained approximately 4 weeks. Special data persistency requests for an additional fee.

### Phred-like quality scores for passing reads

40 = .01% Probability of error

30 = .1%

20 = 1%

10 = 10%

Results delivered typically between 30-10 range.

### Sequencing and Analysis Compute Infrastructure (leave the iron to us)

Raw data storage: Sun X4500 - A Dual Processor, Dual Core 24 TB file Server.

Image processing/Base calling: Sun X4150 – Two 4-core processors, 8G RAM, four 146G Hard drives

Alpheus Variant detection: Four Sun Blade 6000 chassis with ten 6220 Blades, each blade has two dual core processors with two 146G disks and 4G RAM. A total of 160 Cores, 160G of RAM and 11.6 TB disk.

Alpheus DB storage: Sun V490 DB Servers, 6140 SAN with 10 TB for DB disk space, 4 Gb/s Switch fabric

### Turn around time

"Standard"

\*\*\*\*\*

# NCGR Sequencing Center

## FAQ

Sequencing : : Upon receipt of samples, our turn around time for one flowcell (7 lanes) is approximately (depending on cycles)

SE: ~ 6 weeks if you make the libraries

SE: ~ 8 weeks if we make the libraries

PE: ~ 8-10 weeks if you make the libraries

PE: ~ 10-12 weeks if we make the libraries

Remaining results follow shortly on a TBD schedule. Large projects, a custom schedule will be developed.

Alpheus : : Upon receipt of sequencing data (FTP site or sequencing order), the Alpheus web-based analysis system will be set up with a customer defined reference and pipeline alignment runs will be performed on batch units in ~2 weeks.

"Express"

\*\*\*\*\*

For a 20% upcharge, upon receipt of the samples, we can turn-around ~ half the standard time:

SE: ~3 weeks if you prepare the libraries

SE: ~4 weeks if we prepare the libraries

PE: ~4-5 weeks if you prepare the libraries

PE: ~5-6 weeks if we prepare the libraries

Remaining results follow shortly on a TBD schedule. Large projects, a custom schedule will be developed.

### Sample Material Requirements

For mRNA Seq (Whole Transcriptome Shotgun) we can prepare the libraries and you can send us either of these

- total RNA (10 ug)
- polyA RNA (100 nano grams)
- fragmented cDNA (4 ug)

For Whole Genome Shotgun & ChIP

- DNA (5-10ug)
- ChIP DNA (>= 10ng)

### Experimental protocol

\* Fragmentation

1) Nebulization: We only use the nebulizer on very large DNA fragments (e.g. BACs) or genomic DNA. Large pieces of DNA break up efficiently using the nebulizer.

2) Random priming: All of our mRNA-Seq libraries are generated using a concentration of random primer that result in the synthesis of short cDNA fragments. If we are given cDNA instead of RNA, we use the random priming approach as well to generate short fragments. The reason for this is that short

# NCGR Sequencing Center

## FAQ

cDNAs do not shear efficiently in the nebulizer. cDNA libraries generated using this approach would not give even coverage on short cDNAs.

\* mRNA seq protocol (standard Illumina protocol).

- 1) Isolate polyA mRNA from total RNA using two rounds of polyA purification
- 2) Generate 1st and 2nd strand cDNA using random hexamer primers
- 3) End repair performed on cDNA products.
- 4) Add A-base performed on End repair products.
- 5) Illumina adapters ligated onto A-base-added products.
- 6) Ligation products are gel purified on a 2% agarose gel.
- 7) Gel purified products are PCR enriched.
- 8) Libraries are quantified on both a nanodrop and bioanalyzer.
- 9) Flowcell construction and sequence analyses performed using Illumina's standard protocols.

\* Once libraries are generated, they are purified on a gel and a gel slice is excised creating fragments of the same approximate size range:

Single End: 200bp size range INCLUDES Illumina adapter sequences.

Paired End (default): 200bp size range INCLUDES Illumina adapter sequences. Please contact us if you want a larger insert size.

### Minimum order

\$10k

### More information

Sequencing : : <http://sequencing.ncgr.org>

Alpheus : : <http://alpheus.ncgr.org>

### We Deliver

Excellent service and 100% customer satisfaction.

### Contact

Faye Schilkey at [fds@ncgr.org](mailto:fds@ncgr.org) or 505-995-4449 for any questions and pricing.