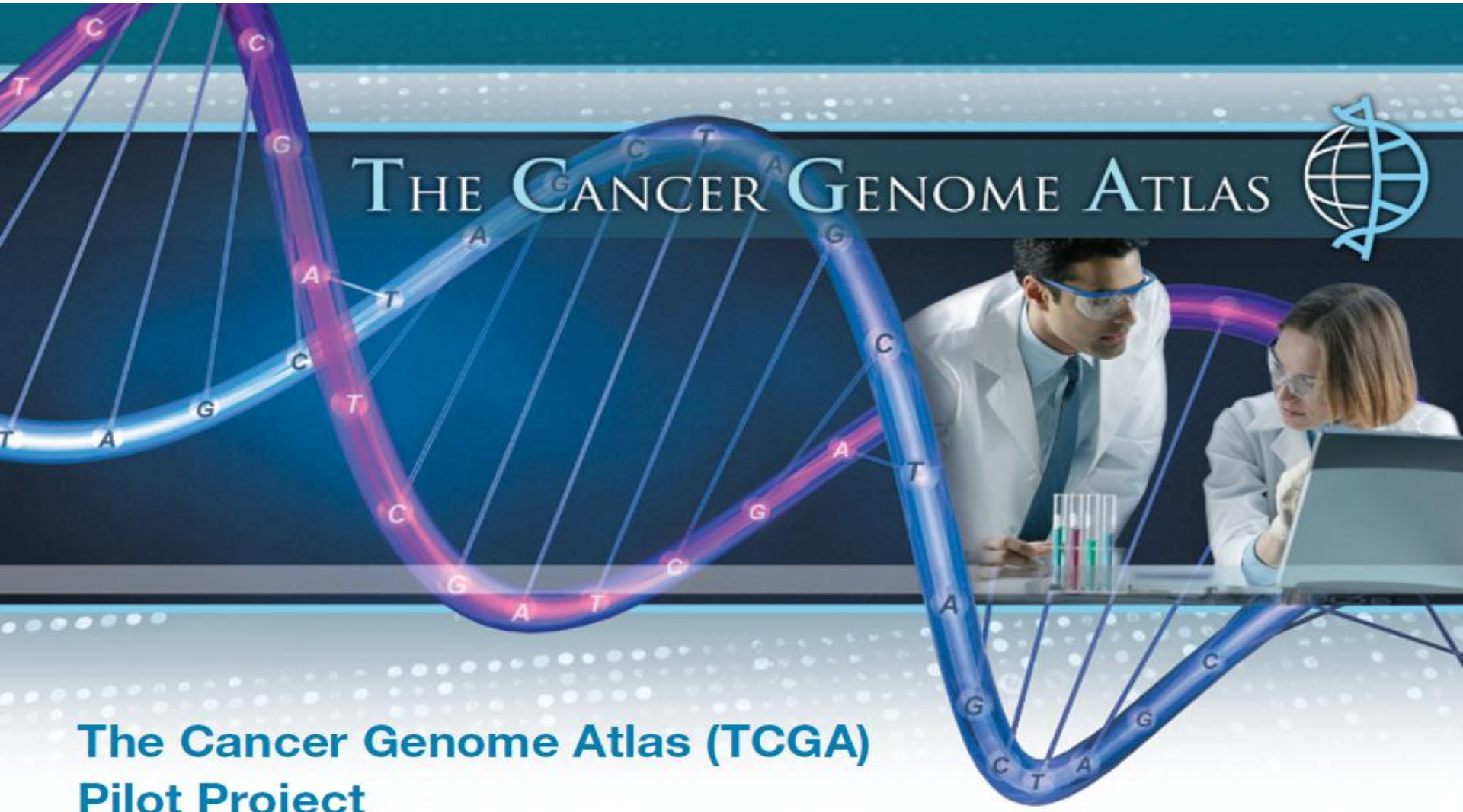
The cancer genome atlas: TCGA data mining

Chi-Heum Cho, M.D., PhD

Keimyung University Department of Obstetrics and Gynecology



The Cancer Genome Atlas (TCGA) Pilot Project

Charting a new course for prevention, diagnosis, and treatment of cancer

What is the Cancer Genome Atlas?

1. The Cancer Genome Atlas (TCGA) is a comprehensive and coordinated effort to accelerate our understanding of the molecular basis of cancer through the application of genome analysis technologies, including large-scale genome sequencing

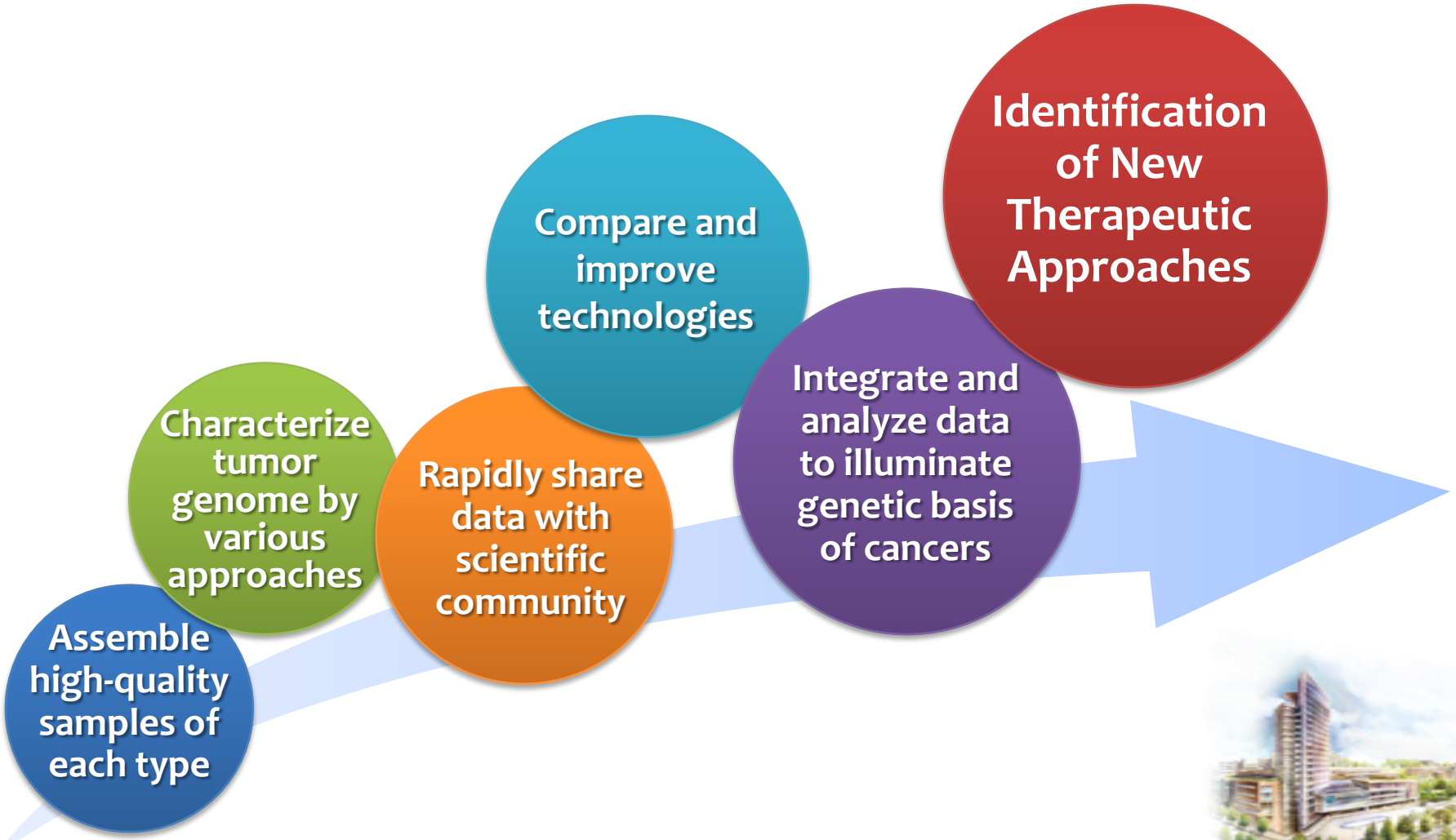
2. TCGA is a joint effort of the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI), which are both part of the National Institutes of Health, U.S. Department of Health and Human Services

Goals of The Cancer Genome Atlas

1. **Comprehensive Catalog of Cancer Genome (Structural)**
   - Know all driver genes in all types of cancer
   - Know how all driver genes correlate with clinical phenotype

2. **Cancer Therapeutic Roadmap (Functional)**
   - Recognize functional pathways in which targets function
   - Know cancer vulnerabilities, as function of cancer genome (targets)
   - Know resistance mechanisms, as function of cancer genome (combinations)

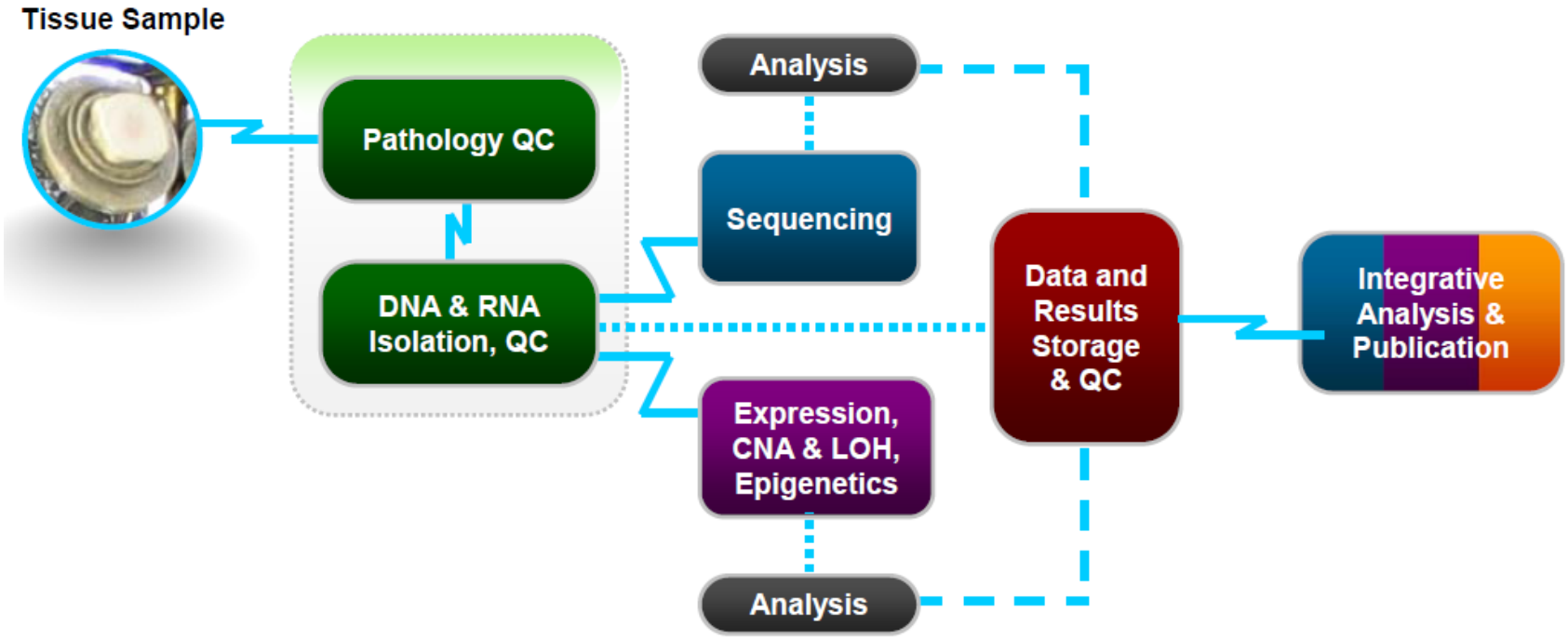
Goals of The Cancer Genome Atlas (TCGA)



TCGA: Pipeline for Comprehensive Characterization

THE CANCER GENOME ATLAS

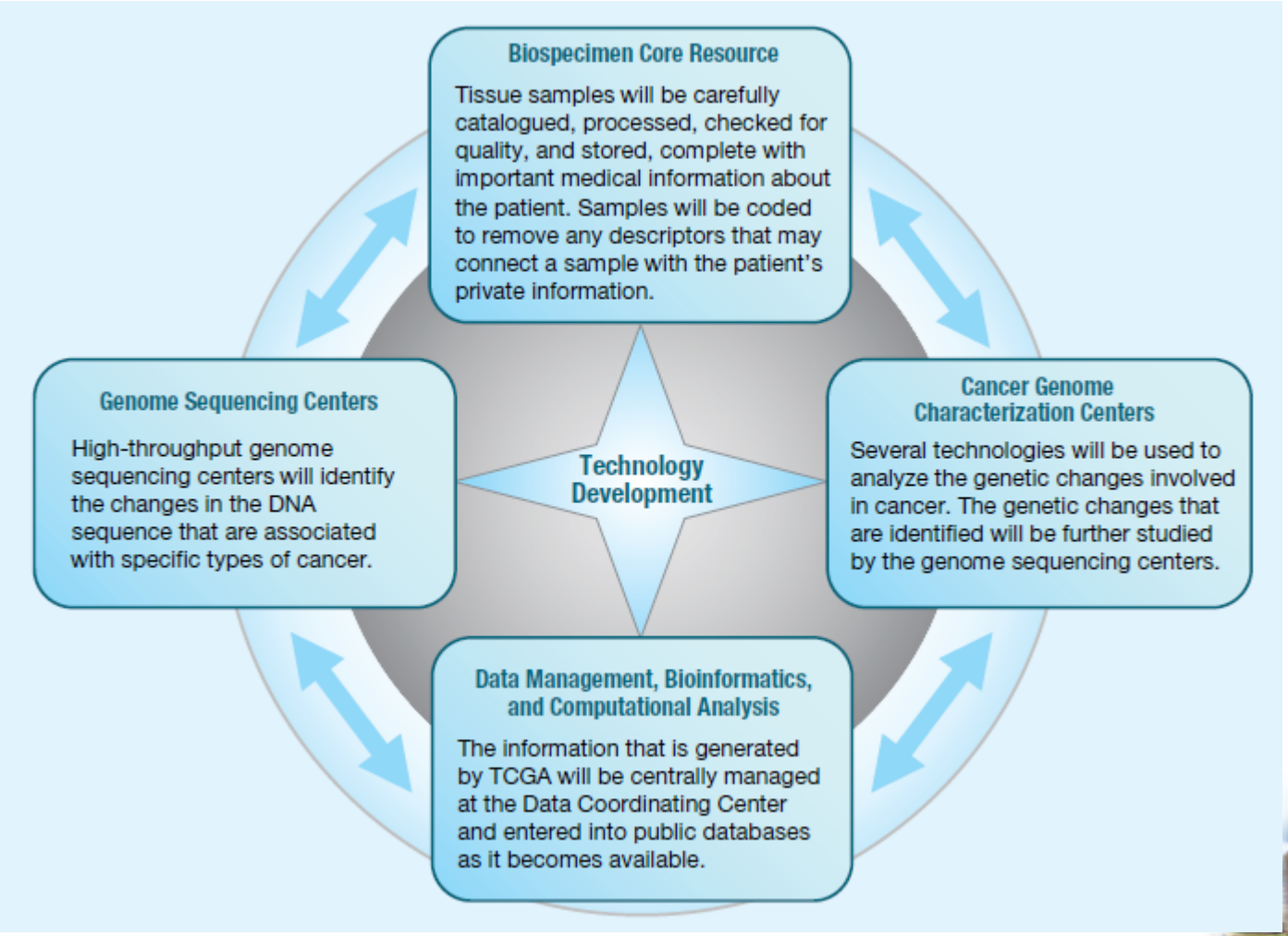
Create comprehensive public catalog of all genomic alterations present at significant frequency for all major cancer types



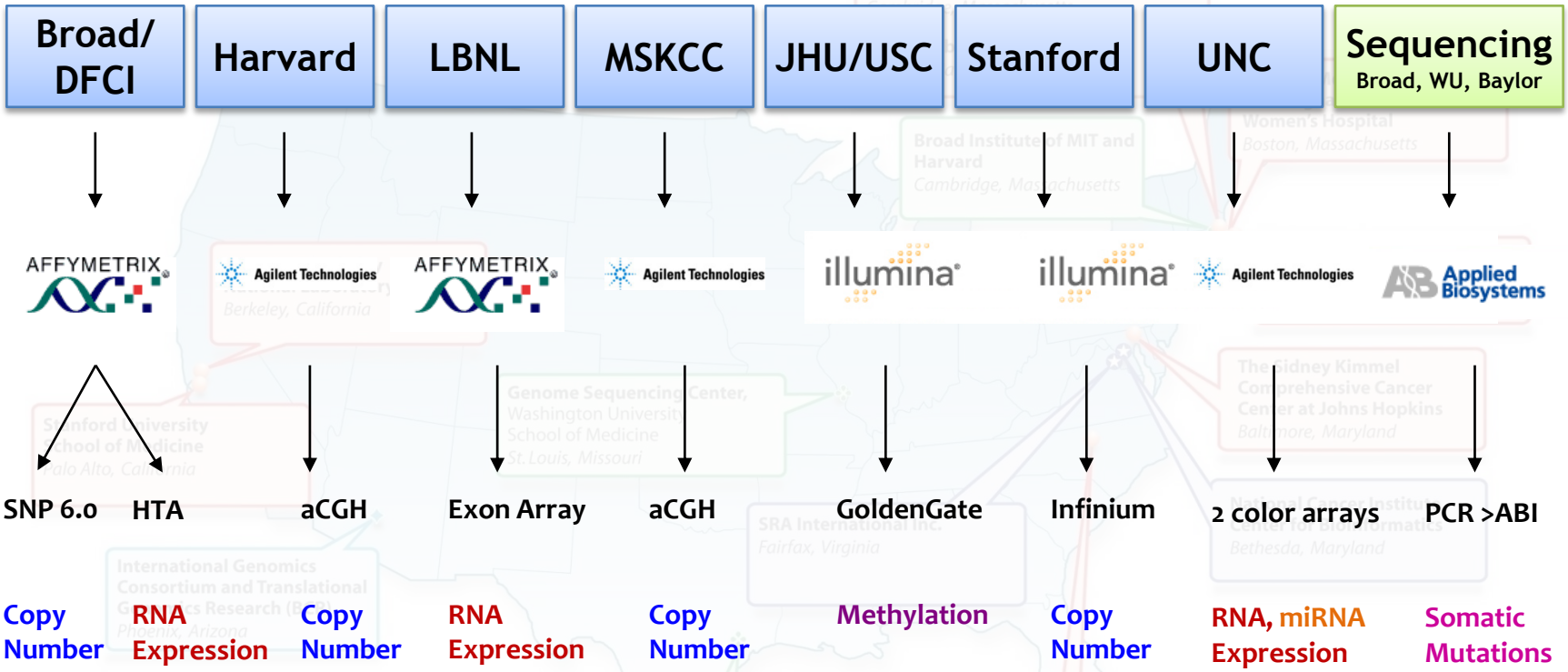
Comprehensive Cancer Genomic Research

1. TCGA researchers identify the genomic changes in more than 20 different types of human cancer
2. TCGA is analyzing hundreds of samples for each type of cancer
3. By comparing the DNA in samples of normal tissue and cancer tissue taken from the same patient, researchers can identify changes specific to that particular cancer
4. By connecting specific genomic changes with specific outcomes, researchers will be able to develop more effective, individualized ways of helping each cancer patient

TCGA process



TCGA : Center Overview & Components



Cross-Platform Data Integration, Comparison

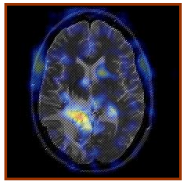
The Cancer Genome Atlas (TCGA):

Starting in 3 Cancers

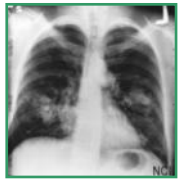
Connecting multiple sources, experiments, and data types

Three Cancers - TCGA

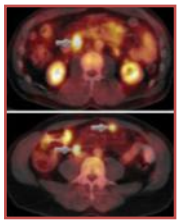
Glioblastoma Multiforme (Brain)



Squamous Carcinoma (Lung)



Serous Cystadenocarcinoma (Ovary)



Biospecimen Core Resource with more than 13 Tissue Source Sites

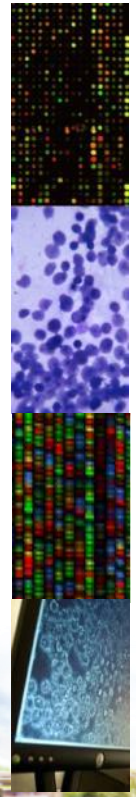
7 Cancer Genomic Characterization Centers

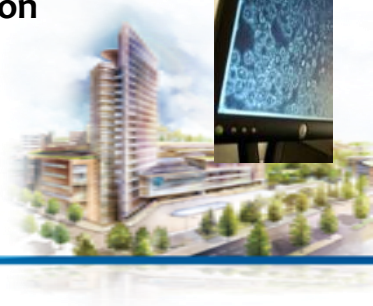
3 Genome Sequencing Centers

Data Coordinating Center

Multiple Data Types

- Clinical diagnosis
- Treatment history
- Histologic diagnosis
- Pathologic status
- Tissue anatomic site
- Surgical history
- Gene expression
- Chromosomal copy number
- Loss of heterozygosity
- Methylation patterns
- miRNA expression
- DNA sequence

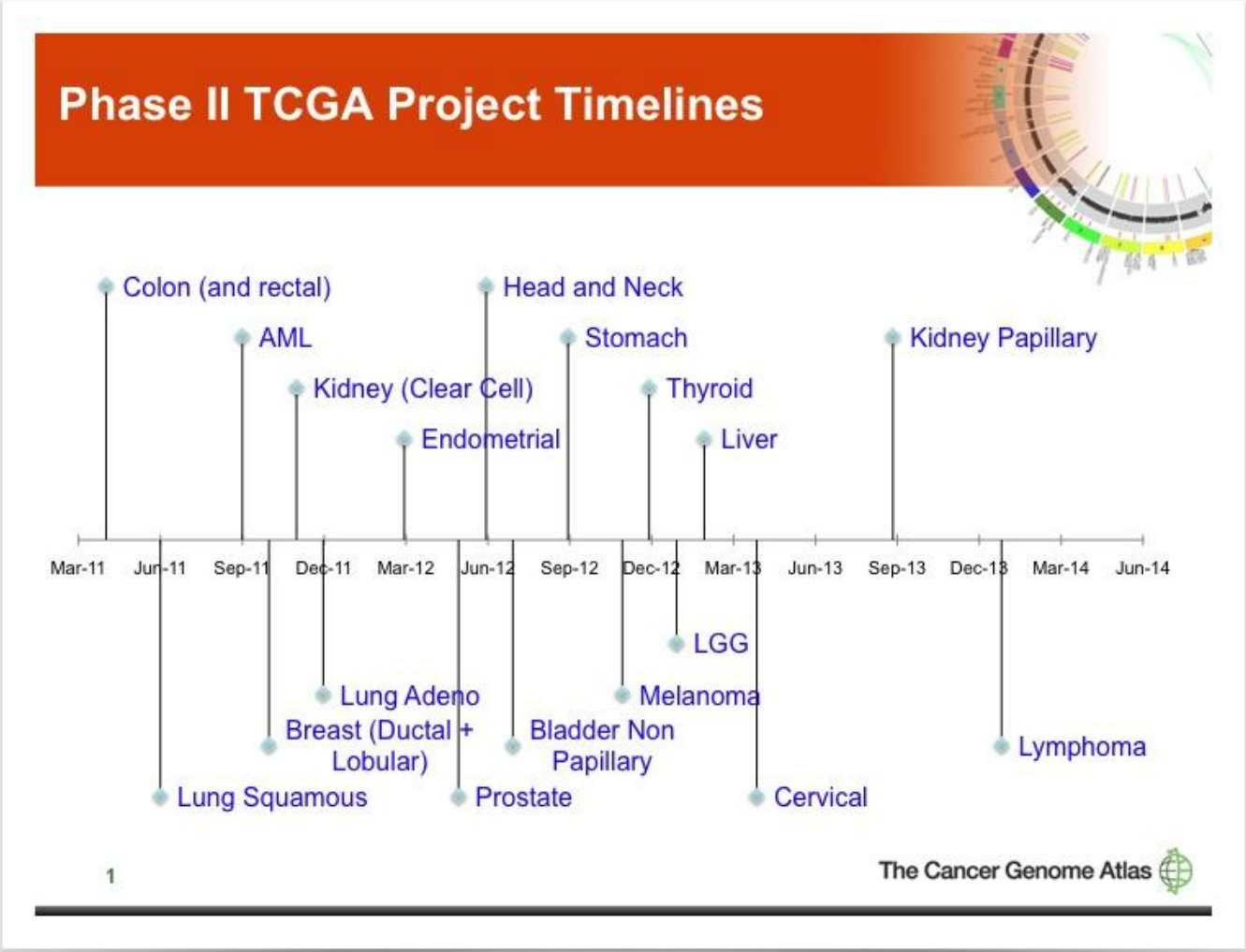




TCGA : Roadmap of Cancer Types & Data Status

TCGA Pilot study (~2011)

- GBM (Brain)
- HGSC (Ovary)
- SCC (Lung)



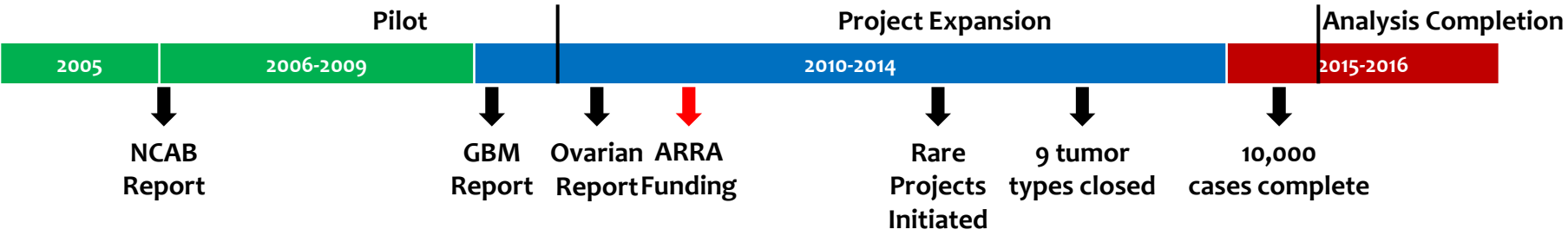
TCGA : Expanding the Enterprise

1. **Pilot project**:
   - FFPE-preserved tissues
   - Mouse models of Human cancers
2. **Projects to study rare tumor types**
   - Smaller numbers yet comprehensive focus of assays and analysis
3. **Integration efforts:**
   - International Cancer Genomics Consortium
   - Interface between TCGA (genomics of cancer samples) and **CPTAC (proteomics of cancer samples)**

TCGA : cancer selected for study

- The Cancer Genome Atlas (TCGA) has chosen cancers for study based on specific criteria that include:
- Poor prognosis and overall public health impact
- Availability of human tumor and matched-normal tissue samples
- **CANCER TISSUES BEING COLLECTED FOR POTENTIAL STUDY** Last Updated: May 16, 2014
- Gyn cancer : cervical, ovarian serous, uterine carcinosarcoma, uterine corpus endometrial ca.

TCGA : Timeline



TCGA : "No Platform Left Behind"

Multiple data types

Clinical diagnosis / Treatment history / Histologic diagnosis / Pathologic report & images / Tissue anatomic site / Surgical history / Gene expression & RNA sequence / Chromosomal copy number / Loss of heterozygosity / Methylation patters / miRNA expression / DNA sequence / RPPA(protein) / Subset for Mass Spec

- *Core Data Set* -

Synoptic path report	**Histology images**	**Required clinical data**	**Whole exome**
SNP 6.0 array	**mRNAseq**	**miRNAseq**	**Methylation array**

Rare tumor Project

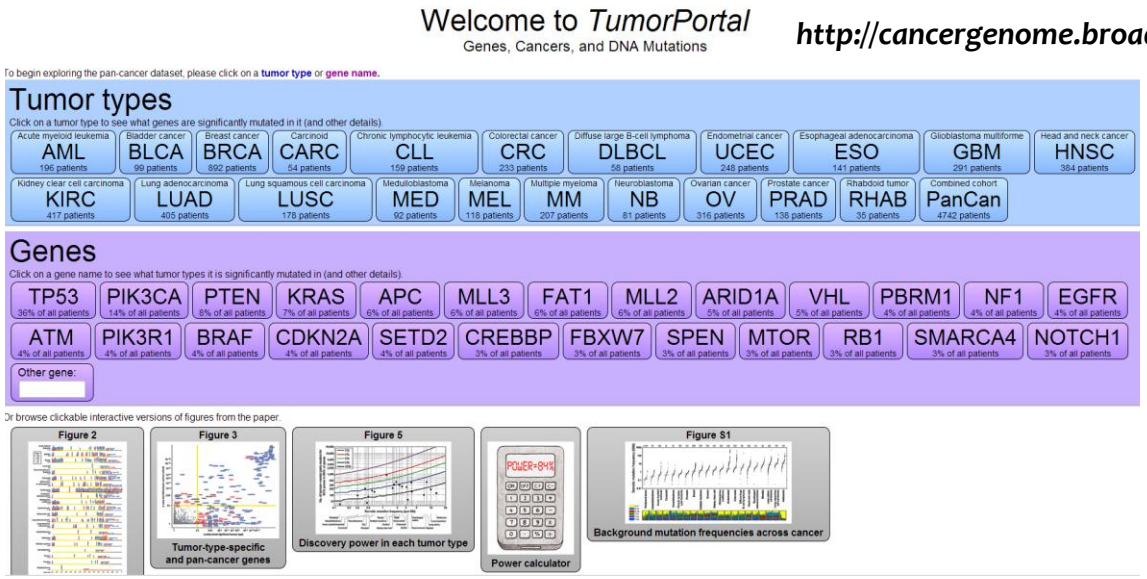
- Adrenocortical Carcinoma
- Chromophobe kidney
- Mesothelioma
- Paraganglioma / Pheochromocytoma
- Uterine Carcinosarcoma
- Thymoma
- Uveal Melanoma
- Testicular Germ Cell
- Cholangiocarcinoma
- Diffuse Large B cell Lymphoma

Ongoing Analysis Work Groups (AWGs)

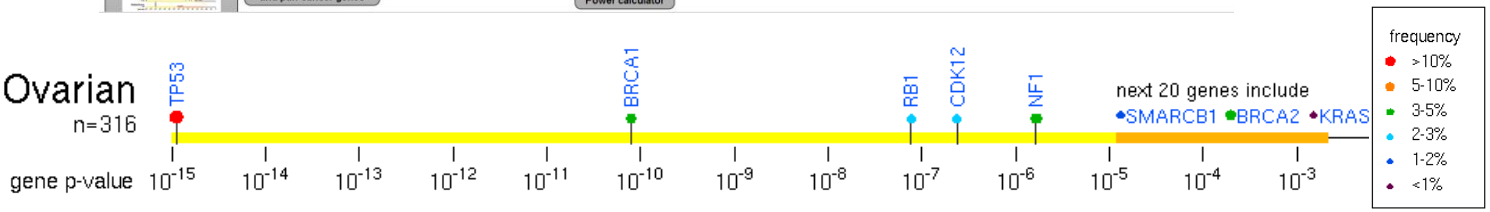
- Melanoma
- Prostate Adenocarcinoma
- Thyroid
- Low Grade Glioma
- Stomach & Esophageal
- Kidney Papillary Carcinoma
- Lung Adenocarcinoma
- Uterine Carcinosarcoma
- Cervical Carcinoma
- Pancreatic Adenocarcinoma
- Adrenocortical Carcinoma
- LGG + GBM
- Breast Lobular Carcinoma

2014 AACR Method Workshops, Jean C. Zenklusen

Finding cancer genes across ~5000 tumor/normal pairs from 21 tumor types



Example :



- **Cancer mutation datasets are complex, with heterogeneity at all levels of the data.**
- **In a typical tumor type, there are** : a few genes mutated at high frequency / many genes mutated at lower frequencies
- ***High-frequency genes*** : account for only a small fraction of all driver mutations have nearly all been discovered
- ***Lower-frequency genes*** : account for the vast majority of all driver mutations affect nearly all patients are still being discovered at a rapid pace

Integrated genomic analyses of ovarian carcinoma

*The Cancer Genome Atlas Research Network**

Nature 2011;474: 609-615

Background

- The Cancer Genome Atlas (TCGA)
  - Clinically annotated HGS-OvCa samples
  - Identify molecular abnormalities that influence pathophysiology, affect outcome and constitute therapeutic targets

- The analyses of 489 HGS-OvCa tumors
  - mRNA expression
  - microRNA (miRNA) expression
  - DNA copy number
  - DNA promoter methylation
  - Whole exome DNA sequence (316 samples)

Methods

- Sample inclusion criteria
  - Newly diagnosed, clinically annotated stage-II–IV HGS-OvCa
  - No prior treatment
  - Each frozen tumor specimen had to have a companion normal tissue specimen
  - >70% tumor cell nuclei and <20% necrosis
- Clinical data collection
  - Demographics
  - Histopathologic information
  - Treatment details
  - Outcome parameters

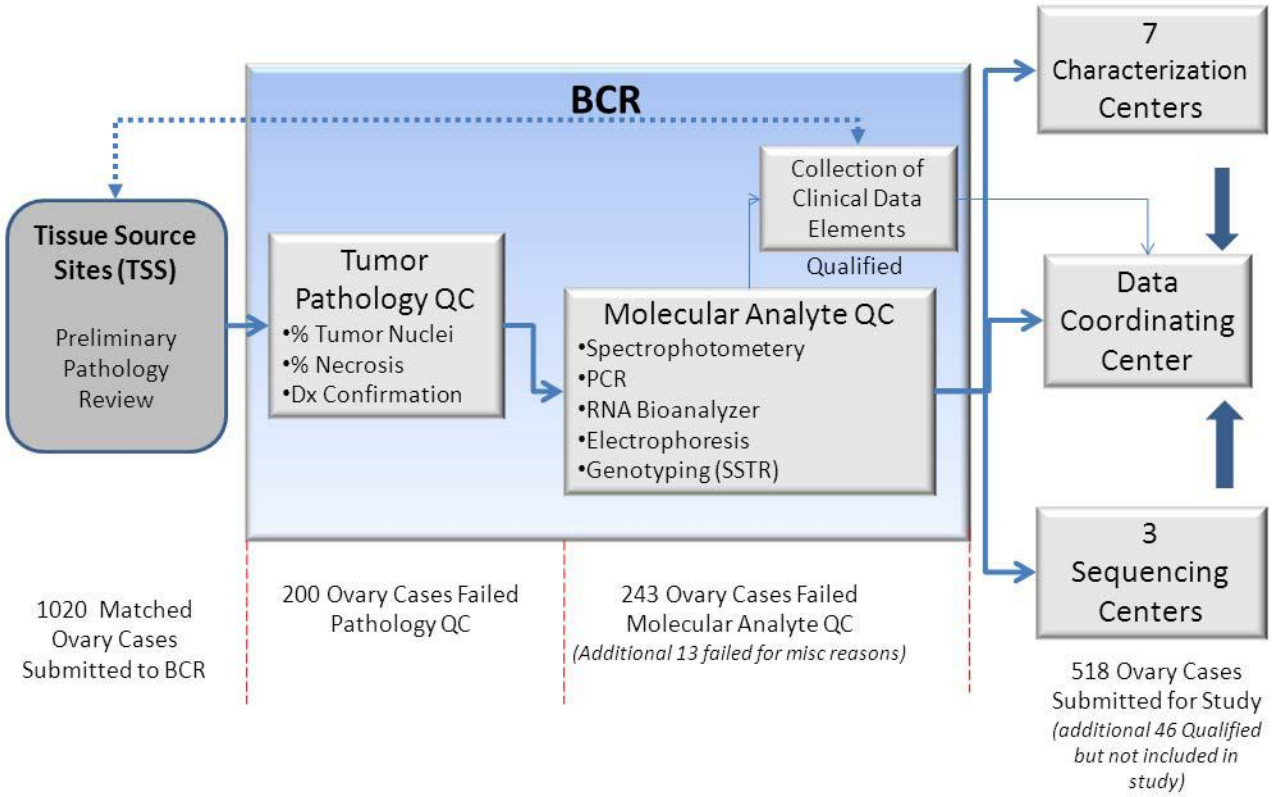
Clinical data analyses @

➤ Univariate analysis of overall and progression-free survival for TCGA ovarian cases

	Progression-free survival			Overall survival		
	HR	95% CI	*P*	HR	95% CI	*P*
Age (years)	1.00	0.99-1.01	0.99	**1.02**	**1.01-1.03**	**0.002**
Grade, 3 vs 2	1.33	0.95-1.86	0.10	1.35	0.94-1.94	0.11
Stage, III vs IV	0.88	0.75-1.04	0.13	0.87	0.74-1.01	0.07
TCGA cohort, training vs validation	1.05	0.94-1.19	0.38	0.99	0.88-1.12	0.91
Platinum status, resistant vs sensitive	**24.28**	**15.9-37.1**	**2.3e-49**	**3.94**	**2.86-5.43**	**6.0e-17**
Surgical outcome, optimal vs suboptimal	0.87	0.66-1.15	0.34	0.77	0.59-1.02	0.06

HR, hazard ratio

Biospecimen processing and quality control



To date, 1020 ovarian cases have been received by the BCR and 564 (55%) have passed quality control. The biospecimens included in this report come from 518 ovarian samples.

The Cancer Genome Atlas Research Network, Nature 2011

Platforms used and data produced

- Coordinated molecular analyses using multiple molecular assays at independent sites were carried out.

➤ **Characterization platforms used and data produced**

Data type	Platforms	Cases	Data access
DNA sequence of exome	Illumina GAIIx*†	236	Controlled
	ABI SOLiD‡	80	Controlled
Mutations present in exome		316	Open
DNA copy number/genotype	Agilent 244K§‖	97	Open
	Agilent 415K§	304	Open
	Agilent 1M‖	539	Open
	Illumina 1MDUO¶	535	Controlled
	Affymetrix SNP6*	514	Controlled
mRNA expression profiling	Affymetrix U133A*	516	Open
	Affymetrix Exon#	517	Controlled
	Agilent 244K**	540	Open
Integrated mRNA expression		489	Open
miRNA expression profiling	Agilent**	541	Open
CpG DNA methylation	Illumina 27K††	519	Open
Integrative analysis		489	Open
Integrative analysis with mutations		309	Open

The data set analysed here is available at the TCGA website (http://tcga-data.nci.nih.gov/docs/publications/ov_2011)

The Cancer Genome Atlas Research Network, Nature 2011

1. Mutation Analysis

Nine genes were identified as significant mutations

- **TP53** was mutated in 303 of 316 samples (**95.9%**)
- **BRCA1 and BRCA2** had germline mutations in **9% and 8%** of cases, respectively, and showed somatic mutations in a further 3% of cases
- Six other statistically recurrently mutated genes were identified (**RB1, NF1, FAT3, CSMD3, GABRA6 and CDK12**)

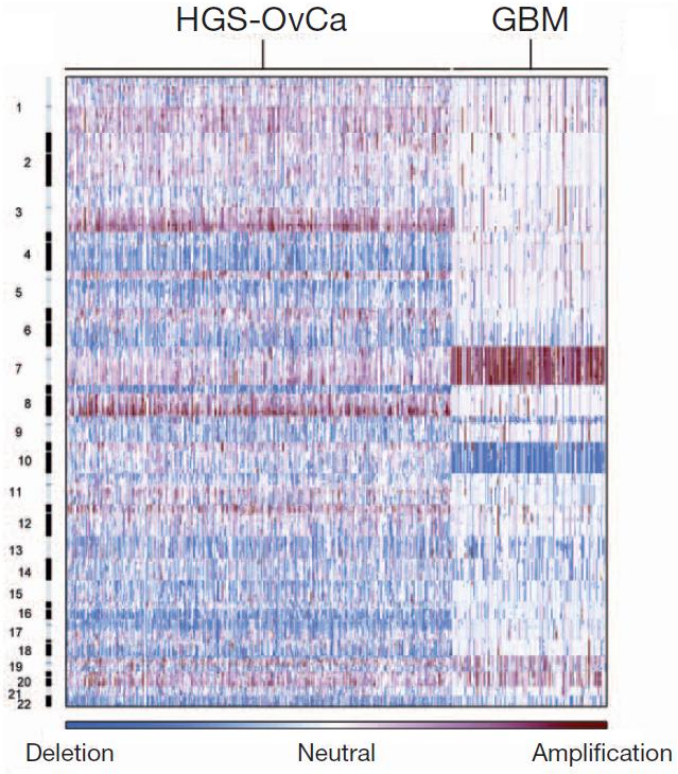
➤ Significantly mutated genes in HGS-OvCa

Gene	No. of mutations	No. validated	No. unvalidated
TP53	302	294	8
BRCA1	11	10	1
CSMD3	19	19	0
NF1	13	13	0
CDK12	9	9	0
FAT3	19	18	1
GABRA6	6	6	0
BRCA2	10	10	0
RB1	6	6	0

The Cancer Genome Atlas Research Network, Nature 2011

2. Copy Number Analysis

➤ **Genome copy number abnormalities**



Somatic copy number alterations (SCNAs) present in the 489 HGS-OvCa genomes were identified

Copy number profiles of 489 HGS-OvCa, compared with profiles of 197 glioblastoma multiforme (GBM) tumours.

Copy number increases (red) and decreases (blue) are plotted as a function of distance along the normal genome (vertical axis, divided into chromosomes).

The Cancer Genome Atlas Research Network, Nature 2011

Copy Number Analysis: Recurrent Focal SCNAs

63 regions of focal amplification were identified

- The most common focal amplifications encoded **CCNE1**, **MYC** and **MECOM**, each of which was **highly amplified in more than 20% of tumors**
- New tightly localized amplification peaks in HGS-OvCa encoded
  - **ZMYND8**(receptor for activated C-kinase)
  - **IRF2BP2**(p53 target gene)
  - **ID4**(DNA-binding protein inhibitor)
  - **PAX8**(embryonic development gene)
  - **TERT**(telomerase catalytic subunit)

50 regions of focal deletions were identified

- The known tumor suppressor genes **PTEN**, **RB1** and **NF1** were in regions of homozygous **deletions in at least 2% of the tumors**

The Cancer Genome Atlas Research Network, Nature 2011

Copy Number Analysis: Possible Therapeutic Targets

Ingenuity System, ClinicalTrials.gov & DrugBank were used to identify possible therapeutic inhibitors of amplified, overexpressed genes were identified.

- ***22 genes that are therapeutic targets***, including **MECOM**, **MAPK1**, **CCNE1** and **KRAS**, are amplified in at least 10% of the cases

Gene	Description
AKT1	v-akt murine thymoma viral oncogene homolog 1
AKT3	v-akt murine thymoma viral oncogene homolog 2
CCNE1	cyclin E1
CDK2	cyclin-dependent kinase 2
EPCAM	epithelial cell adhesion molecule
ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2 (avian)
ERBB3	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)
HSP90AB1	heat shock protein 90kDa alpha (cytosolic), class B member 1
IGF1R	insulin-like growth factor 1 receptor
KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
MAPK15	mitogen-activated protein kinase 15
MECOM	MDS1 and EVI1 complex locus
MSTN	myostatin
NOS3	nitric oxide synthase 3 (endothelial cell)
POLB	polymerase (DNA directed), beta
RHEB	Ras homolog enriched in brain
RICTOR	RPTOR independent companion of MTOR, complex 2
RPTOR	regulatory associated protein of MTOR, complex 1
STAT1	signal transducer and activator of transcription 1, 91kDa
STAT4	signal transducer and activator of transcription 4
TERT	telomerase reverse transcriptase
VEGFA	VEGFA

The Cancer Genome Atlas Research Network, Nature 2011

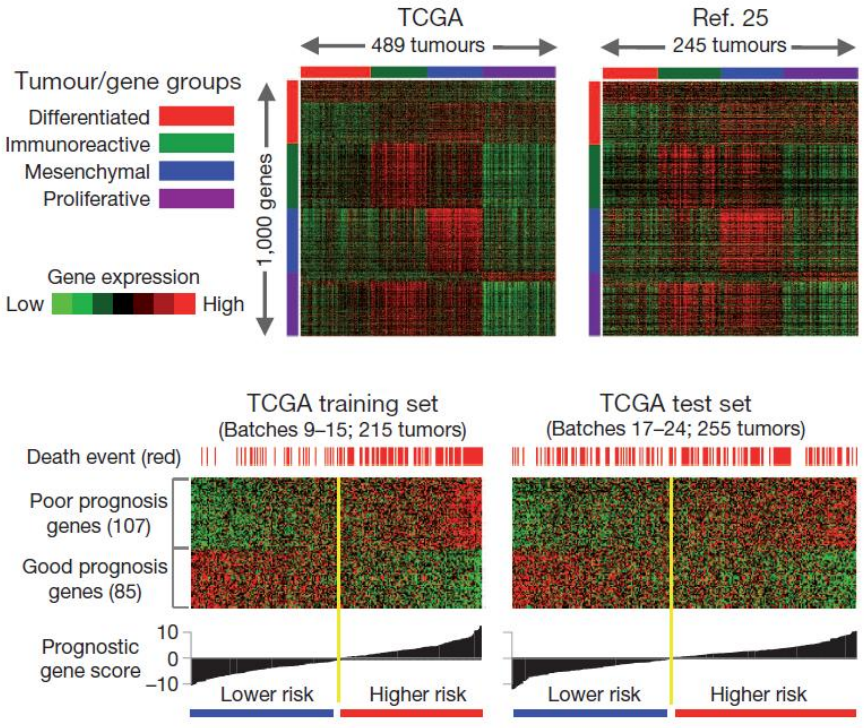
3. mRNA and miRNA Expression Analysis

At least **four robust expression subtypes** exist in HGS-OvCa

- ***Immunoreactive***
- ***Differentiated***
- ***Proliferative***
- ***Mesenchymal***

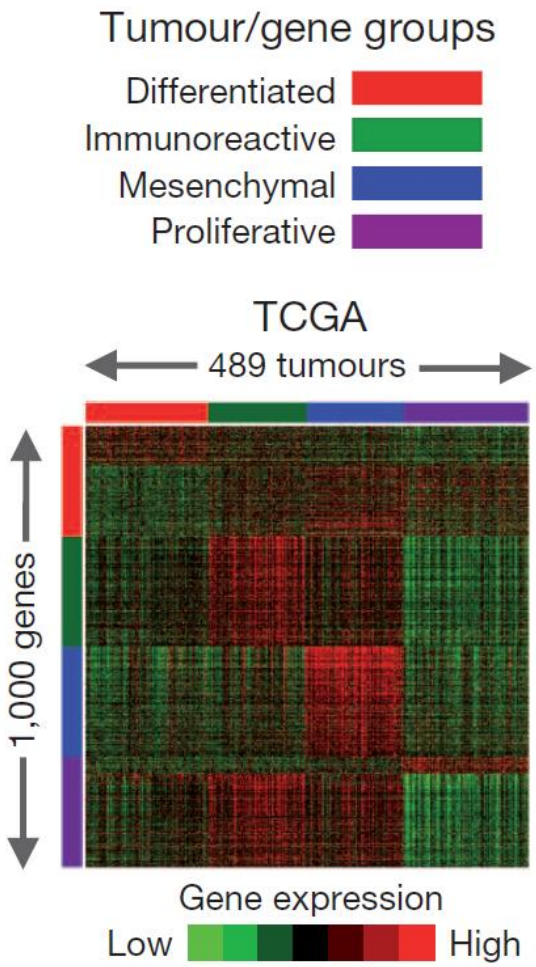
A **193-gene transcriptional signature predictive** of overall survival was defined using the integrated expression data set

➤ **Gene and miRNA expression patterns of molecular subtype and outcome prediction**



The Cancer Genome Atlas Research Network, Nature 2011

mRNA (Gene) Expression Analysis:
Four Clusters @



1. ***Immunoreactive*** *subtype:* T-cell chemokine ligands **CXCL11** and **CXCL10** and the receptor **CXCR3** characterized the immunoreactive subtype

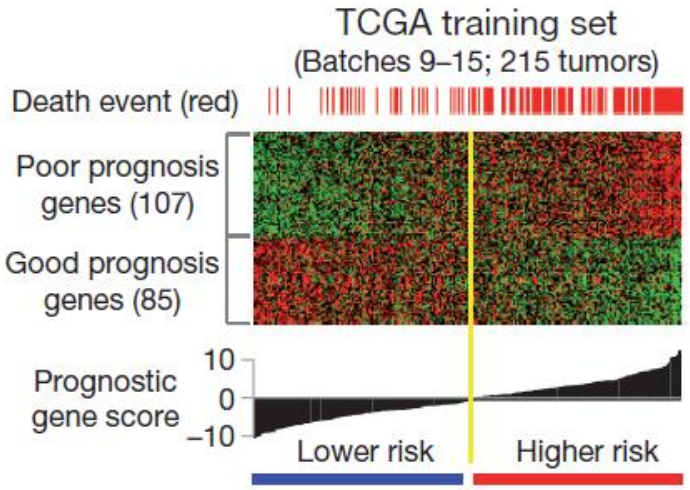
2. ***Proliferative*** *subtype:* High expression of transcription factors such as **HMGA2** and **SOX11,** low expression of ovarian tumour markers (**MUC1** and **MUC16**) and high expression of proliferation markers such as **MCM2** and **PCNA** defined the proliferative subtype

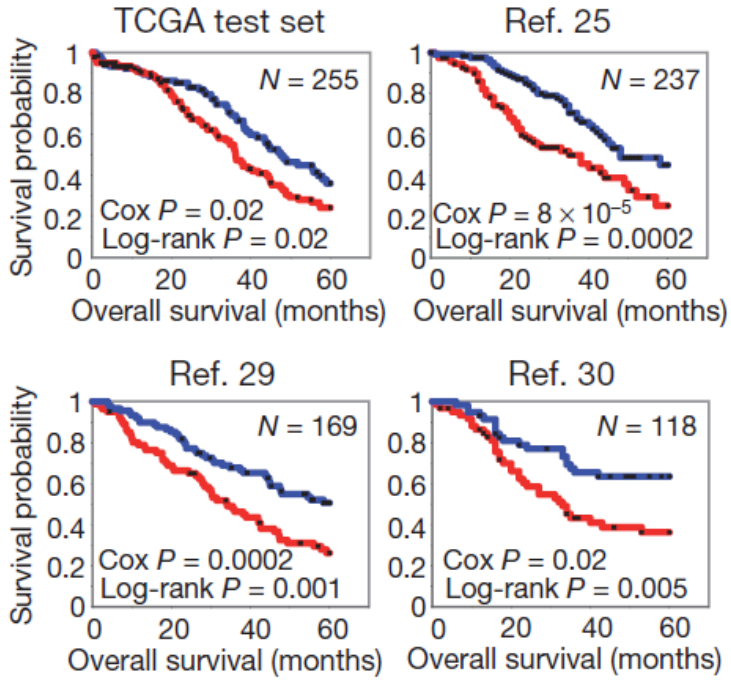
3. ***Differentiated*** *subtype:* The differentiated subtype was associated with high expression of **MUC16** and **MUC1** and with expression of the secretory fallopian tube maker **SLPI**

4. ***Mesenchymal*** *subtype:* High expression of **HOX** genes and markers suggestive of increased stromal components such as for myofibroblasts (**FAP**) and microvascular pericytes (**ANGPTL2** and **ANGPTL1**) characterized the mesenchymal subtype

The Cancer Genome Atlas Research Network, Nature 2011

mRNA Expression Analysis: Correlation with Survival @@





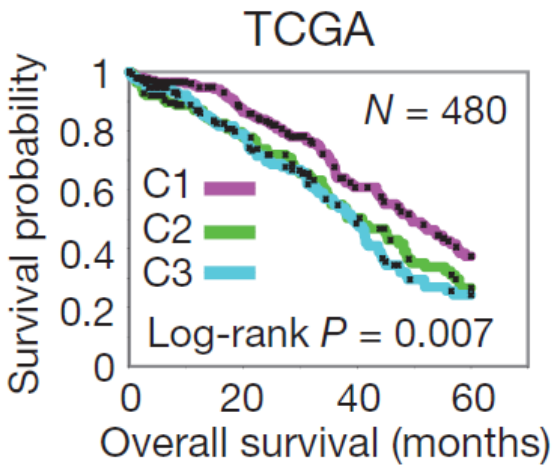
Using a training data set, a prognostic gene signature was defined and applied to a test data set

Kaplan–Meier analysis of four independent expression profile data sets, comparing survival for predicted higher-risk patients versus lower-risk patients.

The Cancer Genome Atlas Research Network, Nature 2011

miRNA Expression Analysis: Overlapping Cluster

miRNA cluster	Gene cluster D	I	M	P
C1	55	48	15	89
C2	40	21	51	29
C3	39	37	43	20



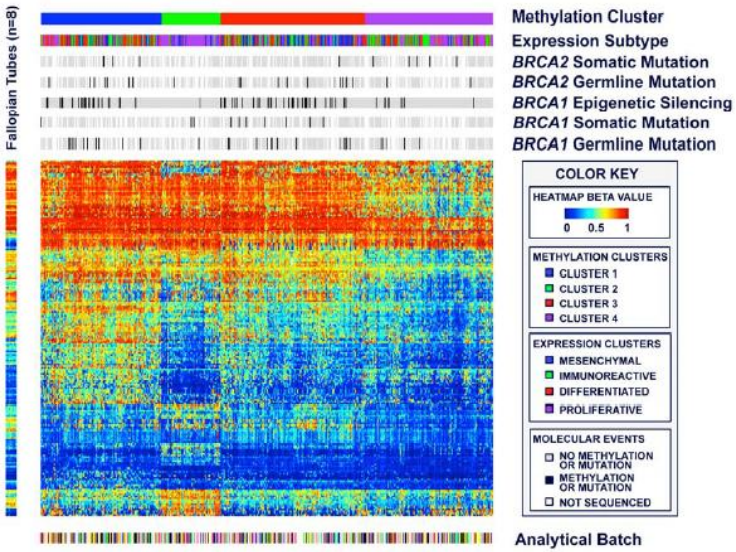
Tumours separated into three **clusters on the basis of miRNA** expression, **overlapping with gene-based clusters** as indicated.

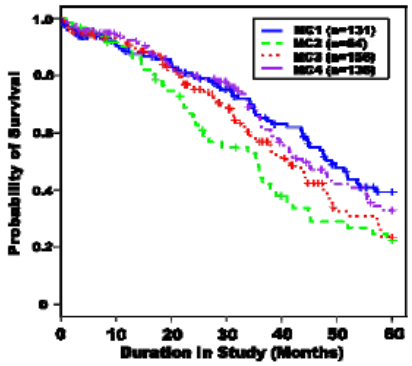
- D, differentiated
- I, immunoreactive
- M, mesenchymal
- P, proliferative

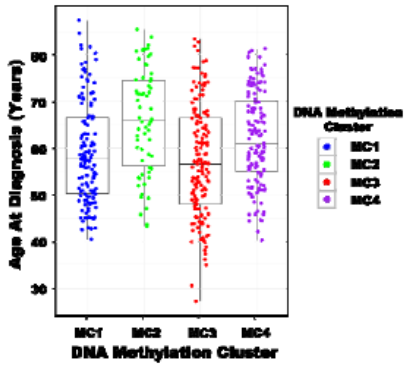
Survival duration differed significantly between **miRNA subtypes:** patients with miRNA subtype-1 tumors survived significantly longer

The Cancer Genome Atlas Research Network, Nature 2011

4. DNA Methylation Analysis







DNA methylation was correlated with reduced gene expression

- **AMT, CCL21** and **SPARCL1** were noteworthy because they showed promoter hypermethylation in the vast majority of the tumors

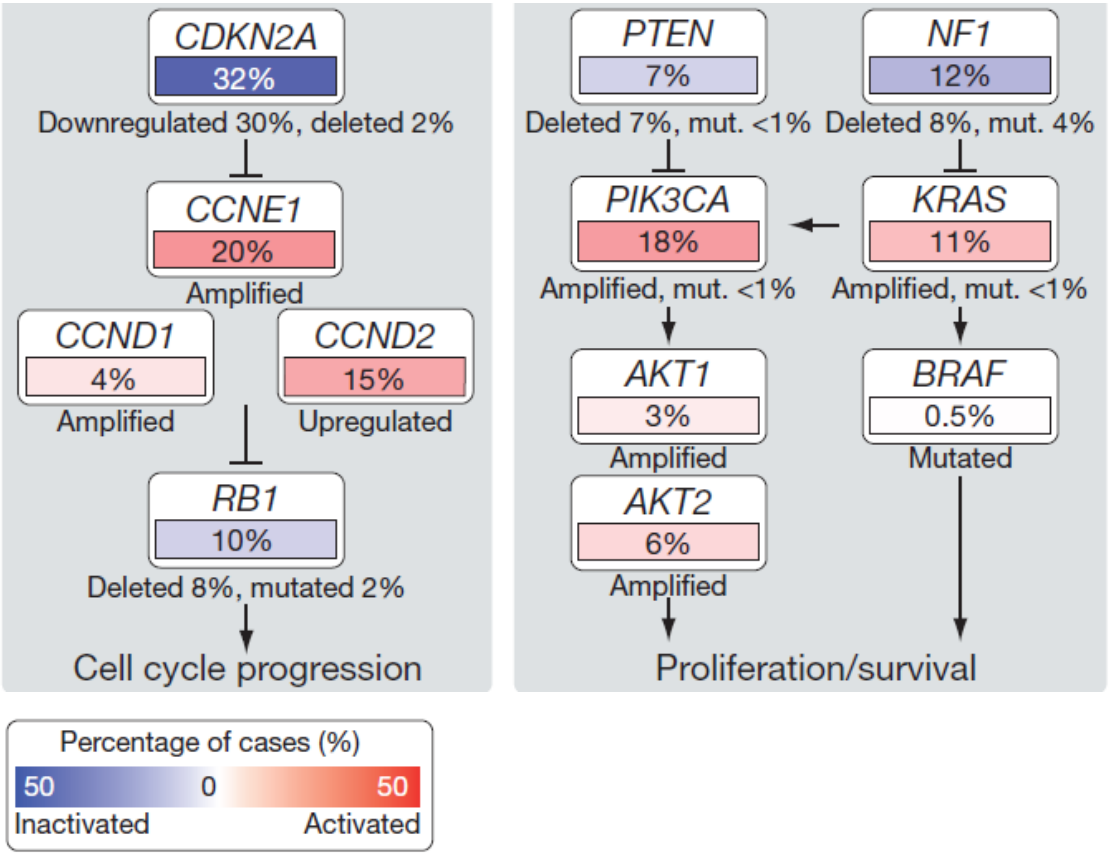
Consensus **clustering of variable DNA methylation** across tumors identified **four subtypes** that were significantly associated with differences in **age, BRCA inactivation** events and **survival**

The Cancer Genome Atlas Research Network, Nature 2011

5. Pathways Analysis: RB1 and PI3K/RAS pathway

➤ **RB signaling:**
67% of cases altered

➤ **PI3K/RAS signaling:**
45% of cases altered

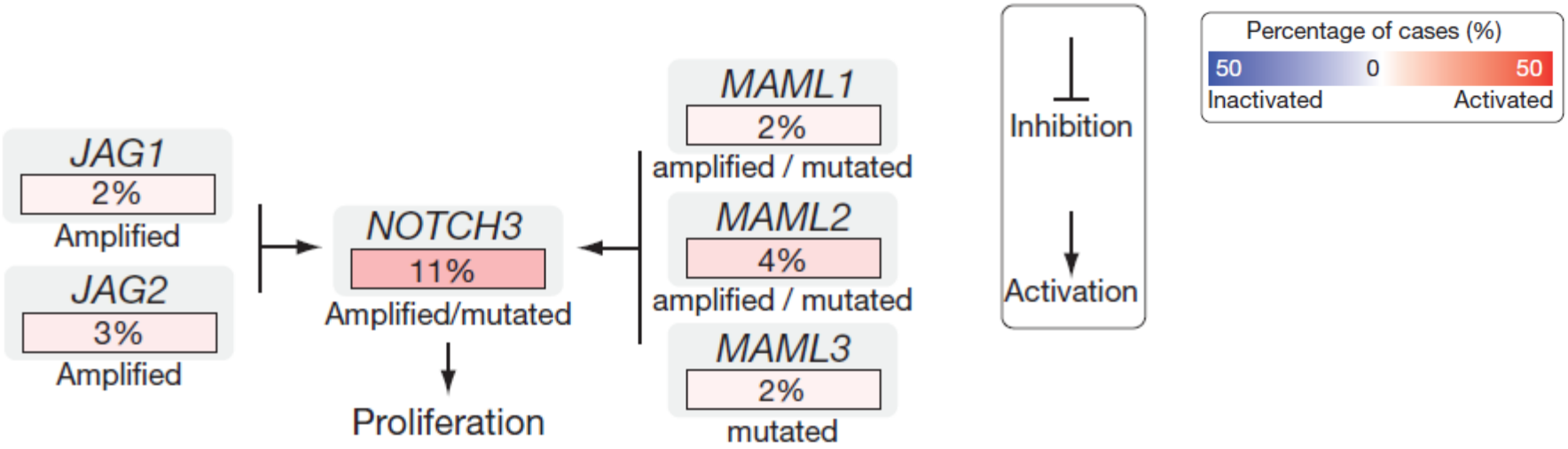


Analysis of the frequency with which known cancer-associated pathways harbored mutations, copy number changes or changes in gene expression showed that the **RB1 and PI3K/RAS pathways** were **deregulated in 67% and 45% of cases,** respectively

The Cancer Genome Atlas Research Network, Nature 2011

Pathways Analysis: NOTCH pathway

➤ **NOTCH signaling:** *22% of cases altered*

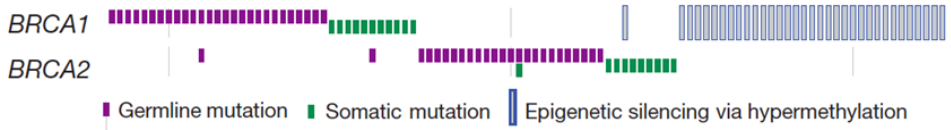


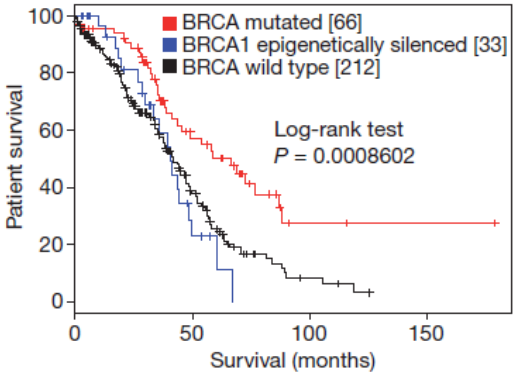
Other analytic method identified several known pathways including the **NOTCH signalling pathway,** which was **altered in 22%** of HGS-OvCa samples

The Cancer Genome Atlas Research Network, Nature 2011

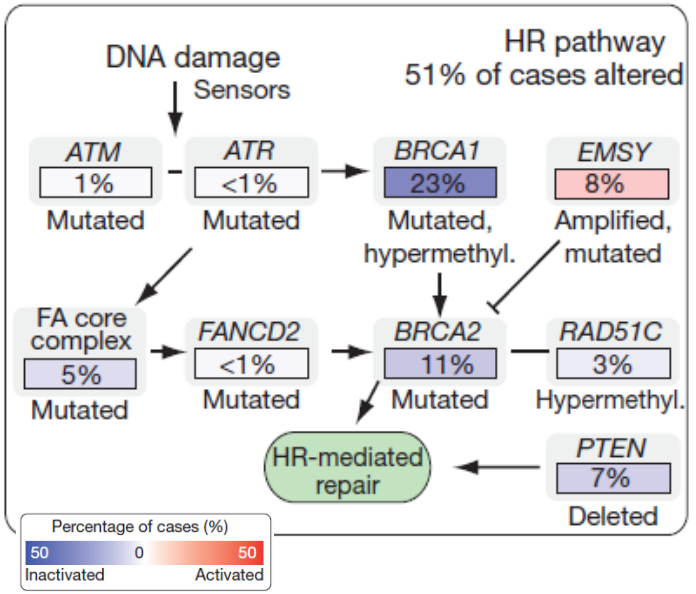
Pathways Analysis: HR pathway

➤ **BRCA1/2**: *20% - mutation, 11% - epigenetic silencing*





➤ **HR pathway**: *51% of cases altered*



BRCA altered cases were 31%

- 20% : germline or somatic mutations in BRCA1/2
- 11% :DNA hypermethylation - epigenetic silencing of BRCA1

Survival analysis showed **better overall survival for BRCA1/2 mutated cases** than BRCA1/2 wild-type cases

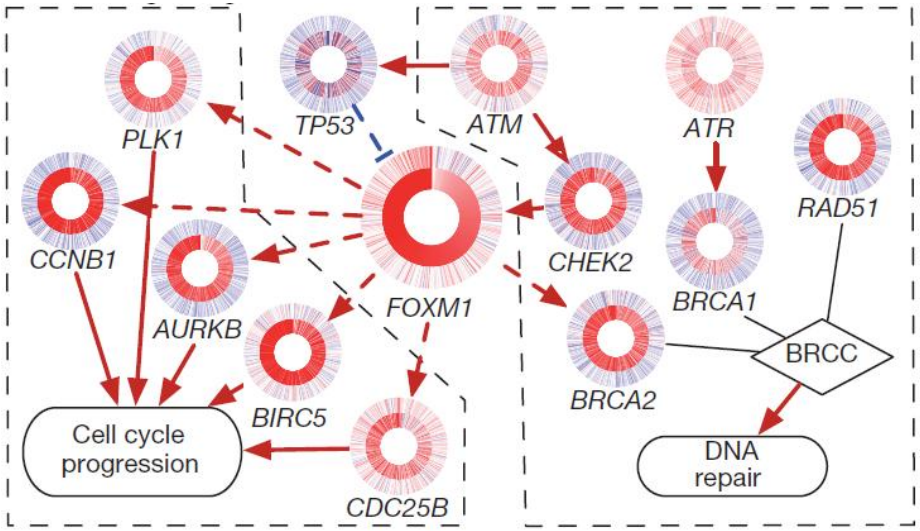
- Notably, epigenetically silenced BRCA1 cases had survival similar to BRCA1/2 wild-type

Overall, homologous recombination defects may be present in approximately half of all HGS-OvCa cases, providing a rationale for clinical trials of PARP inhibitors targeting tumours with these homologous-recombination-related aberrations

The Cancer Genome Atlas Research Network, Nature 2011

Pathways Analysis: FOXM1 pathway

➤ **FOXM1 signaling:** *84% of cases altered*



Each gene is depicted as a multi-ring circle in which its copy number (outer ring) and gene expression (inner ring) are plotted such that each 'spoke' in the ring represents a single patient sample, with samples sorted in increasing order of FOXM1 expression.

The FOXM1 transcription factor network is significantly altered in 87% of cases

- FOXM1 and its proliferation-related target genes, AurB (AURKB), CCNB1, BIRC5, CDC25 and PLK1, were consistently overexpressed but not altered by DNA copy number changes, indicative of transcriptional regulation
- TP53 represses FOXM1 afterDNAdamage, suggesting that the high rate of TP53 mutation in HGS-OvCa contributes to FOXM1 overexpression

The Cancer Genome Atlas Research Network, Nature 2011

Integrated Genomic Analysis of EOC

Integrated genomic analyses of ovarian carcinoma

The Cancer Genome Atlas Research Network*

- **TP53 mutations in almost all tumors (96%)**
- Low prevalence but statistically recurrent **somatic mutations : NF1, BRCA1, BRCA2, RB1, CDK12, CSMD3, FAT3, GABRA6**

- 113 significant **Focal Somatic Copy Number Aberrations**
- **Possible therapeutic targets** of amplified, overexpressed genes (**22 genes**)

- **4 subtypes** based on **mRNA expression**
  - **Immunoreactive** • **Differentiated** • **Proliferative** • **Mesenchymal**
- 3 **microRNA** subtypes, 4 **promoter methylation** subtypes: overlapping with gene-based clusters

➤ **Associated with Survival Duration**

- Analyses of **Pathway**
  - **RB signaling: 67%** of cases altered
  - **NOTCH signaling: 22%** of cases altered
  - **FOXM1 signaling: 84%** of cases altered
  - **PI3K/RAS signaling: 45%** of cases altered
  - **HR pathway: 51%** of cases altered

The Cancer Genome Atlas Research Network, Nature 2011

Genomic analysis of TCGA in Ovarian CA

ARTICLE

doi:10.1038/nature10166

Integrated genomic analyses of ovarian carcinoma

The Cancer Genome Atlas Research Network*

Nature 2011;474: 609-615

- 샘플: 489명의 악성 **난소암** 환자의 종양샘플
- 분석: 엑손 염기서열, DNA 카피수, mRNA 및 miRNA 발현, 프로모터 메틸화

TP53 mutations in almost all tumors (96%)

NF1, BRCA1, BRCA2, RB1, CDK12
Low prevalence but statistically recurrent mutations

*HGSC are characterized by **few driver point mutations***

113 significant Focal Somatic Copy Number Aberrations

amplified, overexpressed genes (22 genes)

Possible therapeutic targets

4 Molecular Subtype based on mRNA expression

- *Immunoreactive subtype*
- *Differentiated subtype*
- *Proliferative subtype*
- *Mesenchymal subtype*

New Subtype of OVCA

Alteration of Pathway

- RB signaling: 67%
- PI3K/RAS signaling: 45%
- NOTCH signaling: 22%
- HR pathway: 51%
- FOXM1 signaling: 84%

ARTICLE

doi:10.1038/nature10166

Integrated genomic analyses of ovarian carcinoma

The Cancer Genome Atlas Research Network*

Nature 2011;474: 609-615

- In molecular characteristics of HGSC, epithelial ovarian cancer is increasingly understood to be a molecularly heterogenous malignancy with

1) no obvious "targetable" driver mutations

2) a myriad of tumor suppressor gene loss or dysfunction &

3) widespread DNA copy number aberrations .

❖ Although molecularly targeted agents are unlikely to replace current therapy soon, several of these agents are showing promise efficacy

- Bevacizumab, an **inhibitor antibody directed toward VEGF**, is showing promise in combination with existing cytotoxics
- **The central importance of HRR (homologous recombination repair) deficiency in HGSCs** renders these tumors sensitive to inhibition of other DNA repair proteins, such as **PARP inhibitors**

Gene-gene interaction network analysis of ovarian cancer using TCGA data

- explore the molecular mechanism of ovarian cancer pathogenesis using TCGA data
- identify the differentially expressed genes (DEGs) between ovarian cancer and normal samples, followed by the function and pathway annotations of the DEGs.
- NetBox software was used to for the gene-gene interaction (GGI) network construction and the corresponding modules identification, and functions of genes in the modules

Ying et al. Journal of Ovarian Research 2013, 6:88

Gene-gene interaction network analysis of ovarian cancer using TCGA data

- identified 332 DEGs, including 146 up-regulated genes which mainly involved in the cell cycle related functions and cell cycle pathway
- 186 down-regulated genes which were enriched in extracellular region par function, and Ether lipid metabolism pathway
- provides a comprehensive bioinformatics analysis of genes, functions, and pathways

Ying et al. Journal of Ovarian Research 2013, 6:88

Analysis and Comparison of Somatic Mutations in Paired Primary and Recurrent Epithelial Ovarian Cancer Samples

- investigate the pattern of somatic point mutations in matched paired samples of primary and recurrent epithelial ovarian cancers using the OncoMap mutation detection protocol
- set of 92 formalin-fixed, paraffin-embedded (FFPE) tumors, consisting of matched paired samples of initially diagnosed and recurrent tumors from 46 epithelial ovarian cancer (EOC) patients
- somatic mutations were found in *CDKN2A*, *KRAS*, *MLH1*, and *TP53*. No differences in mutational status between primary and recurrent samples

Kim et al. PLOS one 2014,

Assessing the clinical utility of cancer genomic and proteomic data across tumor types

- retrospectively predict patient survival using diverse molecular data (somatic copy-number alteration, DNA methylation and mRNA, microRNA and protein expression) from 953 samples of four cancer types from TCGA project
- incorporating molecular data with clinical variables yields statistically significantly improved predictions (FDR < 0.05) for three cancers but those quantitative gains were limited (2.2-23.9%).
- provides a starting point and resources, including an open-access model evaluation platform

Yuan et al. Nature Bioetech2 014,

Integrated genomic characterization of endemetrial carcinoma

*The Cancer Genome Atlas Research Network**

Nature 2013;497(7447):67-73

ARTICLE

OPEN

doi:10.1038/nature12113

Integrated genomic characterization of endometrial carcinoma

The Cancer Genome Atlas Research Network*

Nature 2013 May 2;497(7447):67-73

373 patients with endometrial cancer
- 307 endometrioid
- 53 serous
- 13 mixed histology

- Somatic copy number alterations
- Exome sequence analysis

✓ Multiplatform subtype classifications
✓ Structural aberrations
✓ Pathway alterations

Sample Characteristics

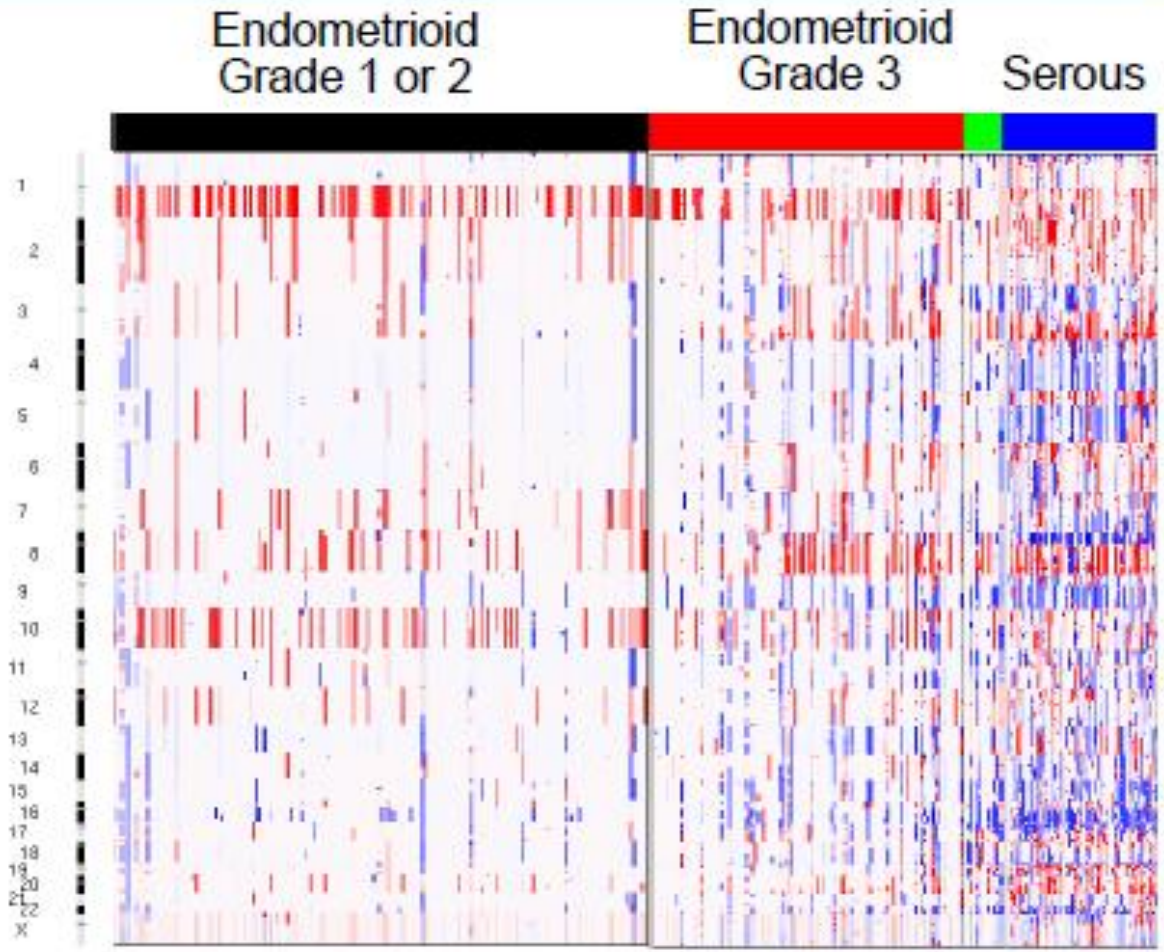
Cohort	Total
Number of patients	373
Age	
Mean, years (STD)	63 (11)
Range	31-90
Recurrent Disease	
Yes	72 (19.3%)
No	279 (74.8%)
Unknown	22 (5.9%)
Vital Status	
Alive	332 (89%)
Dead	39 (10.5%)
Unknown	2 (0.5%)

Stage	EndoGr1	EndoGr2	EndoGr3	MixedGr3	SerousGr3	Total
Stage I	78 (89%)	83 (79%)	70 (63%)	6 (46%)	17 (32%)	254 (69%)
Stage II	3 (3%)	9 (9%)	6 (5%)	2 (15%)	5 (9%)	25 (7%)
Stage III	7 (8%)	12 (11%)	26 (23%)	4 (31%)	25 (47%)	74 (20%)
Stage IV	(0%)	1 (1%)	9 (8%)	1 (8%)	6 (11%)	17 (5%)
Total	88 (100%)	105 (100%)	111 (100%)	13 (100%)	53 (100%)	370 (100%)

The Cancer Genome Atlas

Somatic Copy Number Alterations

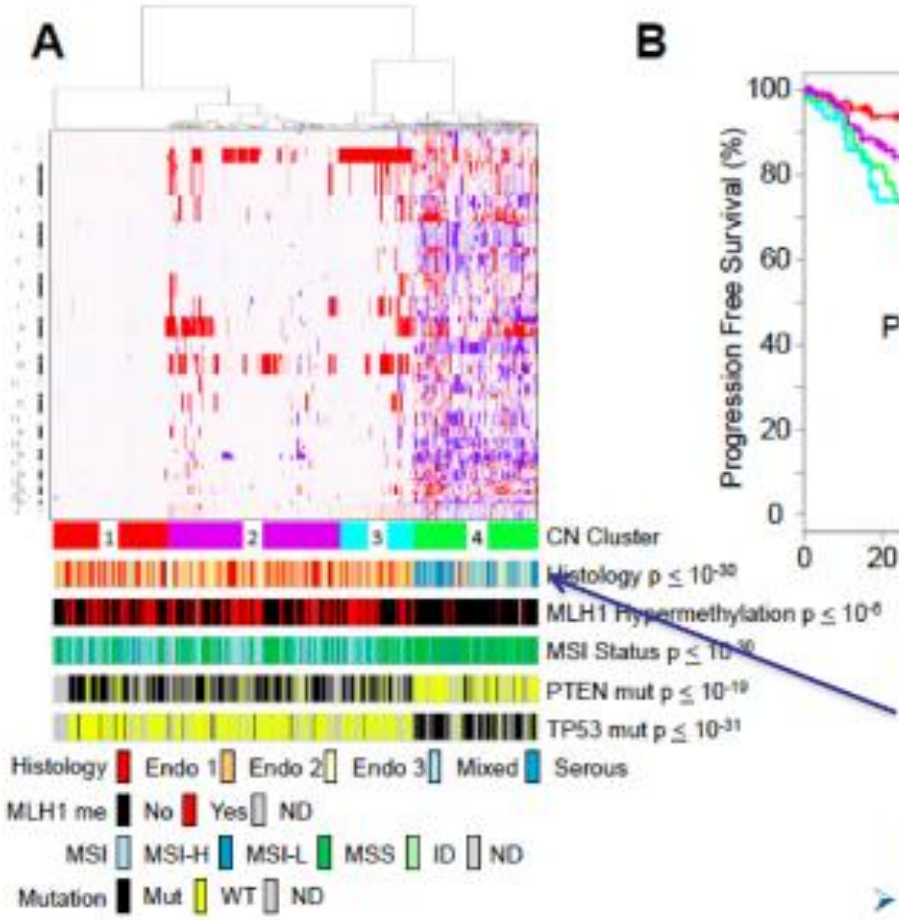
More genomic instability as tumors become less differentiated

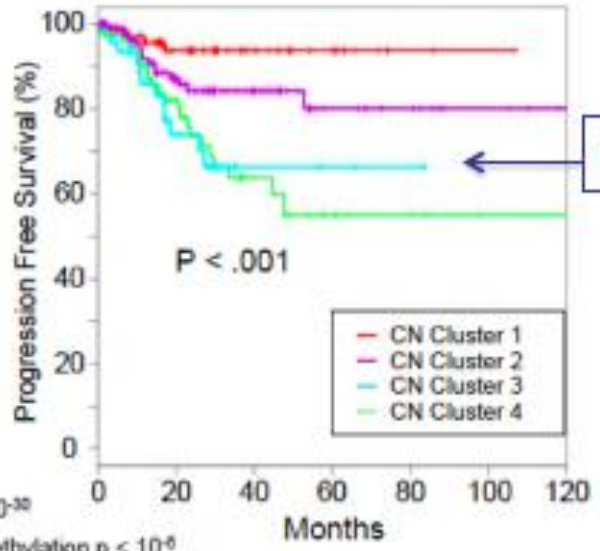


➢ Andrew Cherniack, Broad

The Cancer Genome Atlas

Copy number alteration clusters





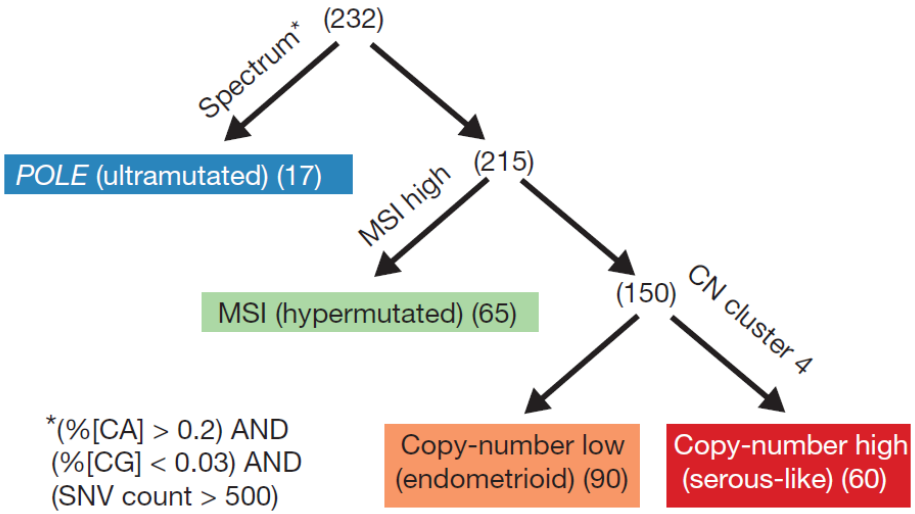
1q amplification

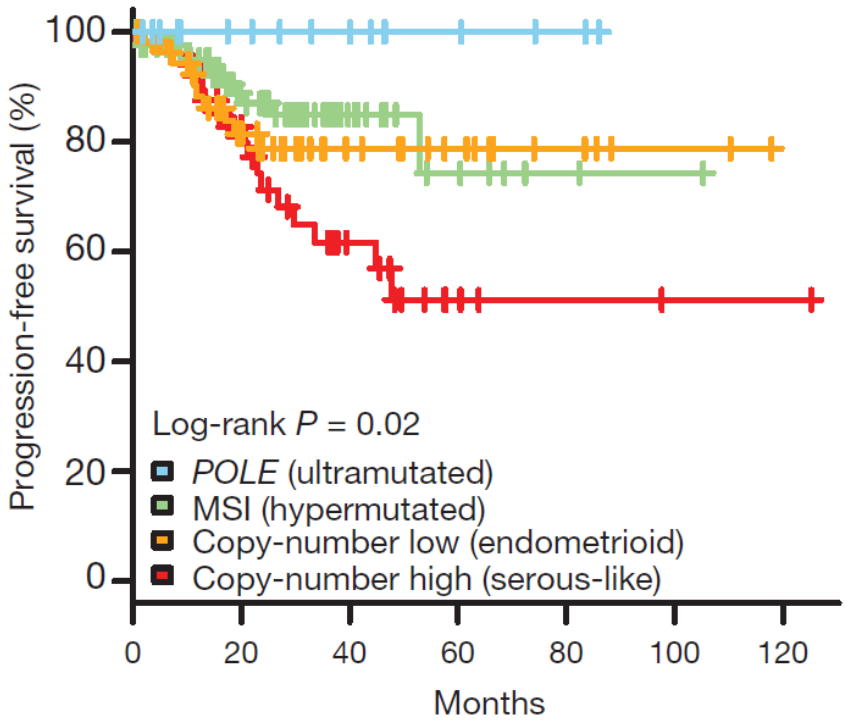
24% of high-grade endometrioid tumors cluster with serous tumors (**serous-like**)

- **Andrew Cherniack, Broad**

The Cancer Genome Atlas

Integrated genomic characterization of endometrial carcinoma





Nature 2013 May 2;497(7447):67-73

Integrated genomic characterization of endometrial carcinoma

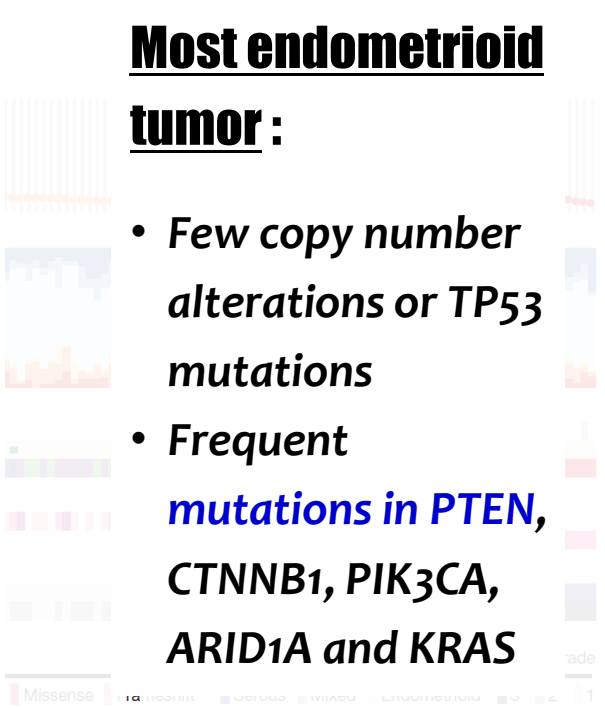
Four genomic-based subtypes of endometrial cancer

POLE ultramutated group	• Unusually high mutation rates • Hotspot mutations (sequences highly susceptible to mutation in the POLE gene
Hypermutated microsatellite instability group	• High mutation rate • Few copy-number alterations • Not carry mutations in the POLE gene
Copy-number low group	• The greatest microsatellite stability • High frequency of mutations in CTNNB1, critical gene for maintaining the linings of organs
Copy-number high group	• Mostly serous tumors, some endometrioid samples • High copy-number alteration; characteristics of serous tumor

Integrated genomic characterization of endometrial carcinoma

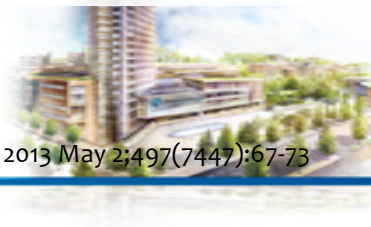
Uterine serous tumor & 25% of high-grade endometrioid tumor :

- *Extensive copy number alterations*
- *Few DNA methylation changes*
- *Low estrogen receptor/progesterone receptor levels*
- *Frequent TP53 mutations*



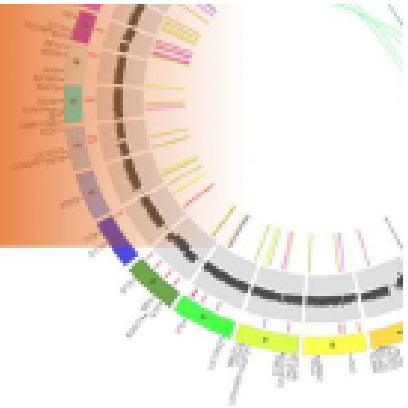
Most endometrioid tumor :

- *Few copy number alterations or TP53 mutations*
- *Frequent mutations in PTEN, CTNNB1, PIK3CA, ARID1A and KRAS*



Nature 2013 May 2;497(7447):67-73

Mutations in select genes

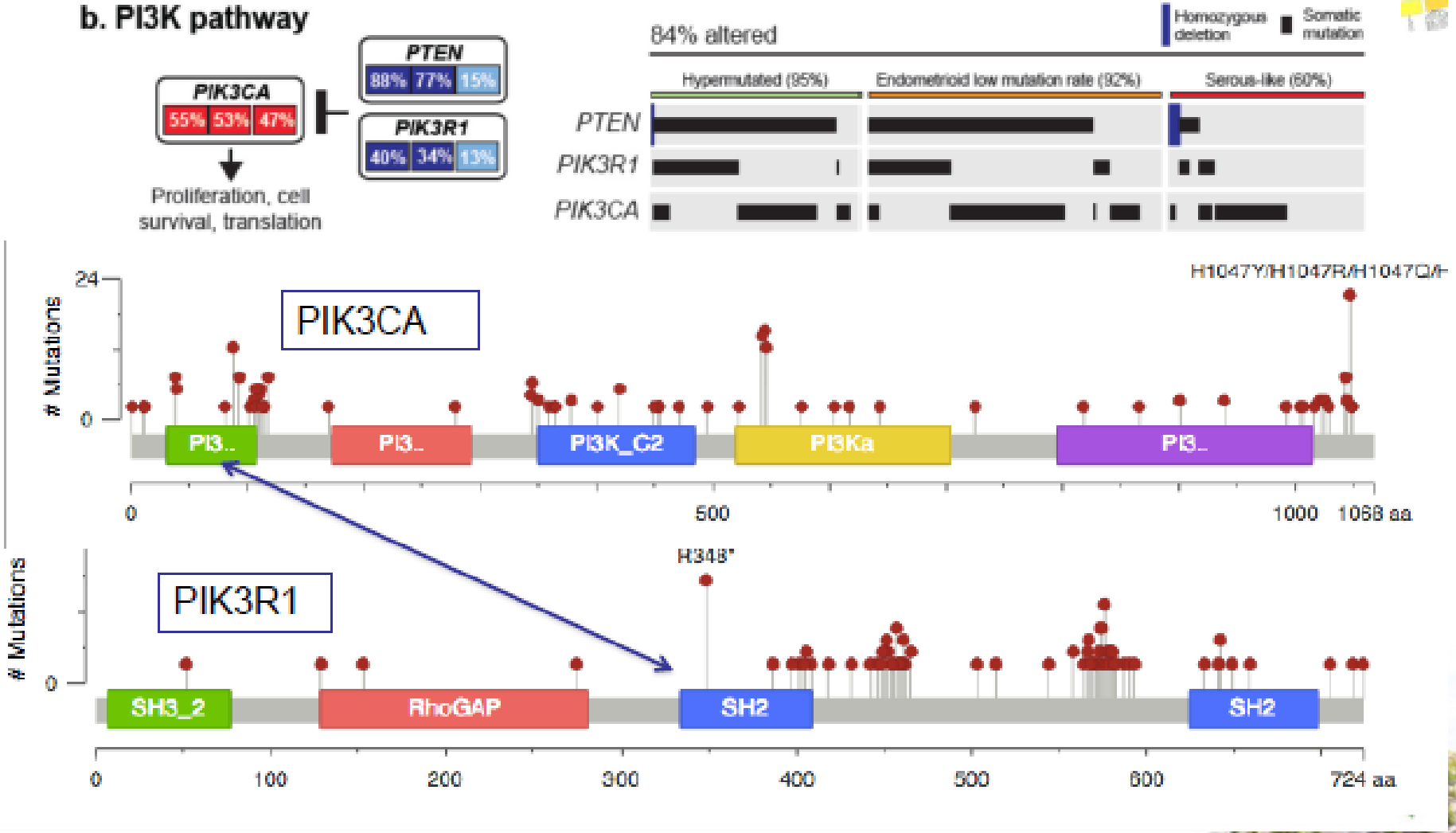


- PTEN mutations are uncommon in Serous cases and very common in low grade Endometrioid cases
- TP53 mutations are uncommon in low grade Endometrioid cases and very common in serous cases
- PIK3CA mutations are distributed across histology and grade
- Higher frequencies than previous reports may be due to more comprehensive sequencing methods

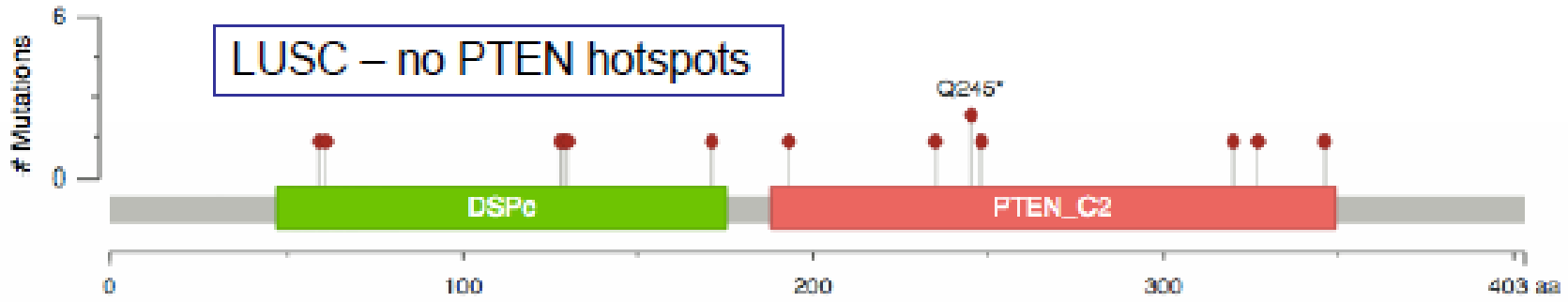
HistologyGrade	PTEN	TP53	PIK3CA	Total
EndoGr1	62 (0.83)	3 (0.04)	43 (0.57)	75
EndoGr2	62 (0.82)	9 (0.12)	38 (0.5)	76
EndoGr3	35 (0.71)	17 (0.35)	30 (0.61)	49
SerousGr3	1 (0.02)	39 (0.89)	19 (0.43)	44
Total	**160 (0.66)**	**68 (0.28)**	**130 (0.53)**	**244**

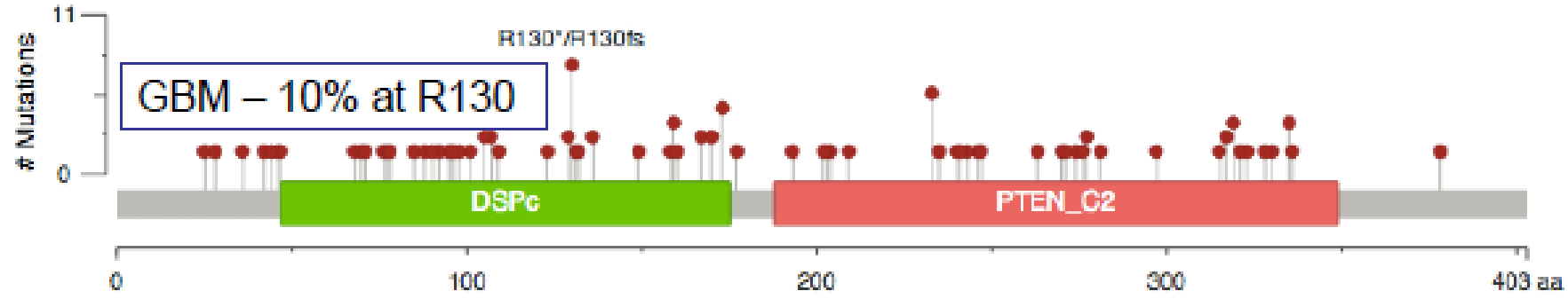
- **Cyriac Kandoth and Li Ding, WashU**

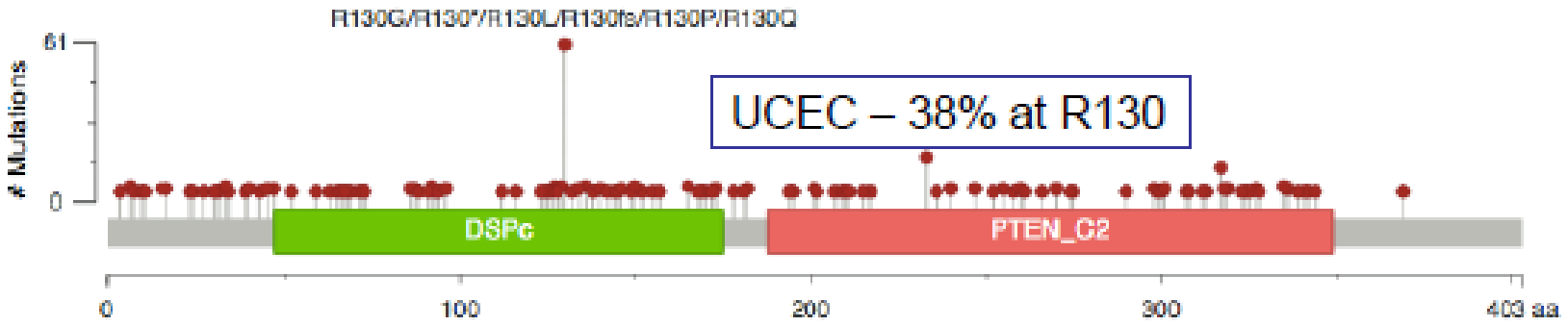
PI3K/AKT – most active in endometrial cancer



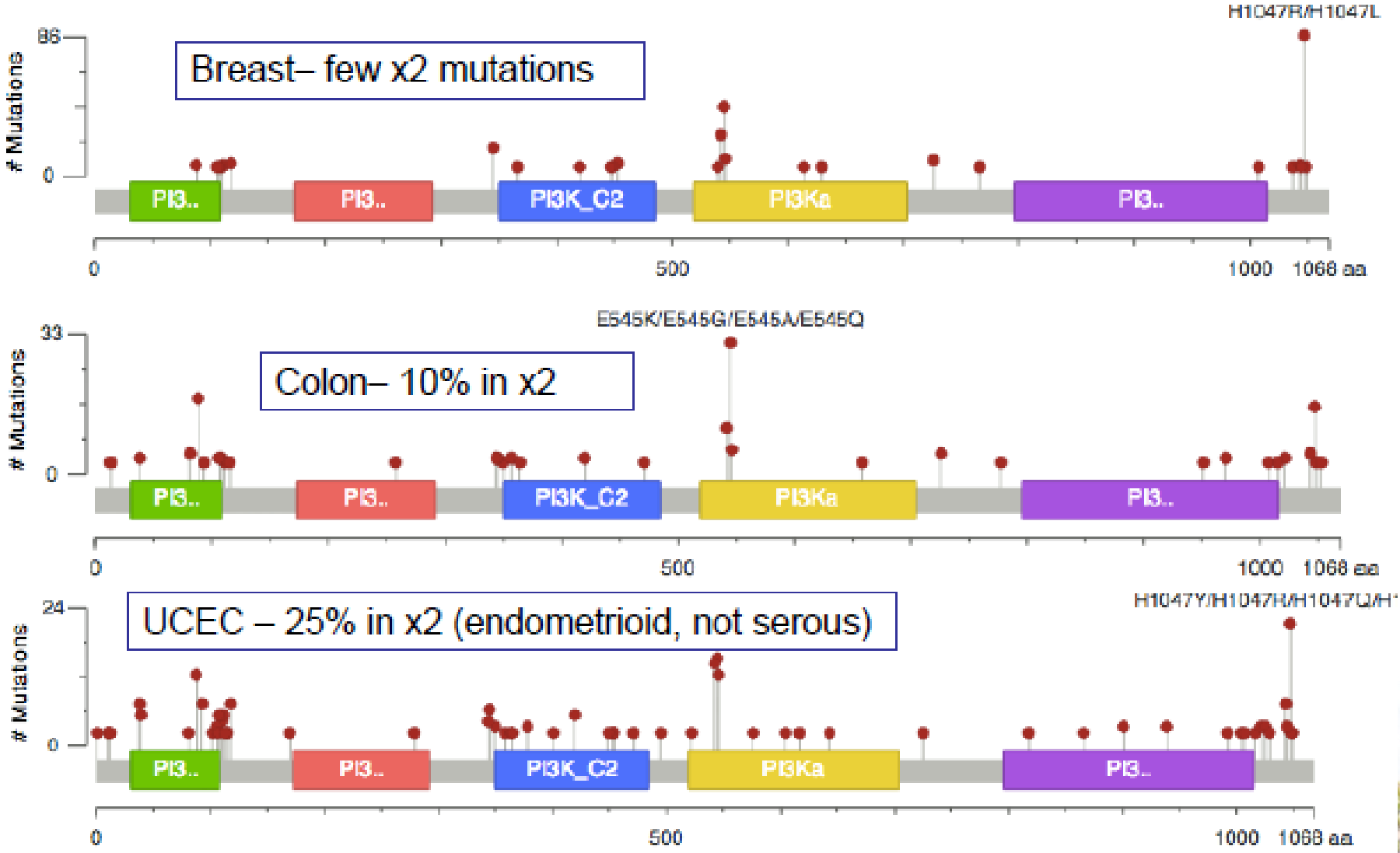
PTEN mutations



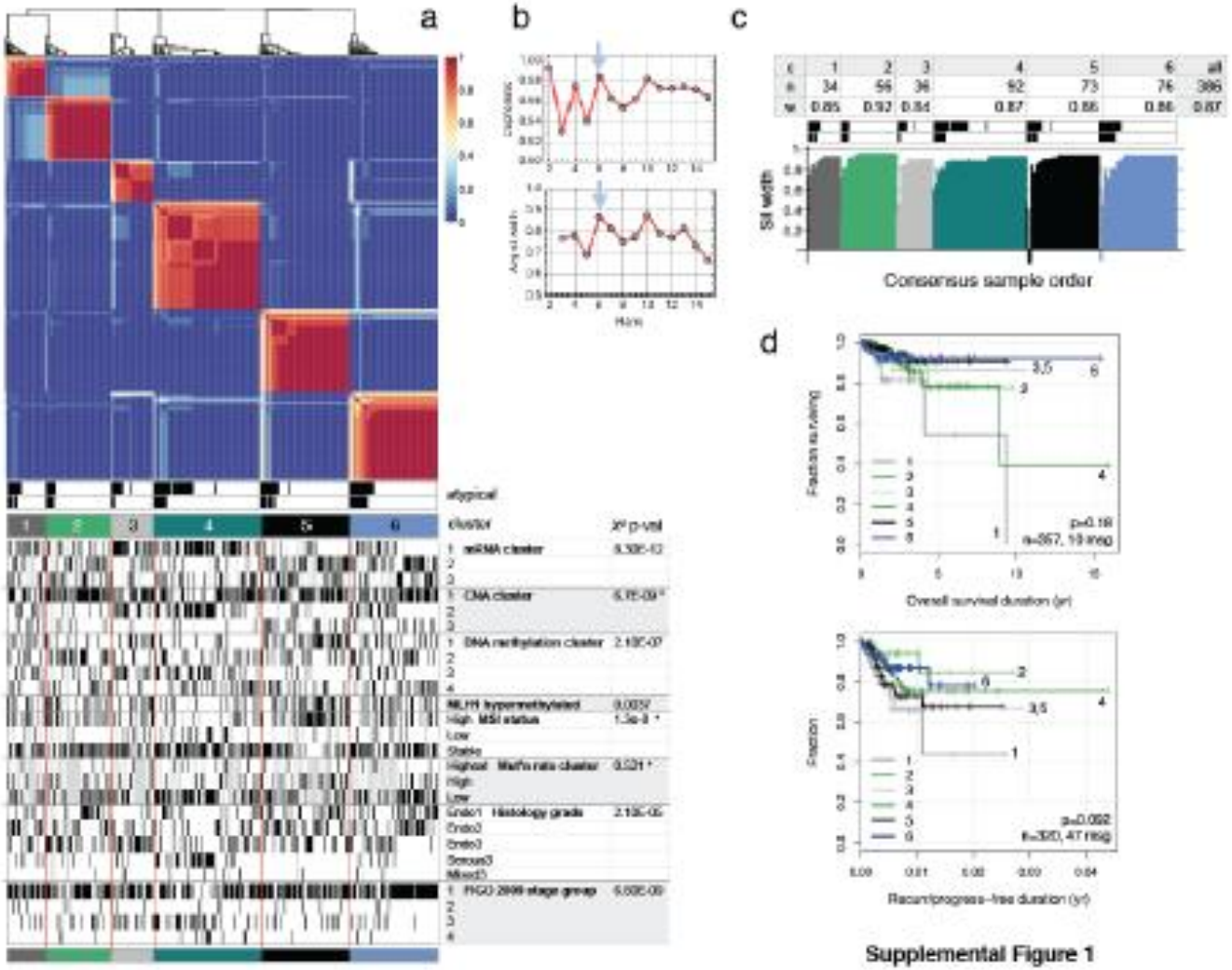




PIK3CA

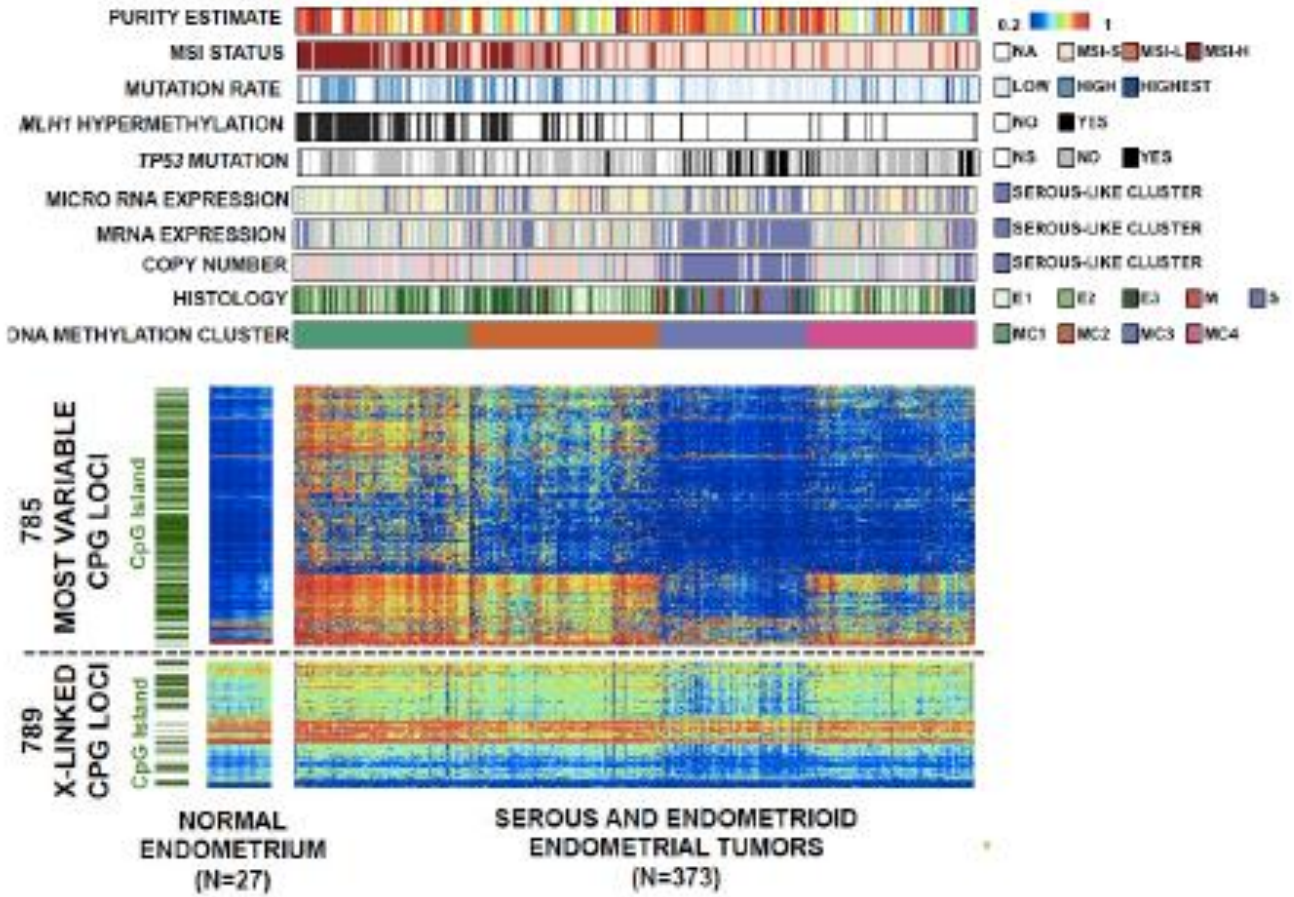

H1047R/H1047L
Mutations
86
0
Breast– few x2 mutations
PI3..
PI3..
PI3K_C2
PI3Ka
PI3..
0
500
1000
1068 aa
E545K/E545G/E545A/E545Q
Mutations
33
0
Colon– 10% in x2
PI3..
PI3..
PI3K_C2
PI3Ka
PI3..
0
500
1000
1068 aa
H1047Y/H1047R/H1047Q/H
Mutations
24
0
UCEC – 25% in x2 (endometrioid, not serous)
PI3..
PI3..
PI3K_C2
PI3Ka
PI3..
0
500
1000
1068 aa

MicroRNA sequencing



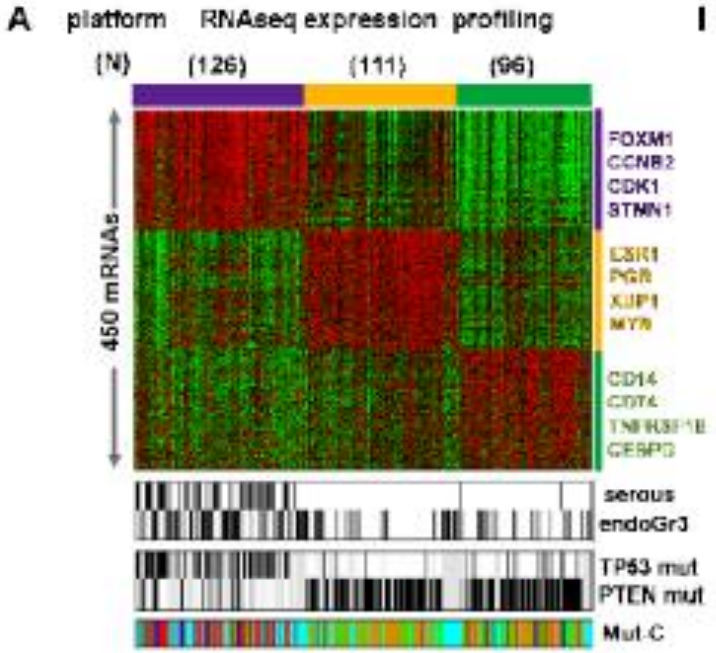
Supplemental Figure 1

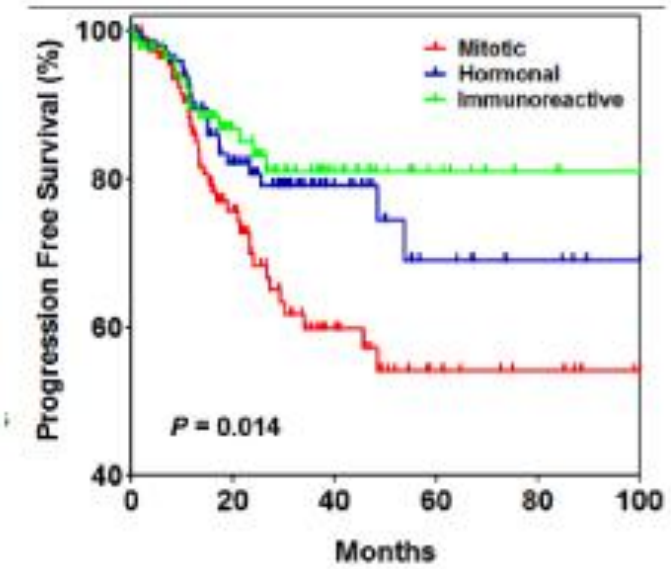
Methylation



The Cancer Genome Atlas

Gene expression clusters



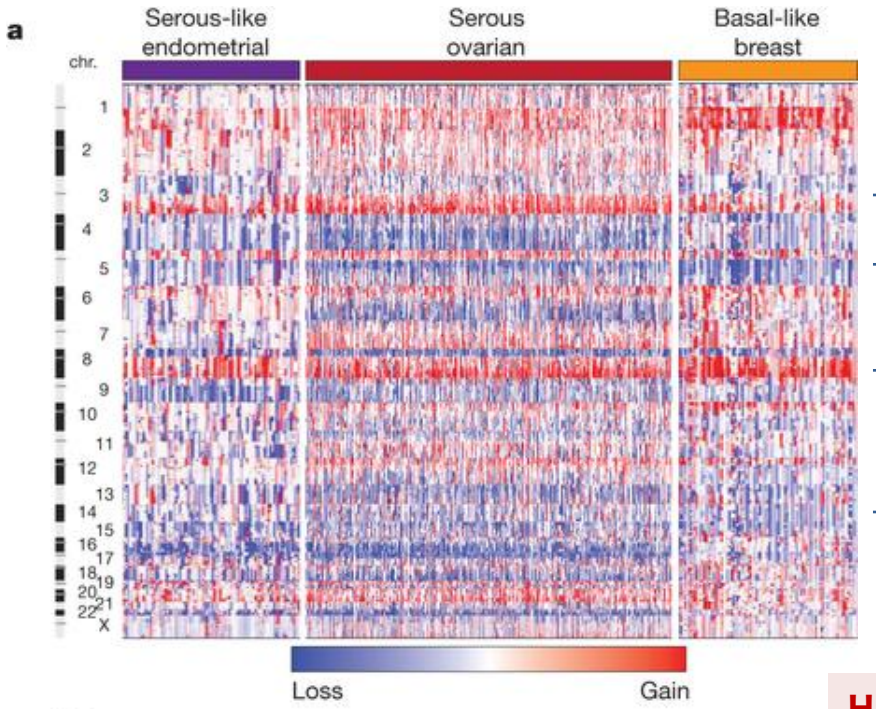


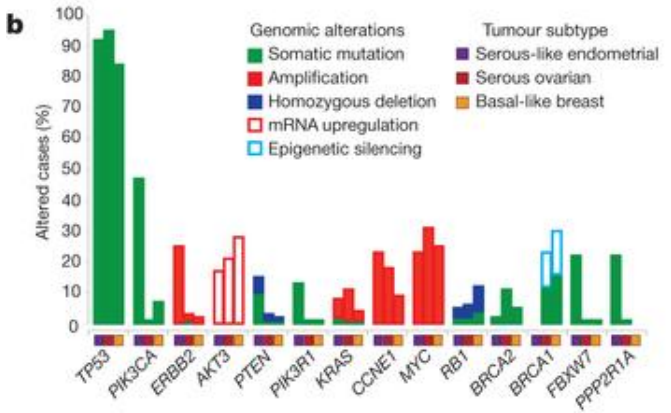
- Mitotic cluster contains serous and serous-like cases
- Hormonal cluster contains samples with greater ER/PR expression
- Immunoresponsive cluster contains immune activated genes

➢ **Wei Zhang and Yuexin Liu, MDACC**

The Cancer Genome Atlas

The Changing View of EmCa:
Discussion from TCGA





Molecular Characteristics

were similar between these three tumor subtypes (uterine serous, HGSOC, basal-like breast) and unsupervised clustering identified relatedness

High frequency of *TP53* mutations

- Uterine serous: 91%
- HGSOC: 96%
- Basal-like breast: 84%

Very low frequency of *PTEN* mutations

- Uterine serous: 2%
- HGSOC: 1%
- Basal-like breast: 1%

Differences included a higher frequency of FBXW7, PPP2R1A and PIK3CA mutations ***in uterine serous***

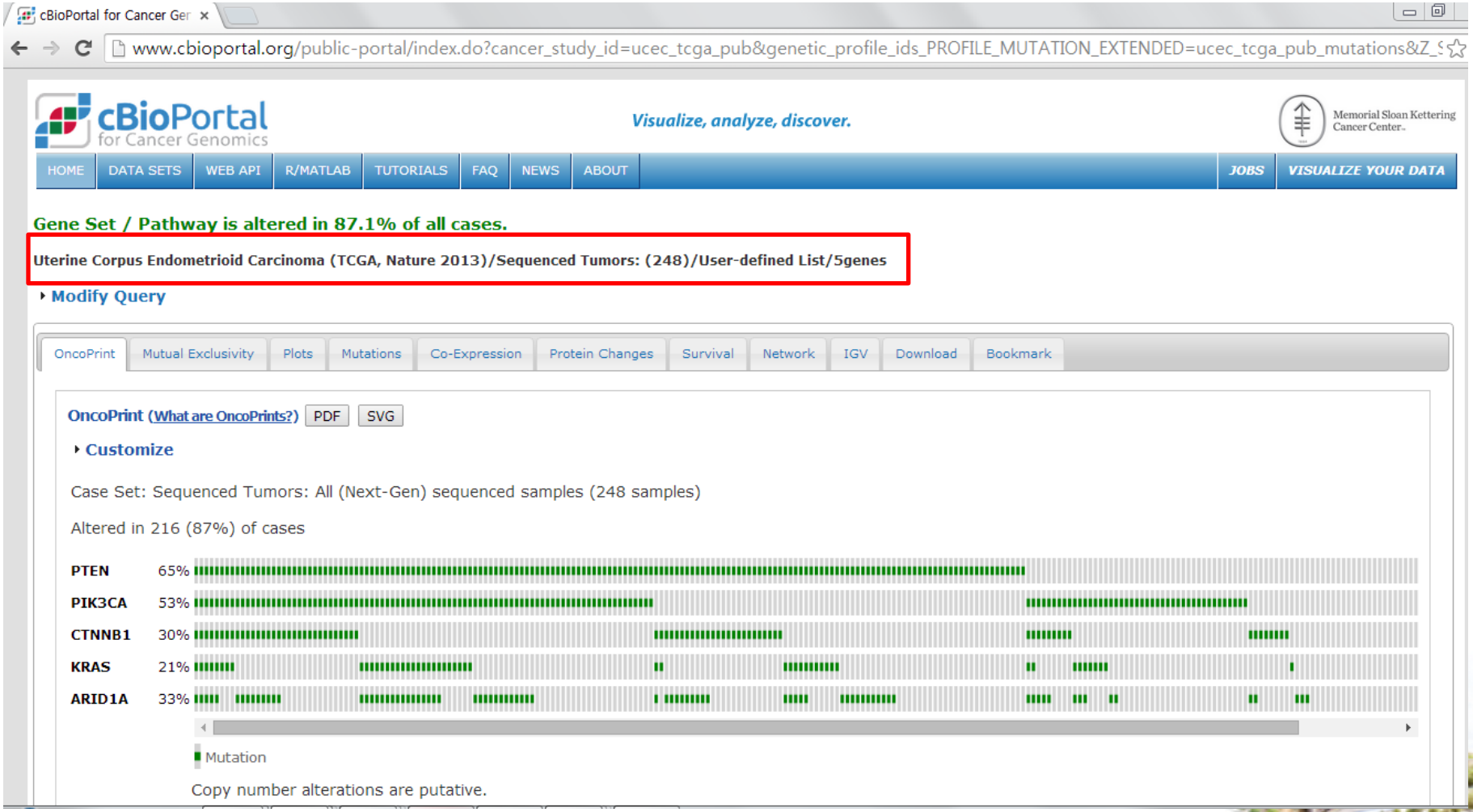
Nature 2013 May 2;497(7447):67-73

Summary:

- Integrated genomic characterization of endemetrial carcinoma -

- Recurrent POLE mutations identified and associated with altered mutation specturm and very high mutation rate
- PI3K/AKT pathway most activated in endometrial – ramifications for targeted inhibition, unique mutation spectra among genes
- Novel genomic stratification may complement or supplant histologic subtyping Has immediate impact on current schizophrenic approaches to adjuvant treatment after hysterectomy
- Warrants re-design of clinical trials with stratification or separation of subtypes
- In the era of 'precision medicine' these finding will help to bring targeted agents to the clinic in a rational manner

Data Mining from TCGA



POLE exonuclease domain mutation predicts long progression-free survival in grade 3 endometrioid carcinoma of the endometrium.

- explore the clinical and pathologic significance of POLE exonuclease domain mutations in high-grade endometrial carcinomas.
- assessed for mutations in the exonuclease domain of POLE and correlated POLE mutation status with clinicopathologic features and molecular parameters
- When analyzed together with published grade 3 endometrioid carcinomas by TCGA, the presence of POLE exonuclease domain mutation was associated with significantly better progression-free survival in univariate (p=0.025)

Meng et al. Gynecol Oncol 2014, 134:15

A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity

- Genes encoding components of the PI3K-AKT-mTOR signaling axis are frequently mutated in cancer, but few mutations have been characterized in mTOR, the gene encoding the mTOR kinase

- generated a comprehensive catalog of mTOR pathway mutations in cancer, identifying 33 MTOR mutations that confer pathway hyperactivation using TCGA
- hyperactivating MTOR mutations display heightened sensitivity to rapamycin both in culture and in vivo xenografts, suggesting that such mutations confer mTOR pathway dependency.

Gravineret al. Cancer discover 2014, 4:553

Multiplatform Analysis of 12 Cancer Types Reveals Molecular Classification within and across Tissues of Origin

- Recent genomic analyses of pathologically defined tumor types identify “within-a-tissue” disease subtypes. However, the extent to which genomic signatures are shared across tissues is still unclear

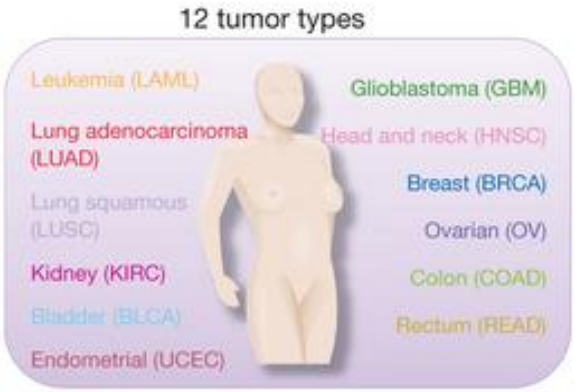
- integrative analysis using five genome-wide platforms and one proteomic platform on 3,527 specimens from 12 cancer types. Lung squamous, head and neck, and a subset of bladder cancers coalesced into one subtype typified by TP53 alterations, TP63 amplifications, and high expression of immune and proliferation pathway genes

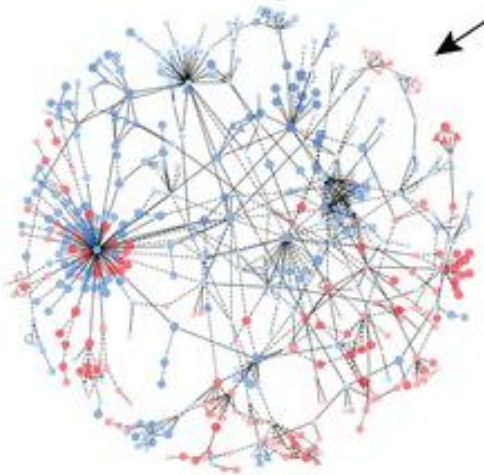
- All data sets are available for data-mining from a unified resource to support further biological discoveries and insights into novel therapeutic strategies..

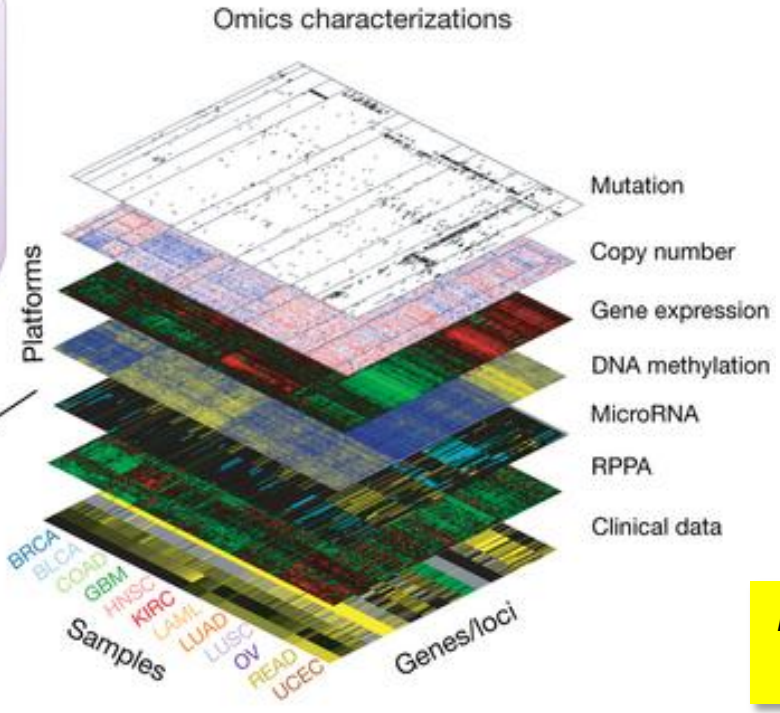
Hoadeley et al. Cell 2014, 158:929

The TCGA Pan-Cancer Project

Integrated data set for comparing and contrasting multiple tumor types







Six types of omics characterization were performed creating a 'data stack'

Maximizing the potential of integrative analysis

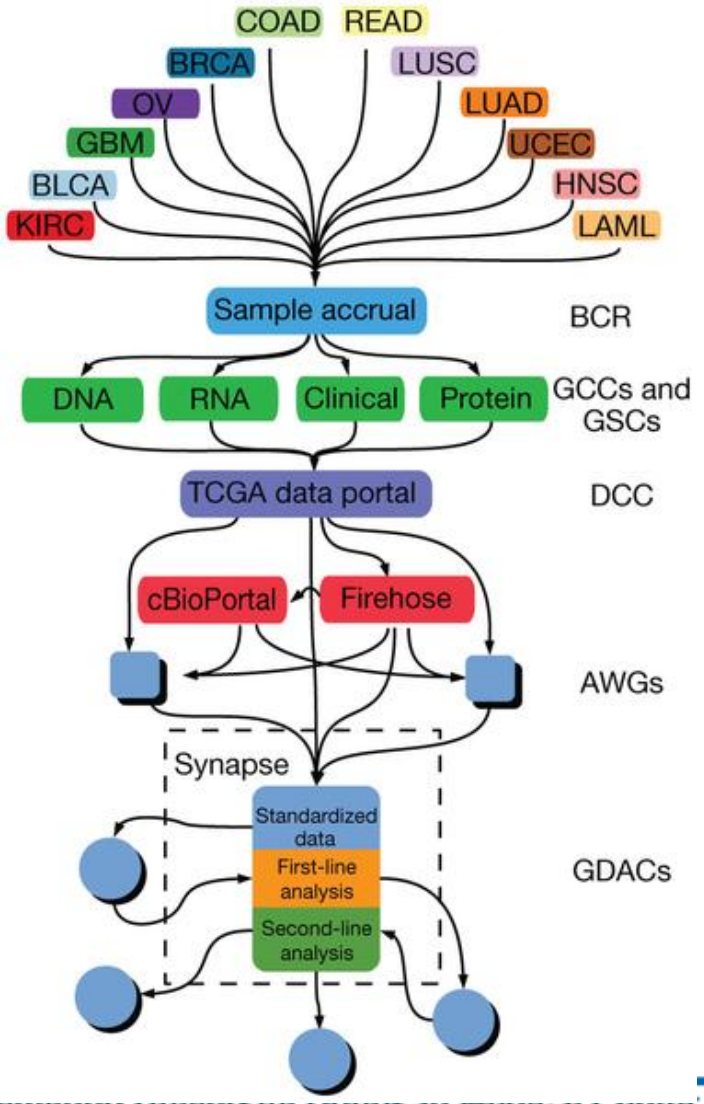
Identification of general trends, including common pathways

Cancer Genome Atlas Research Network, Nat Genet 2013

Data coordination for the Pan-Cancer TCGA project

Data were collected by the Biospecimen Collection Resource (**BCR**) from 12 different tumor types and characterized on 6 major platforms

Data sets were deposited in the TCGA Data Coordination Center (**DCC**) and distributed to the Broad Institute's **Firehose** and the Memorial Sloan-Kettering Cancer Center's **cBioPortal** for various automated processing pipelines.

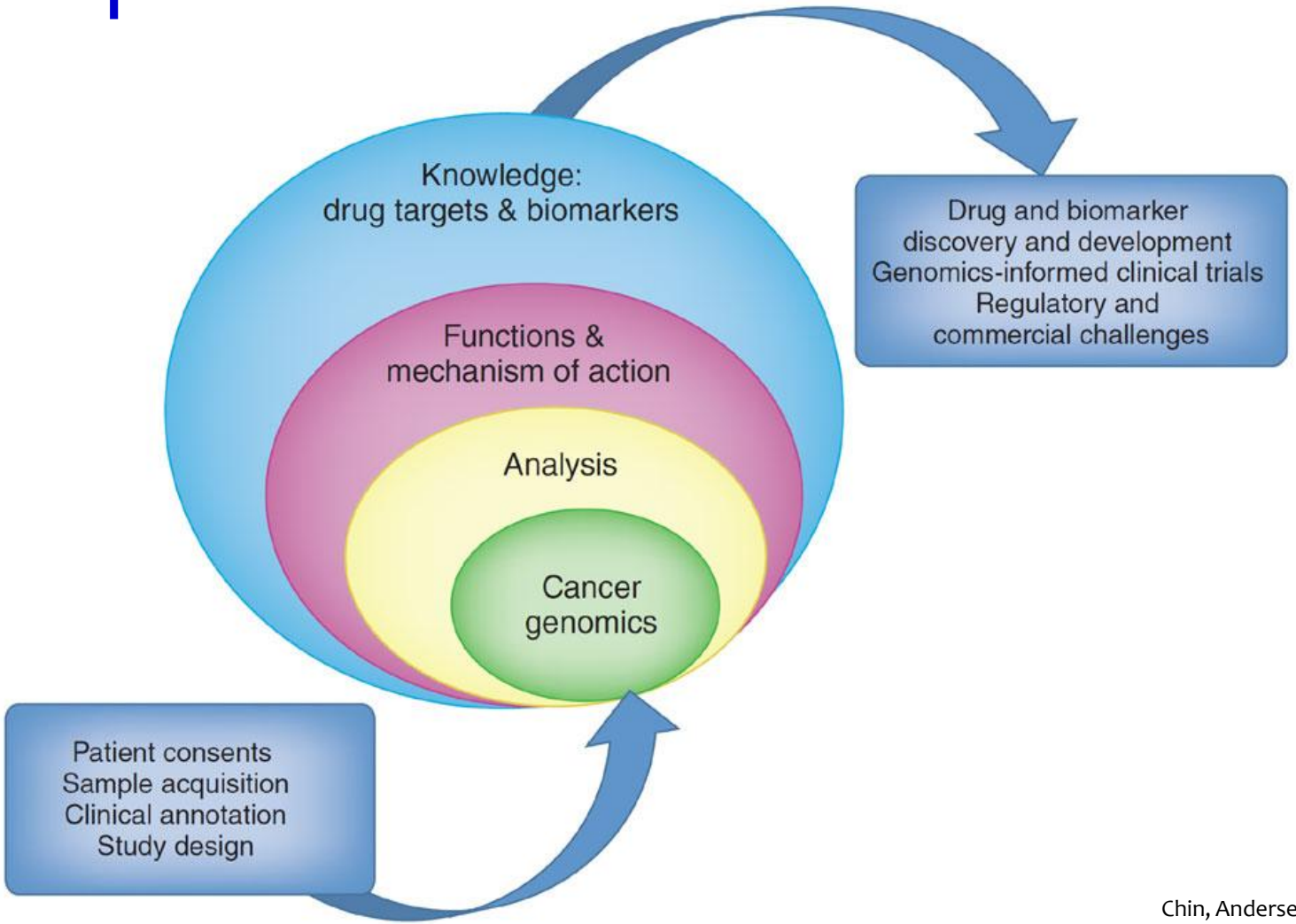


Genome Characterization Centers Genomic Sequencing Centers (**GCCs and GSCs**).

Analysis Working Groups (**AWGs**) conducted ***focused analyses*** on individual tumor types.

Genome data analysis centers (**GDACs**) accessed and deposited ***both data and results through Synapse to coordinate distributed analyses***.

Cancer Genomics: From discovery science to personalized medicine



Chin, Andersen et al. 2011

Thank you for your attention