# The Heflin Center for Genomic Sciences Genomics Core Facility

Molly Bray, Ph.D

Michael R. Crowley, Ph.D

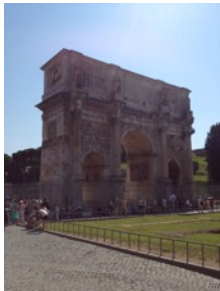
David Crossman, Ph.D

# Our Mission

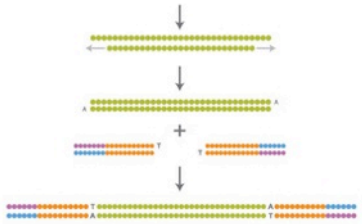
To provide cutting edge genetic and genomic technologies and expertise to the UAB Community

# Microarray Technology and Services

- Affymetrix Resources
  - 3000 7G Scanner - gene expression, genotyping, and promoter arrays
- Illumina Array Resources
  - iScan - High density and high throughput genotyping, gene expression and methylation arrays
  - BeadXpress - Medium density and throughput genotyping and methylation assays
  - Robotic fluidics for sample preparation



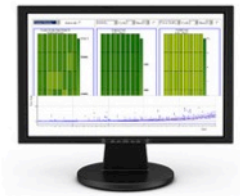
# HiSeq2000

Library Preparation
<6 h (<3 h hands-on)

Cluster Generation
<4 h (<10 min hands-on)

Sequencing by Synthesis
1.5-8 days (<10 min hands-on)



RTA v1.7, CASAVA v1.7
2 days (30 min hands-on)

**HiSeq 2000 Performance Parameters***

| Read Length | Run Time | Output |
|---|---|---|
| 1 × 35 bp | ~1.5 days | 26–35 Gb |
| 2 × 50 bp | ~4 days | 75–100 Gb |
| 2 × 100 bp | ~8 days | 150–200 Gb |

*Sequencing output generated with a PhiX library and cluster densities between 260,000–347,000 clusters/mm² that pass filtering on a HiSeq 2000.

**Throughput**

Up to 25 Gb per day for a 2 × 100 bp run.

**Reads**

Up to one billion clusters passing filter, and up to two billion paired-end reads.

**Performance**

HiSeq 2000 provides the greatest yield of perfect reads and bases greater than Q30:

Greater than 90% bases higher than Q30 at 2 × 50 bp**

Greater than 85% bases higher than Q30 at 2 × 100 bp**

**Human genome at supported cluster densities

Human Genome is ~3 billion base pairs or
3Gb of sequence information

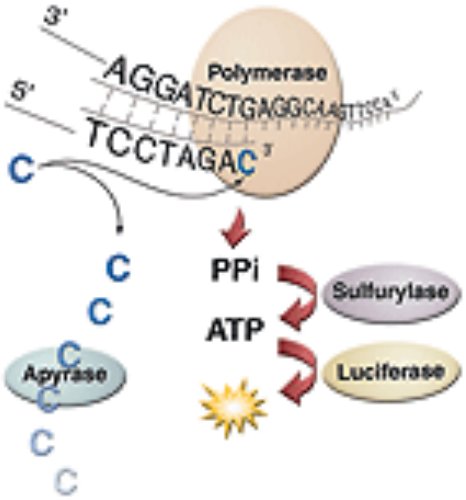
The core runs 4 exomes/lane at 2X50bp reads
generating ~35X coverage/sample

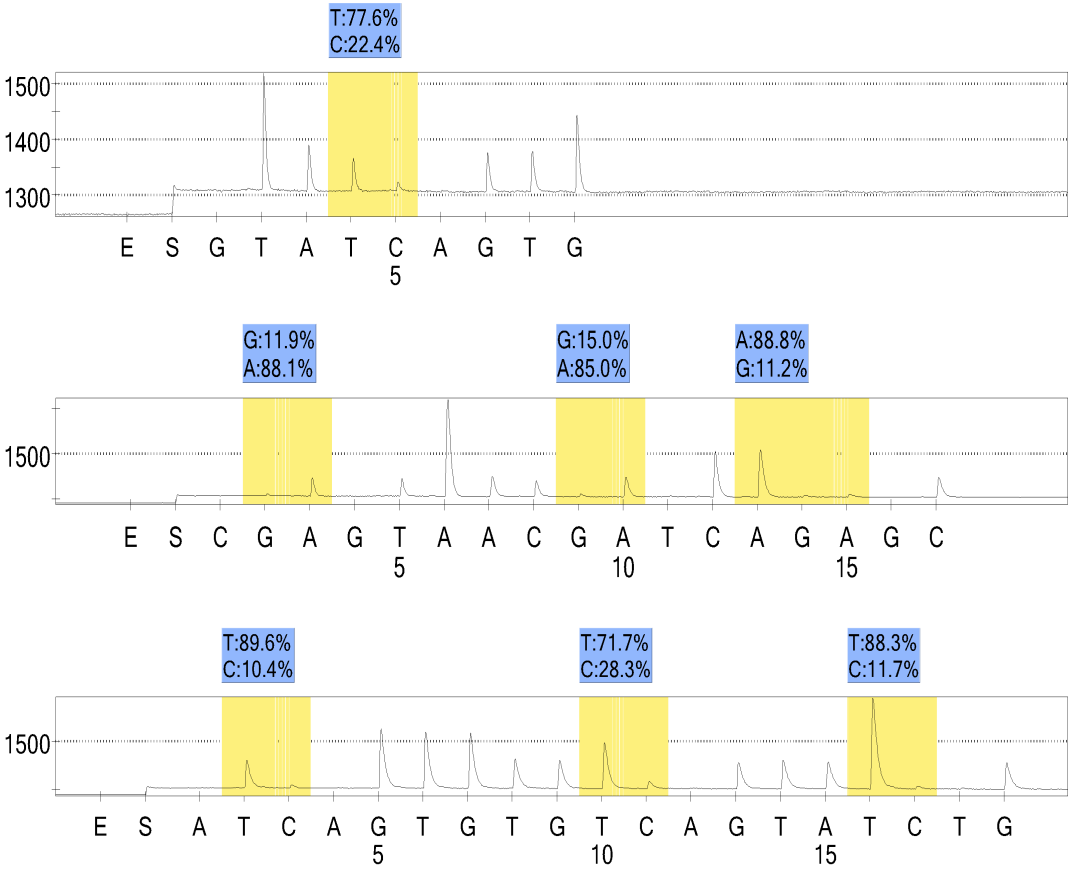
The HiSeq can generate up to 600 billion bases
of sequence information in about 10 days
An equivalent of 6 human genomes at 30X coverage

# Techniques for Epigenetic Profiling

- Whole Genome Bisulfite Sequencing
- Reduced Representational Bisulfite Sequencing (RRBS)
- MeDIP
- ChIP-Seq
- SureSelect Methyl Capture
- Illumina Methyl 450 Arrays

# Pyrosequencing




454 Sequencing is different in that it is whole genome while this pyrosequencing is at the gene/SNP level

# Microarray vs. NGS: Sample Quality and Quantity

| Microarray | | Next Generation Sequencing | |
|---|---|---|---|
| **Assay** | **Amount** | **Assay** | **Amount** |
| Gene expression | ~300 ng total RNA | mRNA-Seq | 1µg of Total RNA |
| Methylation | 500 ng DNA | Methyl-Seq | 100ng* of ChIP'd DNA |
| ChIP arrays | ChIP'd DNA | ChIP-Seq | 100ng* of ChIP'd DNA |
| Genotyping | 200-400 ng DNA | Whole exome | 2-3µg DNA |
| | | Whole genome | 1µg DNA |

*Minimum amount

All assays work best with good quality, un-degraded sample.

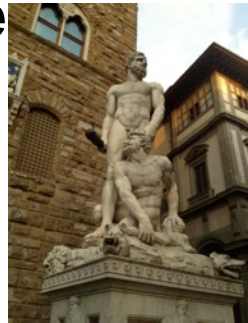
Illumina has developed "rescue" reagents for FFPE samples that can greatly increase yield.

# Microarray vs. NGS: Cost Comparison

| Microarray | | Next Generation Sequencing | |
|---|---|---|---|
| **Assay** | **Cost*** | **Assay** | **Cost*** |
| Gene expression | $245-$650 | mRNA-Seq | $650§ |
| Methylation | $365 | Methyl-Seq | $650-2200† |
| ChIP-Chip | $550 | ChIP-Seq | $650-2200† |
| Genotyping | $80-$620 | Whole exome | $800§ |
| | | Whole genome | $5,000ª |

*Prices include labor and consumables and are subject to change
§Price reflects running 28 samples per flowcell on HiSeq2000
†Prices reflect running several samples per lane v. one sample per lane
ªPrice is for running 3 genomes per flowcell

**NGS analysis is always going to provide more comprehensive data than an off-the-shelf microarray.**



# Thank You

Contacts:

Mike Crowley
mcrowley@uab.edu
5-2013

David Crossman
dkcrossm@uab.edu

Molly Bray
mbray@uab.edu