5-2015

The Cancer Genome Atlas (TCGA): Breast and Ovarian Cancers

Laura Ann Riccio
University of Rhode Island, lriccio@my.uri.edu

The Cancer Genome Atlas (TCGA): Breast and Ovarian Cancers

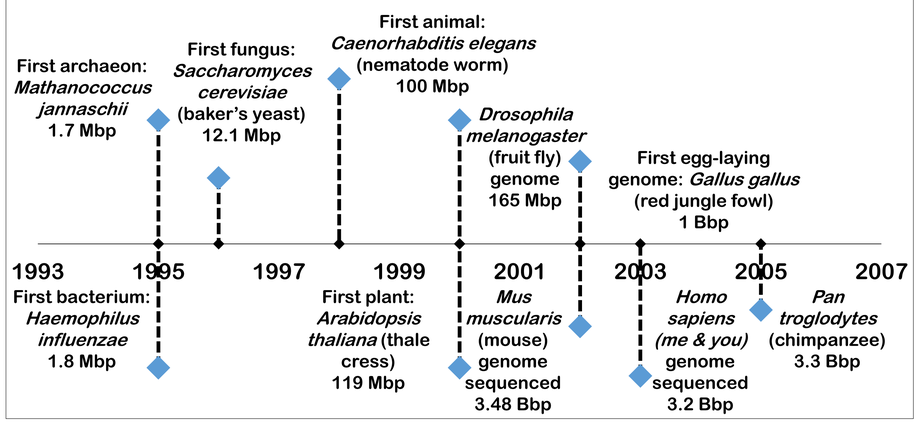
Laura Ann Riccio

Faculty Sponsor: Dr. Niall G. Howlett

Introduction to Genomics

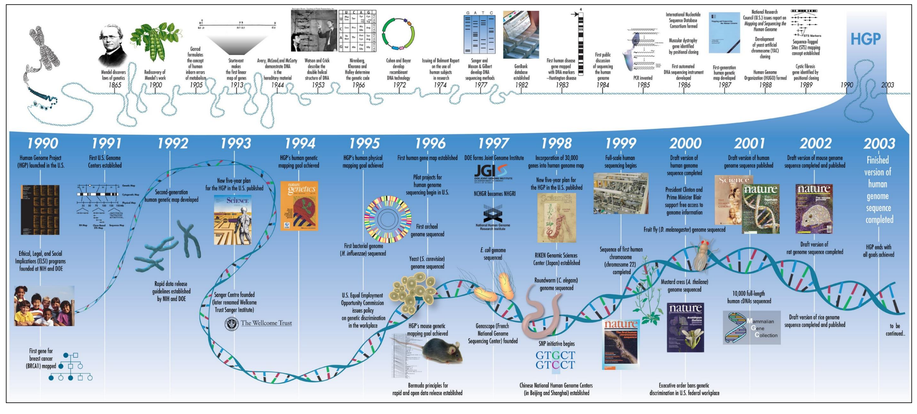
A genome is the complete ordered sequence of DNA bases (A, C, G, T) which make up all of the protein- and RNA-coding genes, and regulatory sequences necessary for the construction of an organism. The field of genomics began in the 1970's when Walter Fiers and his team in Ghent, Belgium sequenced the genome of the M2 bacteriophage.

Significant Genomic Sequencing Events from 1993-2005



Human Genome Project

HGP was an international, collaborative research program with the goal to sequence and map the full human genome of 3 billion base pairs to gain a better understanding of all of the genes that are present. It was initiated in 1990 and a draft sequence was published in 2003, paving the way for the development of new genomics-based research projects.



The Cancer Genome Atlas

The National Institutes of Health (NIH) and the National Human Genome Research Institute (NHGRI) initiated a pilot project called The Cancer Genome Atlas (TCGA) in 2006. The overall goal of TCGA is to catalog all of the significant genomic changes in the major types and subtypes of cancer. It is hoped that this catalog of information will serve as a critical resource for the prevention, diagnosis, and treatment of these cancers.

Key: Estrogen receptor (ER) and human epidermal growth factor receptor 2 (HER2) are cell surface proteins that bind to hormones such as estrogen and epidermal growth factor, respectively, in tissues such as the uterus and breast. Progesterone receptor (PR) is a protein found in the cytoplasm of breast epithelial cells, that binds to the hormone progesterone. These proteins regulate cell division and proliferation.

TCGA Breast & Ovarian Cancers

825 samples of breast cancer tumor & normal tissue

489 samples of stage II-IV high-grade serous ovarian cancer (HGS-OvCa) tumor & normal tissue

Whole Genome/Exome Sequencing
mRNA, miRNA, protein expression analysis
DNA methylation analysis
Chromosome copy-number variation analysis

Genes Mutated in HGS-OvCa

Genes	Type of Mutation(s)	No. of Mutated Samples	Percent of Mutations
TP53	Missense	302	96%
RB1	Missense	6	2%
NF1	Missense	13	4%
BRCA1	Germline, somatic	11	3%
CSMD3	Missense	19	6%
CDK12	Nonsense, indel, missense	9	3%
FAT3	Missense	19	6%
GABRA6	Missense	6	2%
BRCA2	Germline, somatic	10	3%

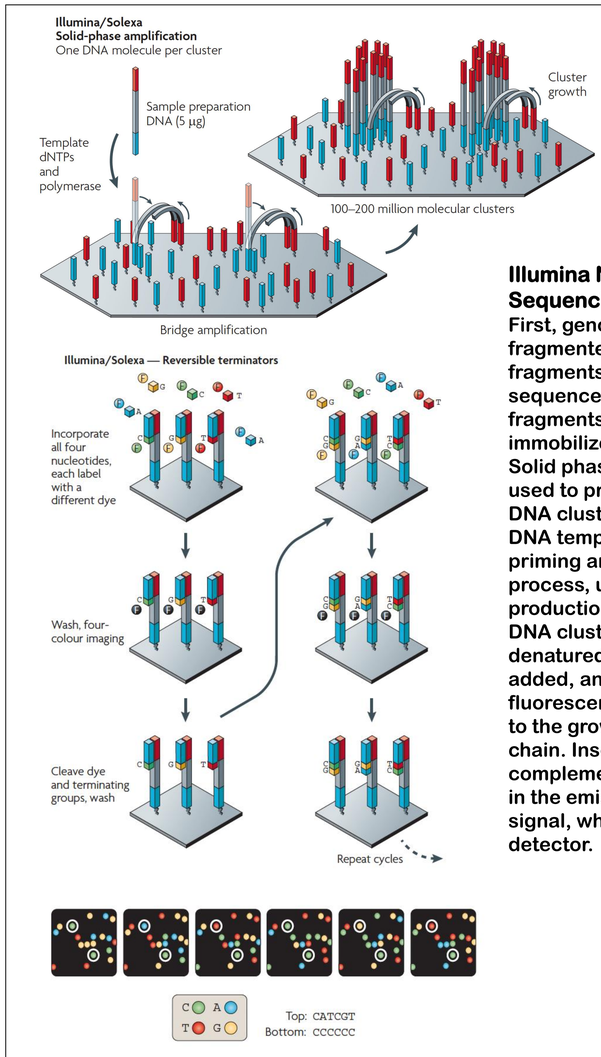
Genes Mutated in Breast Cancer mRNA Subtypes

Genes	Luminal A (n=225)	Luminal B (n=126)	HER2-enriched (n=57)	Basal-like (n=93)
PIK3CA	45%	29%	39%	9%
MAP3K1	13%	5%	4%	0%
GATA3	14%	15%	2%	2%
TP53	12%	29%	72%	80%
CDH1	9%	5%	5%	0%
MAP2K4	7%	2%	2%	0%
PTEN	4%	4%	2%	1%
AKT1	4%	2%	2%	0%
RB1	0.4%	3%	0%	4%
MLL3	8%	6%	7%	5%
TBX3	3%	4%	0%	1%
RUNX1	5%	2%	4%	0%
CBFB	2%	2%	2%	0%
AFF2	1%	2%	5%	4%
PIK3R1	0.4%	2%	4%	0%
PTPN22	0.4%	2%	5%	0%
PTPRD	2%	4%	4%	1%
NF1	2%	4%	0%	2%
CTCF	4%	2%	2%	1%
FOXA1	2%	2%	2%	0%
SF3B1	3%	0%	4%	1%
NCOR1	5%	2%	0%	2%
CDKN1B	1%	1%	2%	0%

Luminal A	Luminal B	HER2E	Basal-like
ER+	ER+	ER-	ER-
PR+	PR+	PR-	PR-
HER2-	HER2+	HER2+	HER2-

Next-Generation Sequencing

The Illumina next-generation sequencing platform was used for both high-grade serous ovarian cancer (HGS-OvCa) and breast cancer. It uses a sequencing by synthesis (SBB) technology which incorporates four fluorescently-labeled nucleotides to sequence DNA clusters.



Illumina Next-Generation DNA Sequencing

First, genomic DNA is purified and fragmented into <500 bp fragments. Universal adaptor sequences are ligated to the fragments. Fragments are then immobilized on a solid support. Solid phase amplification is then used to produce clonally amplified DNA clusters. A single-stranded DNA template is created through a priming and bridge amplification process, ultimately resulting in the production of clonally amplified DNA clusters. DNA clusters are denatured, adaptor primers are added, and DNA polymerase adds fluorescently labeled nucleotides to the growing polynucleotide chain. Insertion of a complementary nucleotide results in the emission of a fluorescent signal, which is collected by a detector.

What Has Evolved From This?

One of the most significant findings from TCGA was that basal-like breast cancer subtype is most similar to HGS-OvCa, containing similar types and frequencies of mutations. These results suggest that the two cancers may have a similar molecular origin, and may be responsive to similar therapies.

Predominant similarities:

- High frequency of *TP53* mutations
- Inactivation of *BRCA1*
- Amplification and high expression of *cMYC*
- Loss of *RB1*

References

David R. Bentley (2000) *Hum. Mol. Genet.* **9**(16): 2353-2358.
Michael L. Metzker (2010) *Nature Rev. Genet.* **11**: 31-46.
The Cancer Genome Atlas Research Network (2012) *Nature.* **490**(7418):61-70.
The Cancer Genome Atlas Research Network (2011) *Nature.* **474**(7353):609-615.