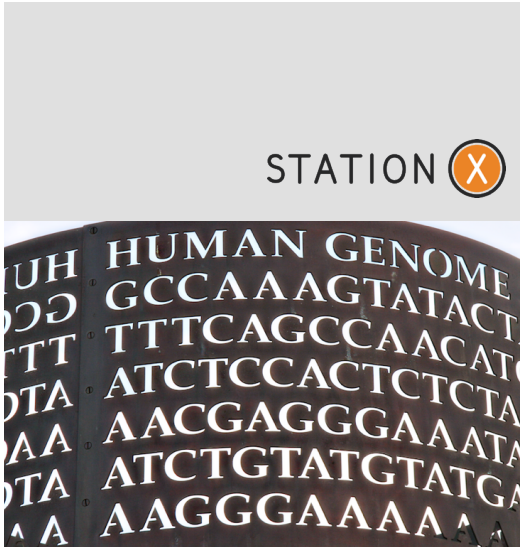
Icahn School of Medicine at Mount Sinai

Data Mining TCGA Breast and Ovarian Exomes for Novel Susceptibility Markers

JOHN A. MARTIGNETTI, MD, PhD

ICAHN SCHOOL OF MEDICINE AT MOUNT SINAI
DEPARTMENTS of GENETICS AND GENOMIC SCIENCES, PEDIATRICS, OBSTETRICS/GYNECOLOGY & REPRODUCTIVE SCIENCES and
ONCOLOGICAL SCIENCES

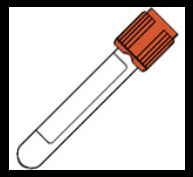


The Problem:

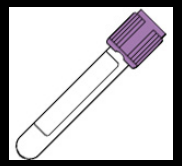
OVARIAN CA	BREAST CA
~ 20,000 cases	~ 210,000 cases
~ 14,000 deaths	~ 41,000 deaths
1/70	1/8

ALL GYN / ONC PATIENTS TREATED AT MOUNT SINAI* ARE INVITED TO PARTICIPATE

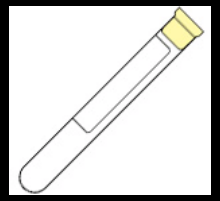
Specimen acquisition: Mount Sinai BioRepository



Serum



PBMC
DNA, RNA



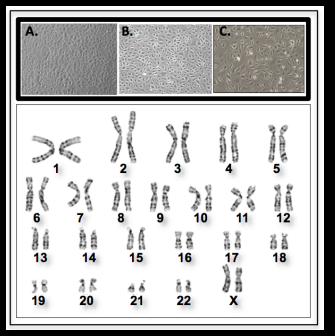
Lymphoblastoid
Cell culture

Blood

Tumor

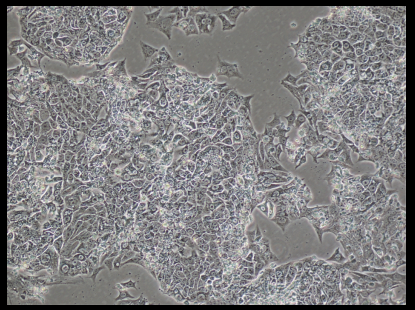
Ascites

Primary Cell Culture

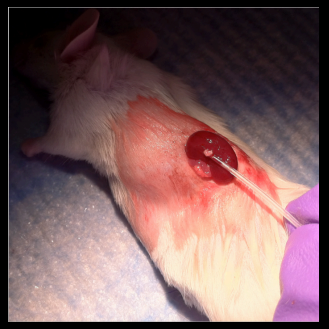


Proteomics

Primary Cell Culture



Animal Models

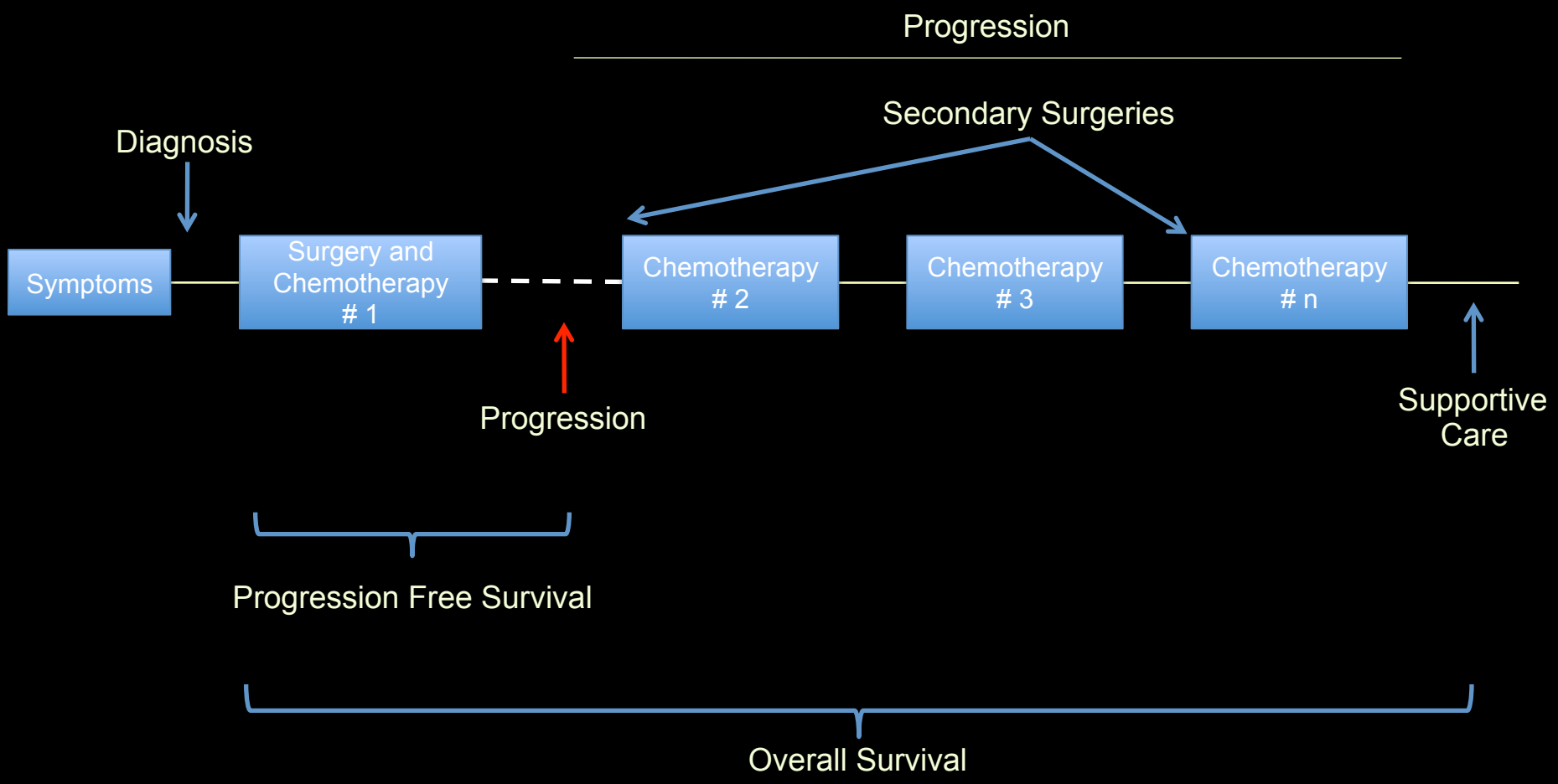


Next Generation Sequencing

digitized record

Mount Sinai Hospital – Beth Israel Medical Center – Elmhurst – Englewood Hospital - Mount Sinai Queens – Beth Israel Brooklyn - Roosevelt Hospital – St. Luke's Hospital

Ovarian Cancer: Natural History



Important strides in detecting increased risk for some forms of cancer.

RATIONALE FOR THE APPROACH: FAMILY-BASED STUDY TO IDENTIFY OVARIAN AND BREAST CANCER SUSCEPTIBILITY GENES

- Family history is the strongest single predictor of a woman's chance of developing breast and / or ovarian cancer(s).

- While BRCA1/2 mutations still represent the strongest known genetic predictors, they are responsible for less than 50% of all families containing two or more cases in first-degree relatives and explain less than 50% of the excess familial cancer risk.

- Genetic studies seeking to identify breast and ovarian cancer susceptibility genes have therefore focused on those families with a high incidence of cancer across multiple generations. Avoids many of the technical, clinical, statistical issues associated with GWAS.

- Overcomes issues of “rare” alleles: no bias against identification of rare disease-causing alleles. Some families will harbor a “private” mutation; whereas others may share a gene.

Patient-centric / Family-centric bias.

Things may become evident on the population level which were not evident when viewed in isolation.

MATERIALS & METHODS

- To identify these novel susceptibility genes, we sequenced germline DNA of selected families with hereditary breast and ovarian cancer lacking deleterious BRCA1/2 mutations

Breast Cancer: > 70 families with 3 or more affected (Univ. of Chile)

[a number of families with male affecteds]

Ovarian Cancer: 72 families with > 2 affecteds (Roswell Park)

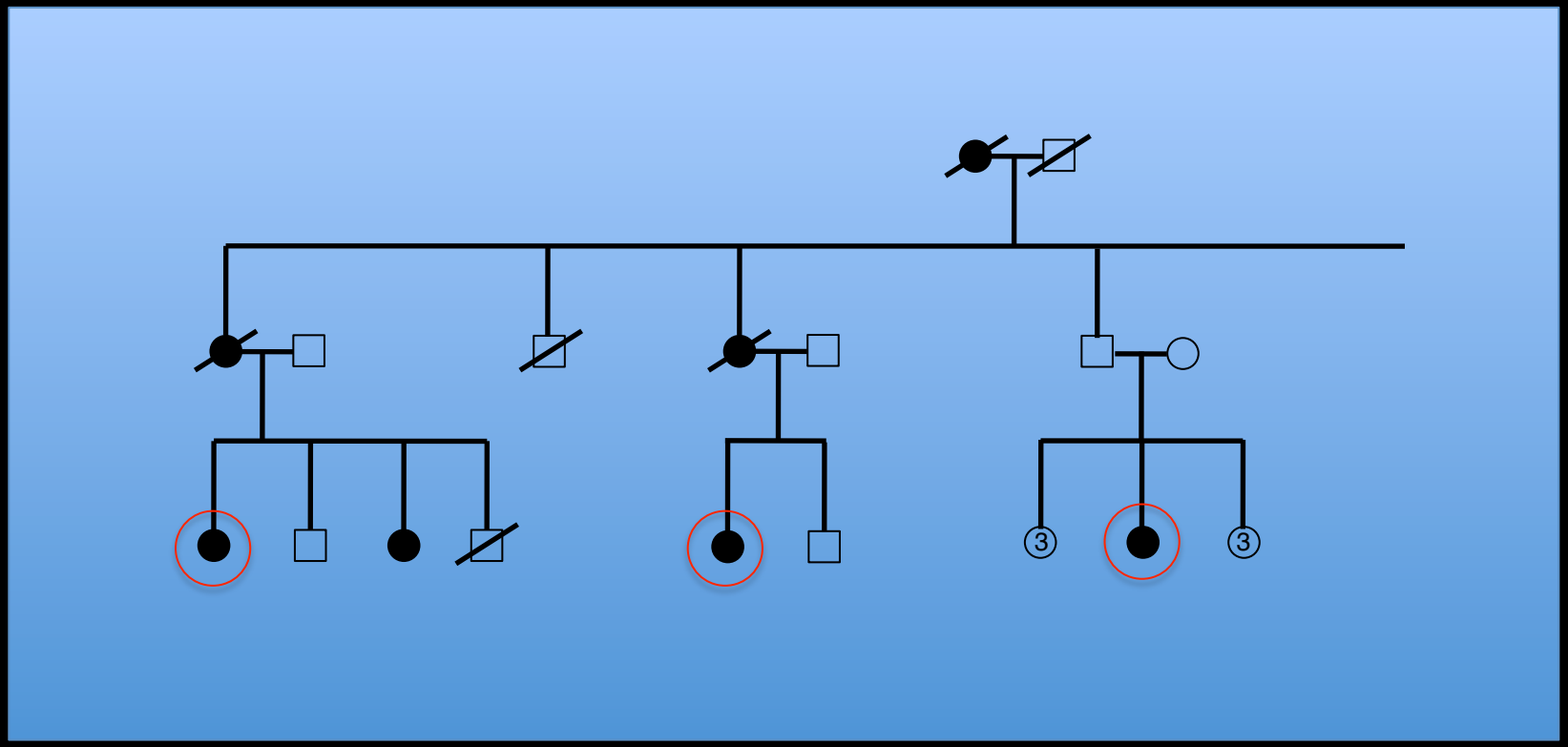
- 21 Exomes were sequenced at the Icahn School of Medicine at Mount Sinai using Illumina sequencing technology and Agilent SureSelect Exome Capture protocol with on-target coverage depths ranging from ~80x to 250x

- Read alignments and small variations called by a standard BWA-GATK bioinformatics pipeline

- Variants were annotated, visualized, and analyzed within GenePool™ (Station X, Inc., San Francisco, CA)

REPRESENTATIVE PEDIGREES

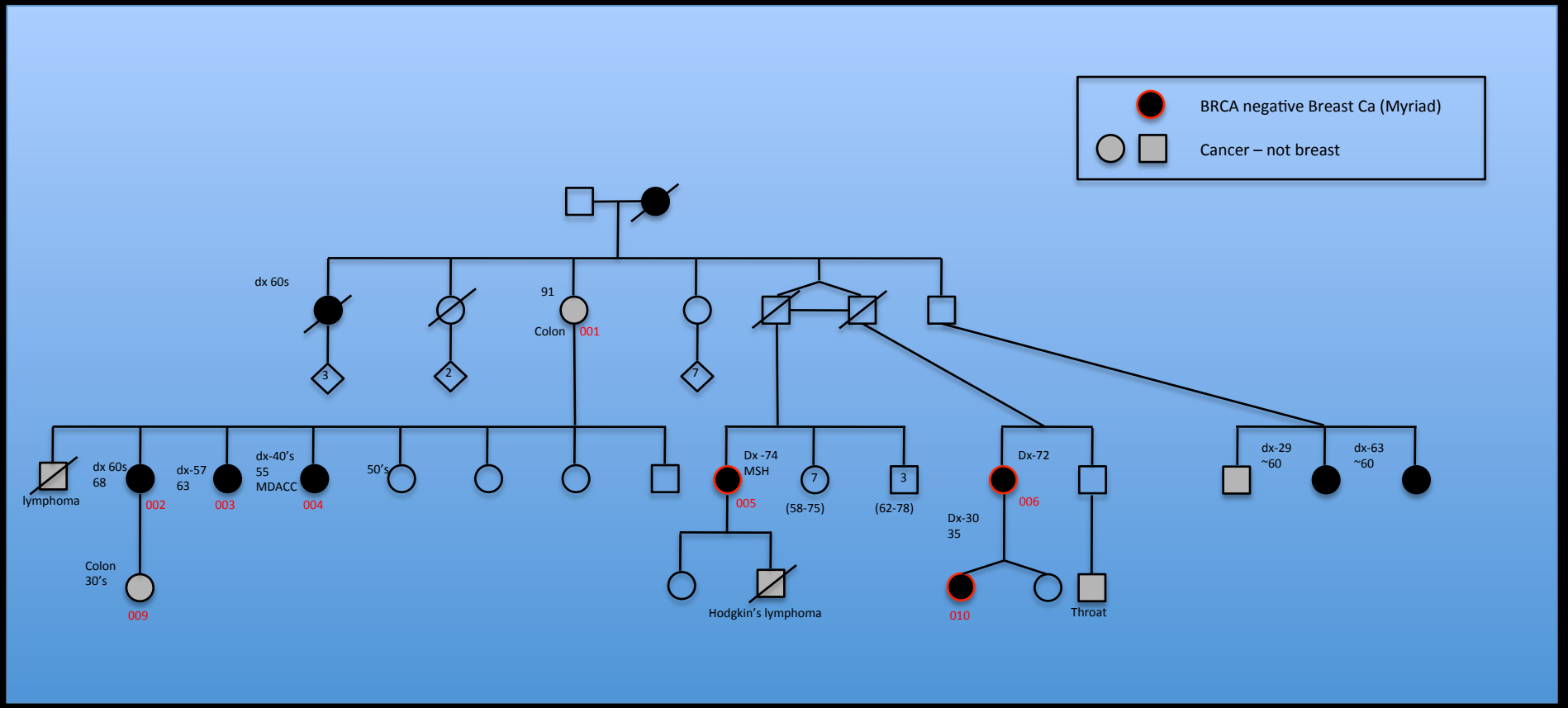
Ovarian Cancer Family 311



16 - age of onset 48; 06 - age on onset 43; 55 - age onset 25

REPRESENTATIVE PEDIGREES

NYBR 01

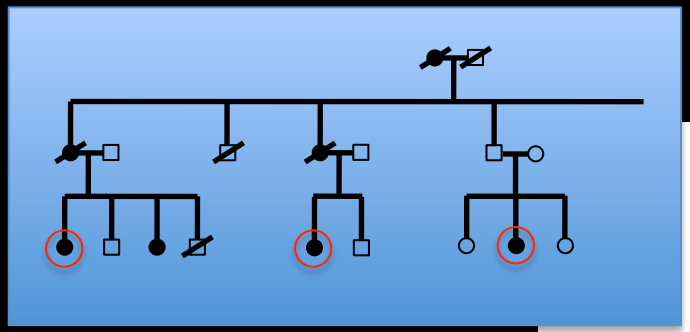


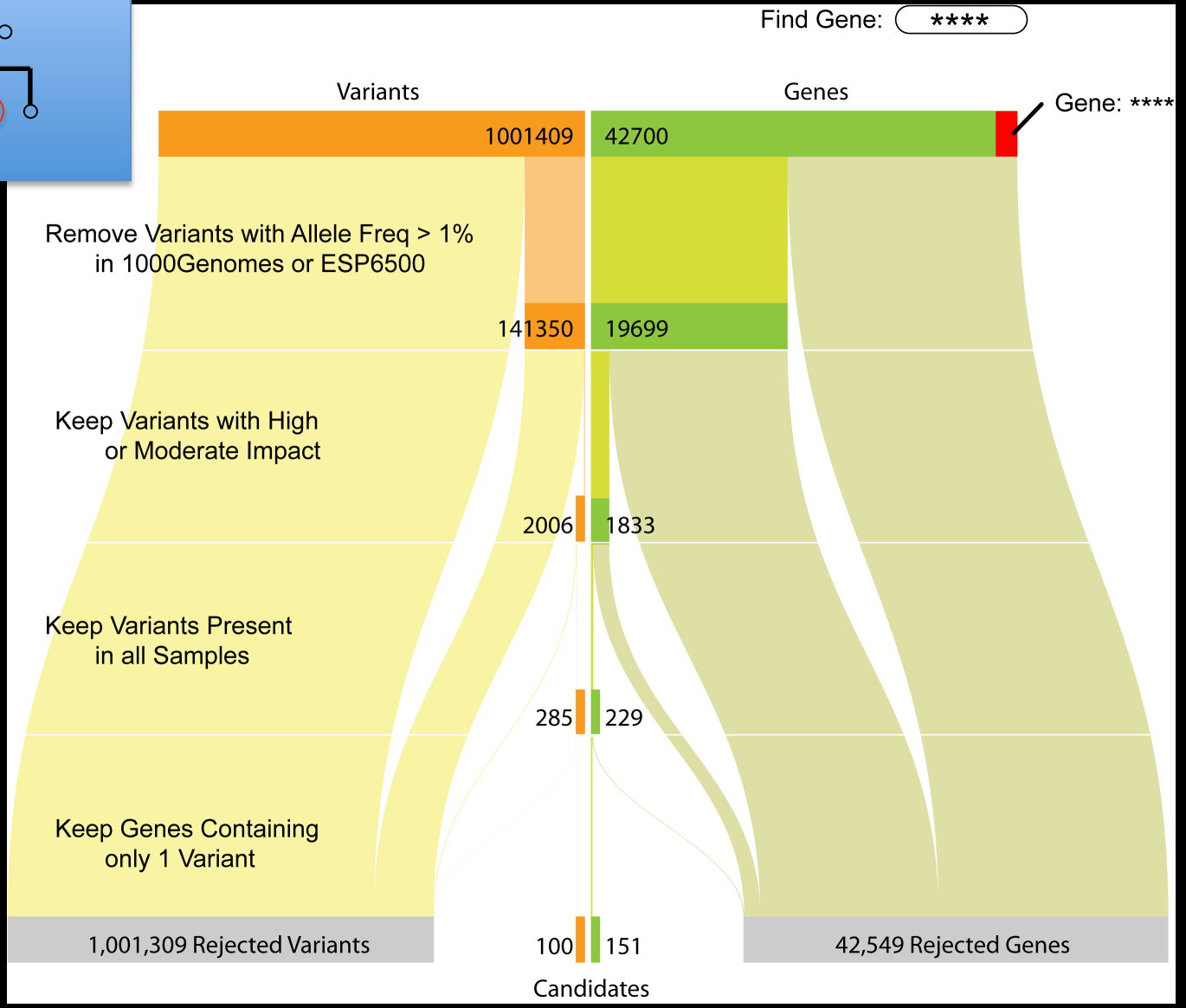
Average Coverage Per Base for BRCA1 and BRCA2 Exons

Indication	Sample	BRCA1	BRCA2
Ovarian	F311_006	128.39	143.39
Ovarian	F311_016	174.21	136.53
Ovarian	F311_055	161.05	97.03
Ovarian	F1515_001	236.49	168.29
Ovarian	F1515_017V2	157.48	132.64
Breast	F135P1	127.68	130.50
Breast	F135P2	160.97	161.67
Breast	F147P1	124.69	126.11
Breast	F151P1	127.82	122.68
Breast	F249P1	129.42	131.01
Breast	F249P2	143.60	145.50
Breast	F269P1	117.05	112.96
Breast	F272P1	107.20	114.55
Breast	F272P2	116.71	121.40
Breast	F294P1	150.71	154.91
Breast	F310P1	149.59	150.58
Breast	F310P2	112.94	112.66
Breast	F45P1	132.37	130.95
Breast	NYB01	150.12	82.09
Breast	NYB03	112.58	61.73
Breast	NYB10	15.96	5.56

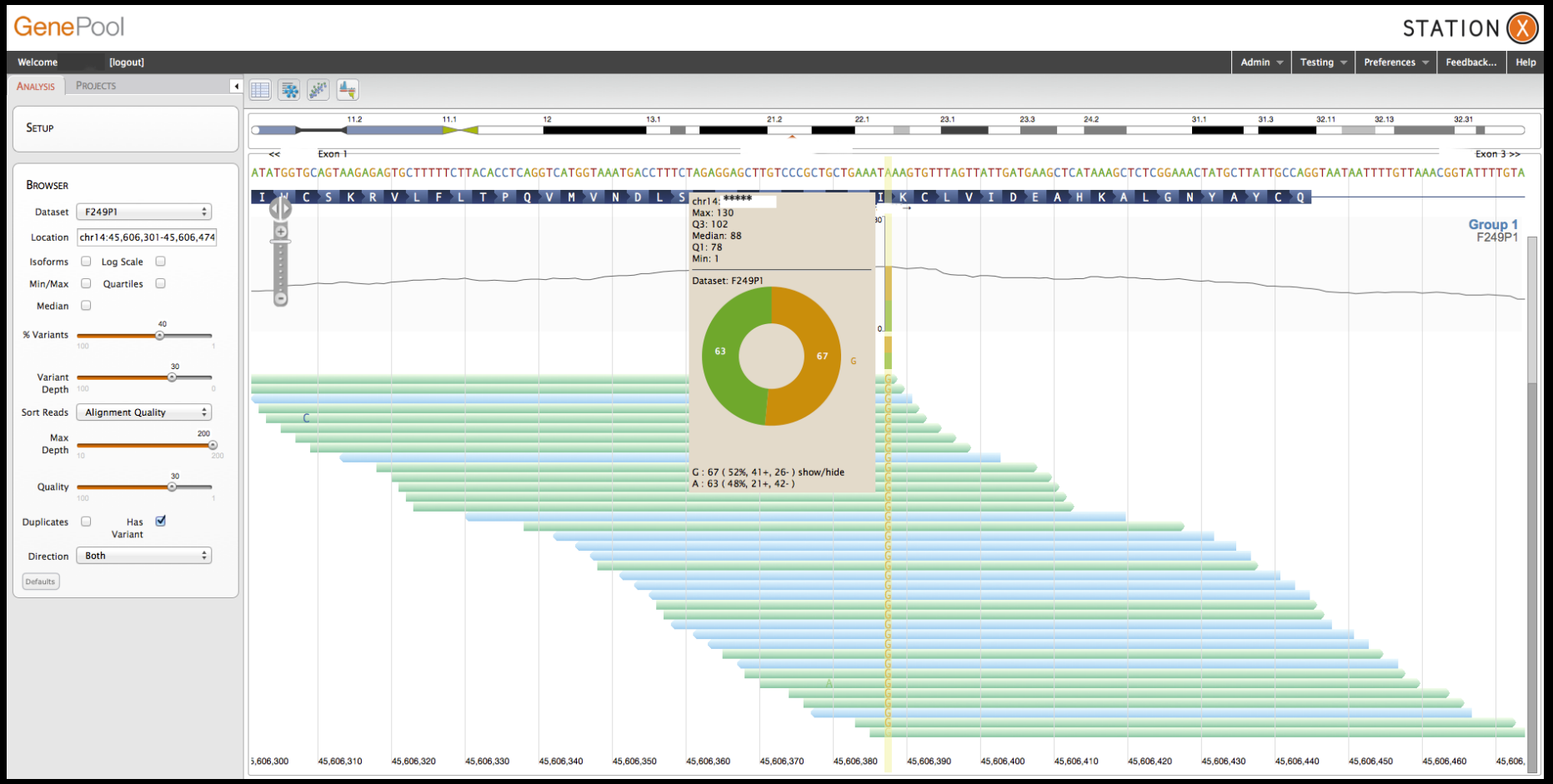
BRCA1 mutation – Family excluded from further analysis

Selecting for Candidate Mutations: Germlines of Related Ovarian Cancer Patients

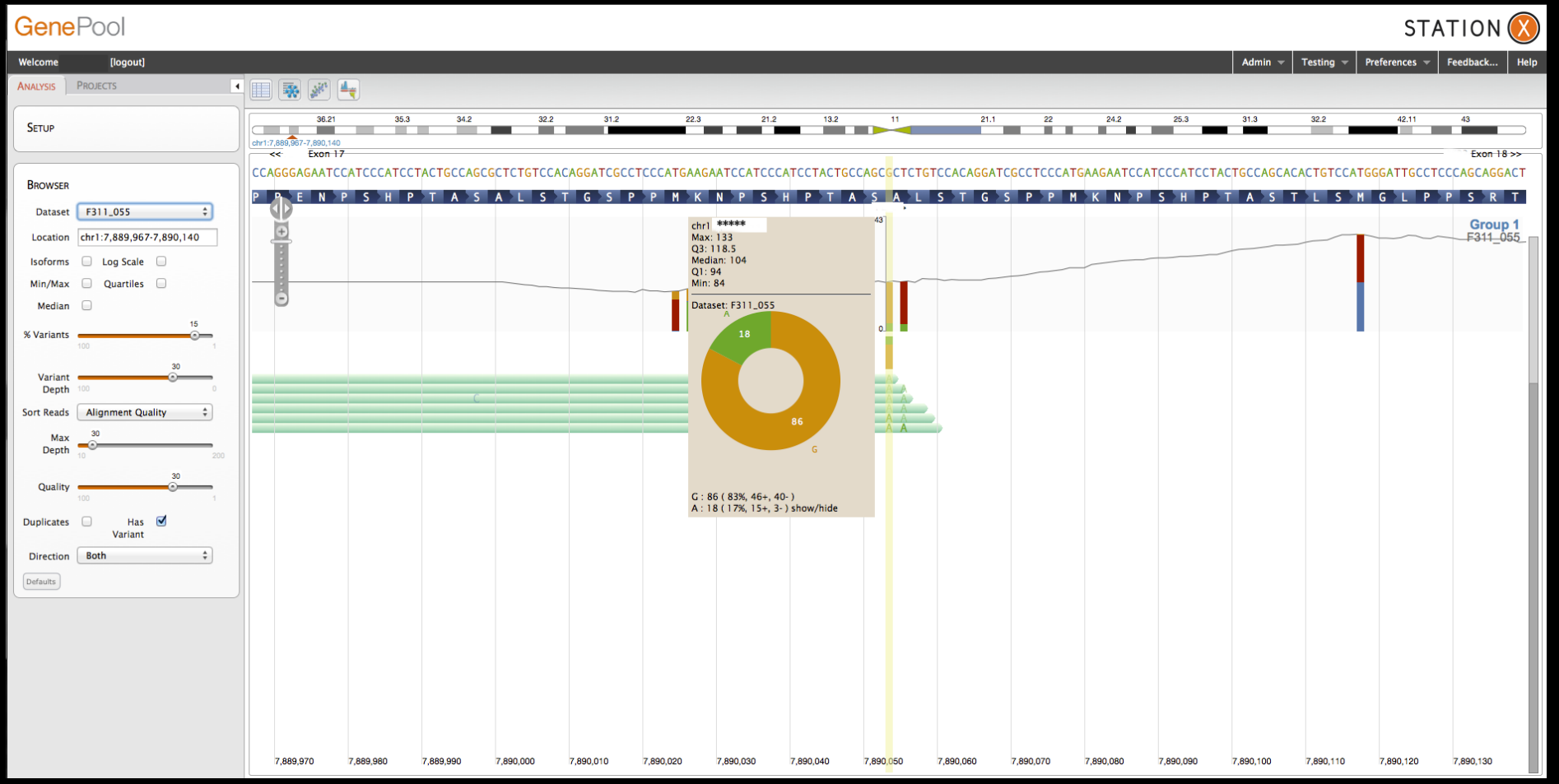




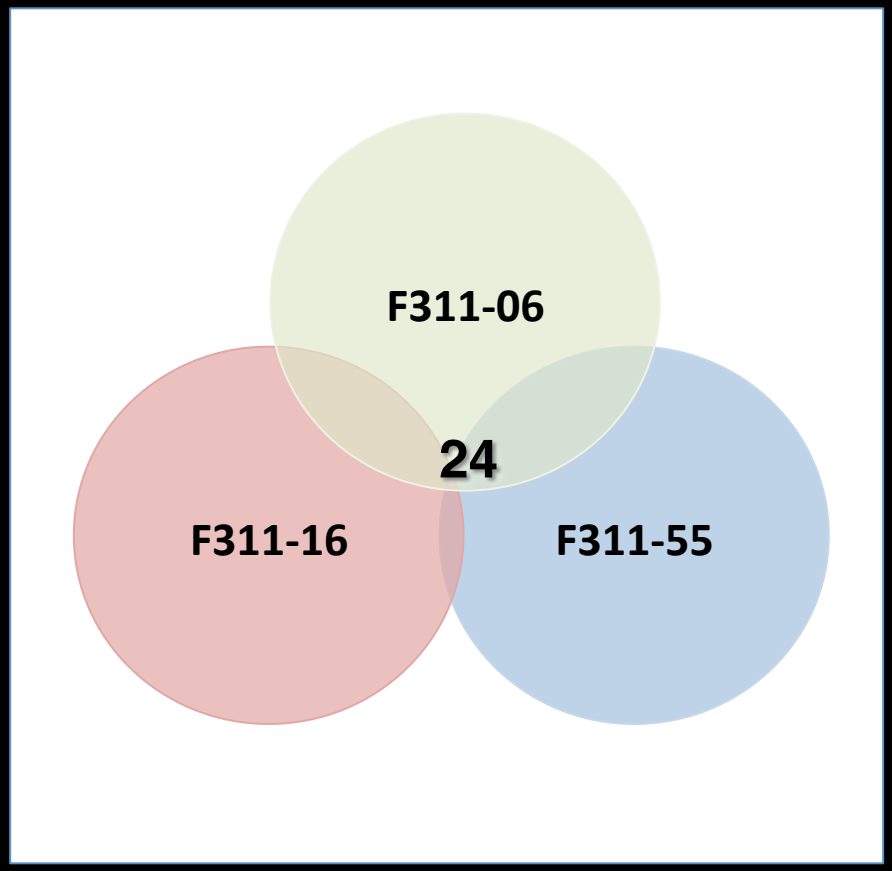
Variants Likely to Validate

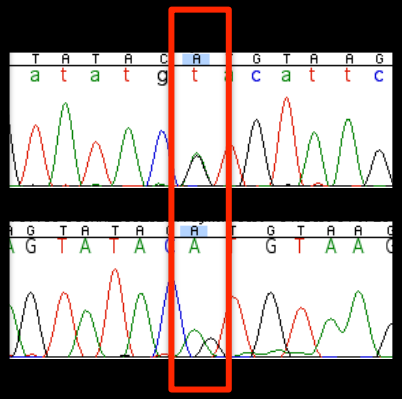


Variants Unlikely to Validate



Fam311: 24 Candidate OvCA Genes





All validated
by Sanger Sequencing

TCGA: OvCA Germline Variant Landscape

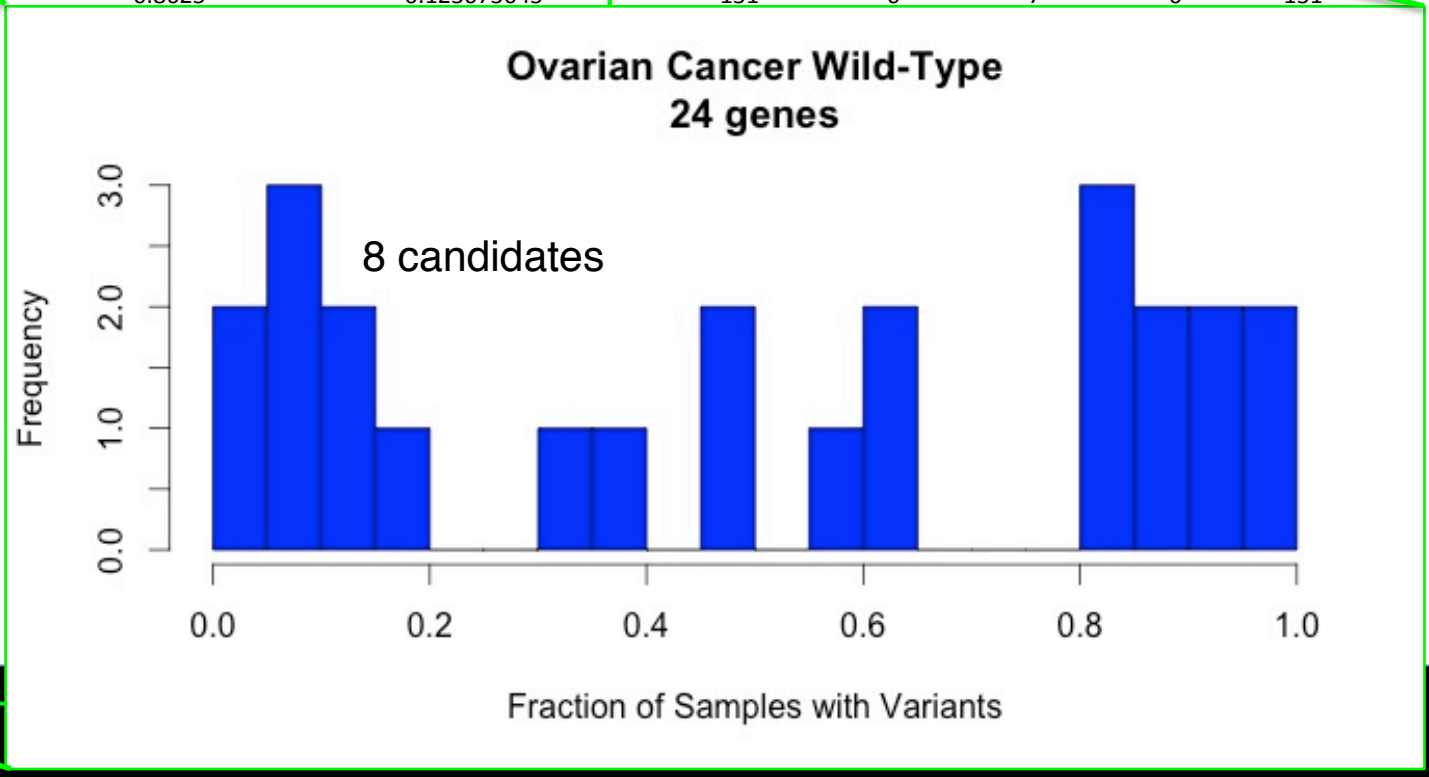
OvCA Candidates

Nam	Fraction of Samples	Average Number of Variants	Average per Kilobase	# Het	# Hom	# CompHet	# High	# Moderate
	0.8125	4.04375	0.703138585	130	47	125	8	130
	0.925	3.94375	3.858855186	148	0	148	2	148
	0.0375	0.04375	0.018784886	6	0	1	0	6
	0.81875	0.8625	0.123673645	131	0	7	0	131
	0.8875	1.975	1.683716965	142	4	134	39	141
	0.475	1.03125	0.442027432	76	0	58	1	76
	0.05	0.05	0.02293578	8	0	0	0	8
	0.10625	0.1125	0.031610003	17	0	1	0	17
	0.9625	6.5625	99.43181818	154	67	146	55	154
	0.825	4.6375	0.106756446	132	0	132	5	132
	0.875	2.21875	3.138260255	140	0	83	37	137
	0.56875	0.625	0.641683778	88	4	8	1	91
	0.075	0.075	0.021676301	12	0	0	0	12
	0.1	0.1125	0.031380753	16	0	2	0	16
	0.93125	1.78125	2.410351827	149	0	112	0	149
	0.46875	0.5	1.098901099	67	8	5	72	8
	0.39375	0.4	0.344827586	52	11	1	59	5
	0.34375	0.55	0.779036827	39	16	23	35	52
	0.05625	0.05625	0.019932672	8	1	0	0	9
	0.975	1.7125	1.082490518	140	31	65	47	156
	0.1875	0.2	0.115008626	30	0	2	0	30
	0.14375	0.15	0.043277553	22	1	1	0	23
	0.65	1.18125	0.281786737	104	0	48	3	102
	0.65	1.89375	2.108853007	103	5	96	0	104

TCGA: OvCA Germline Variant Landscape

OvCA Candidates

Nam	Fraction of Samples	Average Number of Variants	Average per Kilobase	# Het	# Hom	# CompHet	# High	# Moderate
	0.8125	4.04375	0.703138585	130	47	125	8	130
	0.925	3.94375	3.858855186	148	0	148	2	148
	0.0375	0.04375	0.018784886	6	0	1	0	6
	0.81875	0.8625	0.123673645	131	0	7	0	131
	0.8875							
	0.475							
	0.05							
	0.10625							
	0.9625							
	0.825							
	0.875							
	0.56875							
	0.075							
	0.1							
	0.93125							
	0.46875							
	0.39375							
	0.34375							
	0.05625							
	0.975							
	0.1875							
	0.14375							
	0.65							
	0.65							

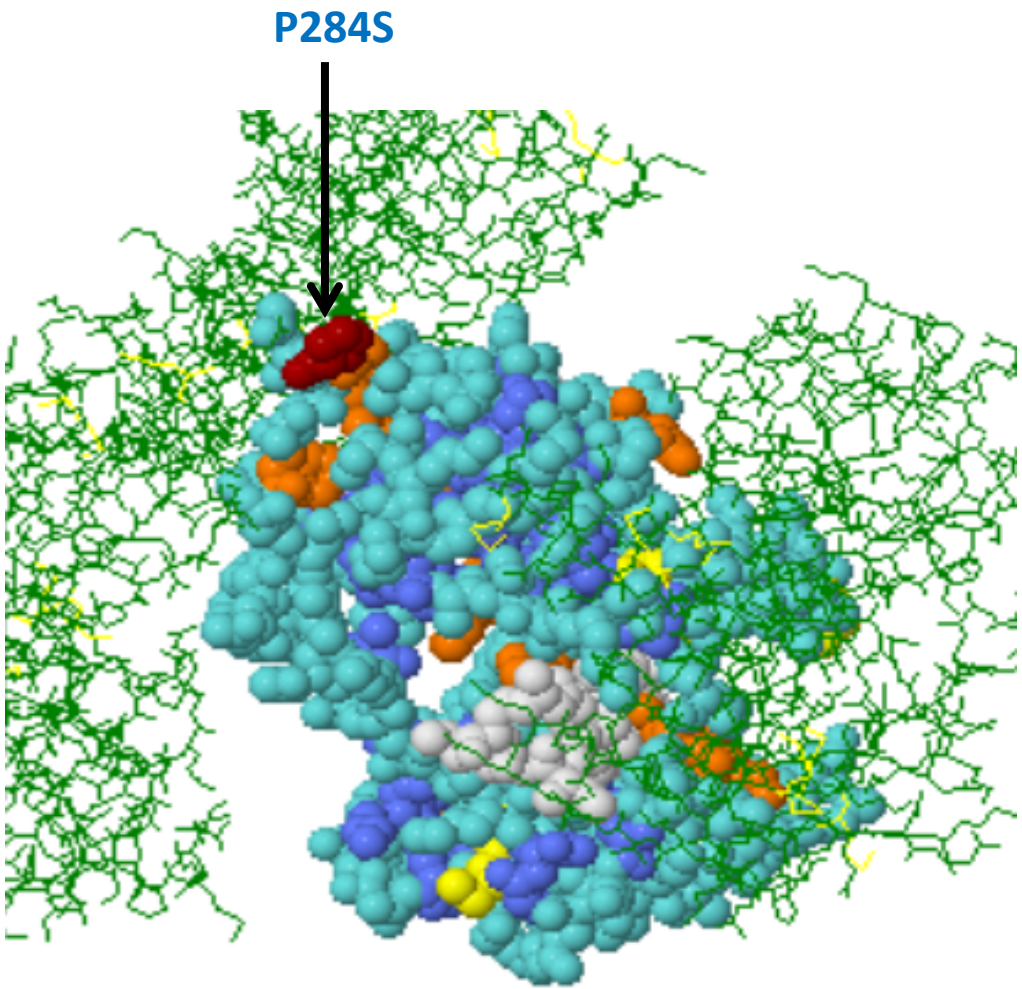


5 of these had the specific variant in frequencies of 1-2% in our WT BRCA subpopulation – 2 of these were ENRICHED from the general population

Parallel lines of support: Functional Impact of Mutations

MA mutationassessor.org
functional impact of protein mutations
release 2

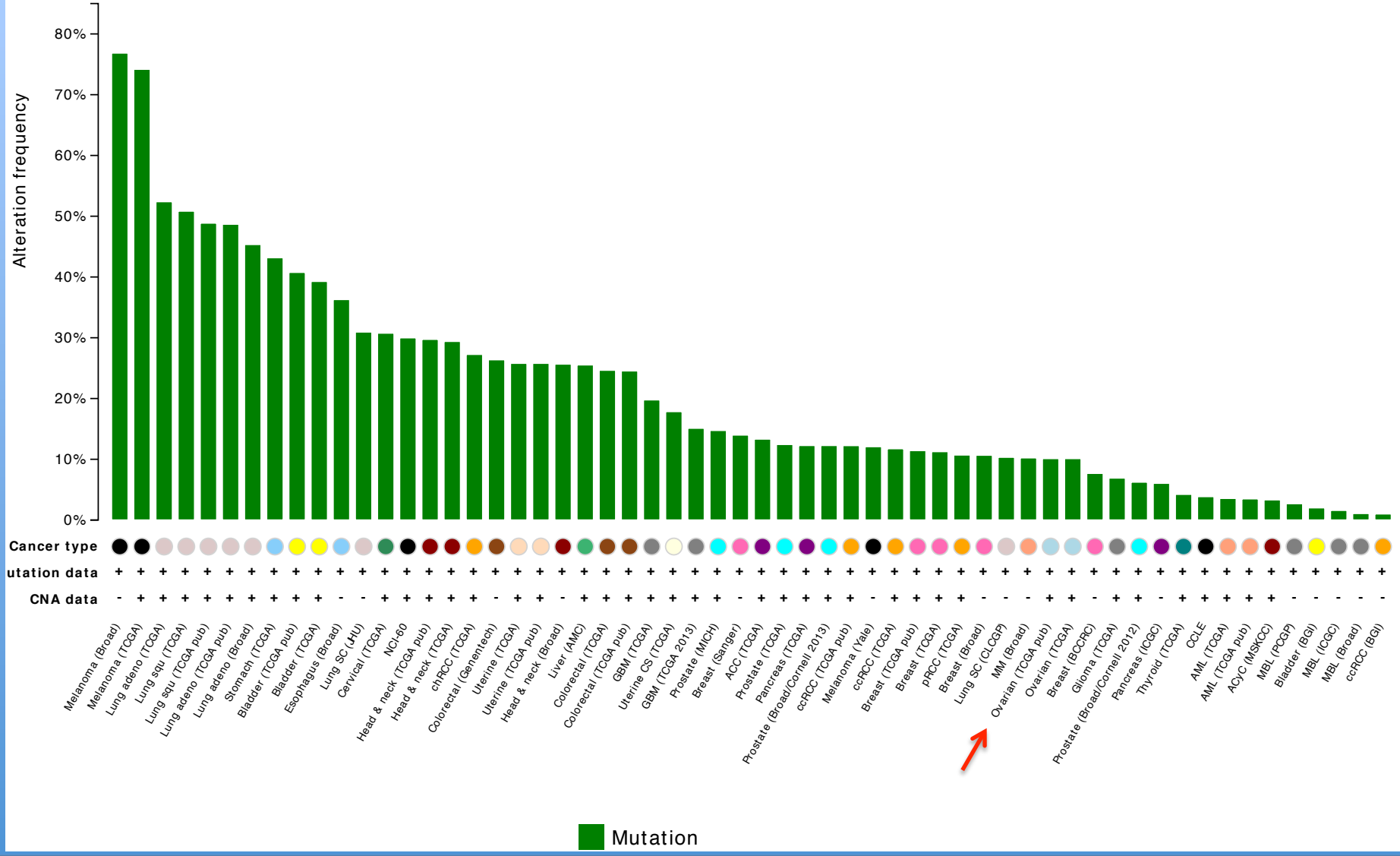
cBio@MSKCC
papers | about | changes | how it works



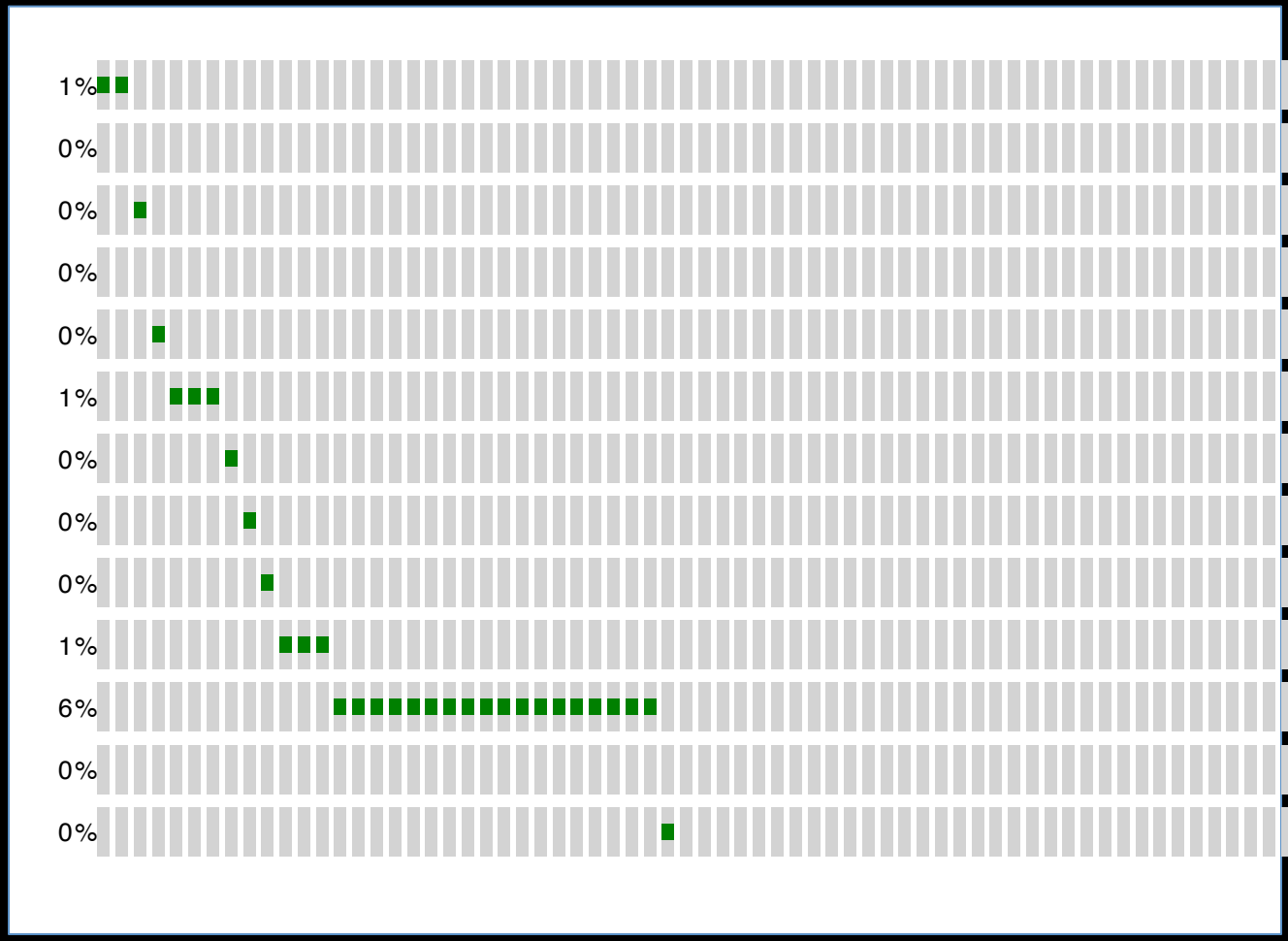
The population frequency of this variant is ~0.7% in European and ~0.6% in American populations (1000 genomes)

7 variants are assessed as functional:
6 variants are likely result in loss of function; 1 variant is a potentially new type of "switch of function" mutation;
5 variants do not have population frequency in 1000 genomes variant frequencies;
2 variants have minor population frequencies (~1%);
3 variants affect genes which are involved in cancer

TCGA: PAN Cancer ANALYSIS – cBIO PORTAL



TCGA Case Set – Serous OvCA (TCGA, Nature 2011): 316 samples



CANDIDATE OVCA3: MUTATION PROFILE

A RARE PRIVATE ALLELE

NO MUTATIONS IDENTIFIED IN ENTIRE OVCA3 GENE IN THE 70 OVCA FAMILIES ….however……

Ovarian TCGA:	E210D
	N310S
	D739H
Breast:	R955G
	L51V
	D383H
	S930C

nature COMMUNICATIONS

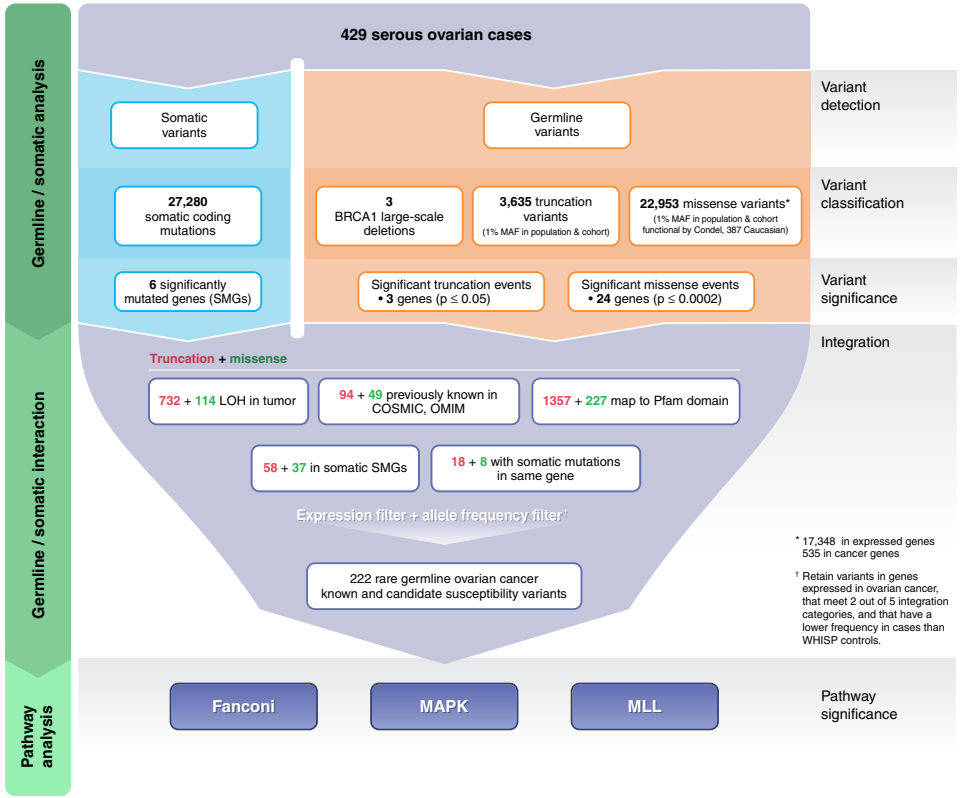
ARTICLE

Received 20 Sep 2013 | Accepted 19 Dec 2013 | Published 22 Jan 2014

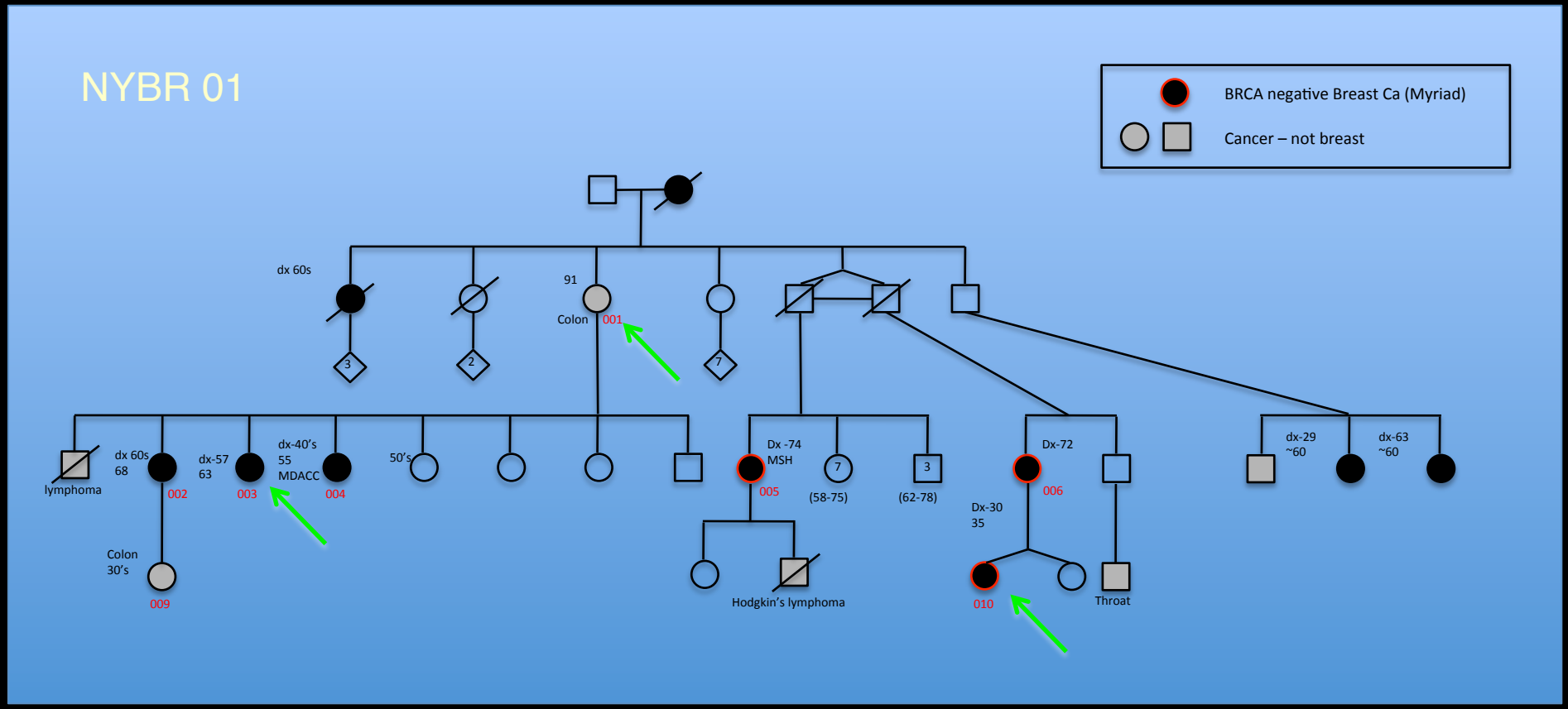
DOI: 10.1038/ncomms4156

Integrated analysis of germline and somatic variants in ovarian cancer

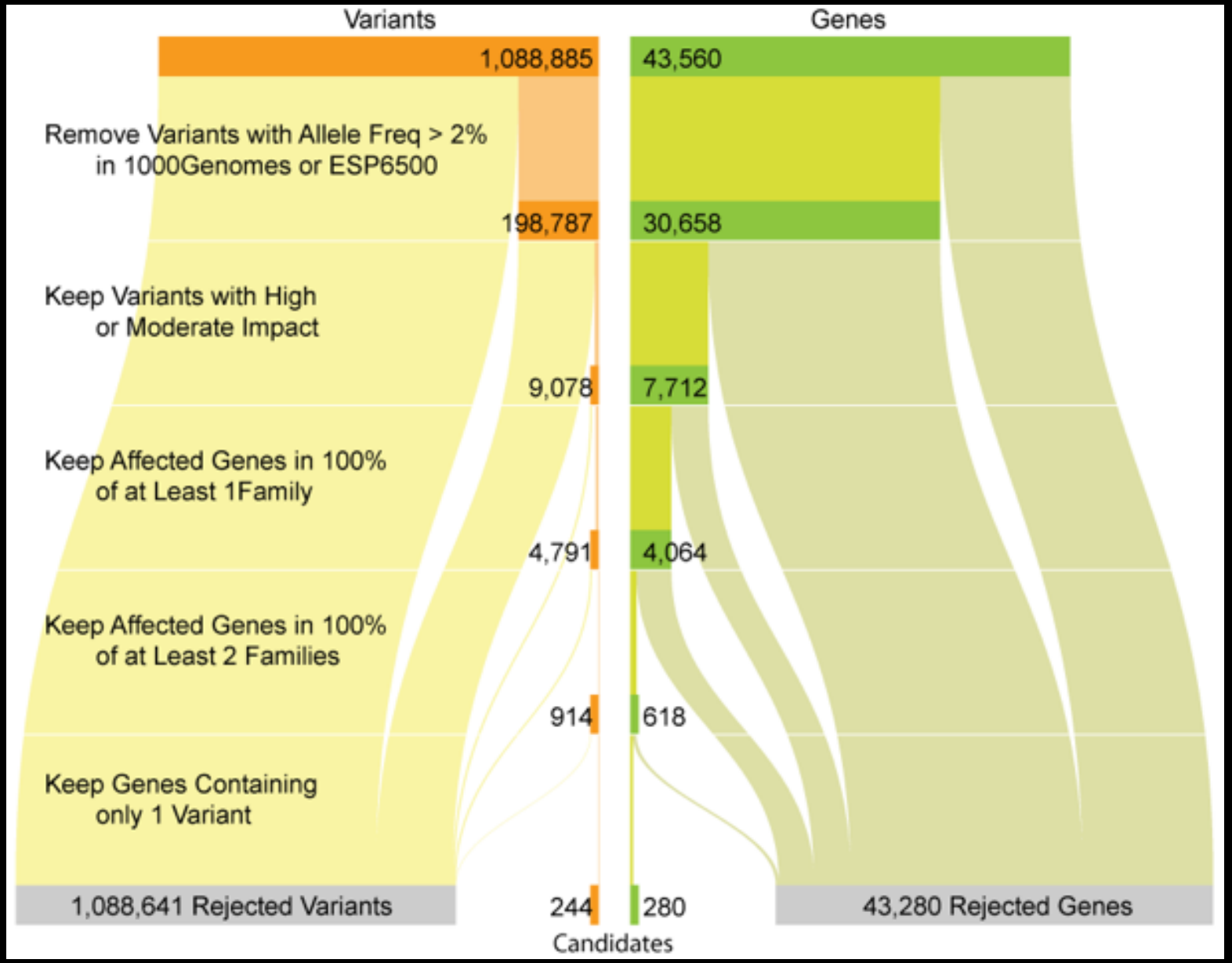
Krishna L. Kanchi[1,*], Kimberly J. Johnson[1,2,3,*], Charles Lu[1,*], Michael D. McLellan[1], Mark D.M. Leiserson[4], Michael C. Wendl[1,5,6], Qunyuan Zhang[1,5], Daniel C. Koboldt[1], Mingchao Xie[1], Cyriac Kandoth[1], Joshua F. McMichael[1], Matthew A. Wyczalkowski[1], David E. Larson[1,5], Heather K. Schmidt[1], Christopher A. Miller[1], Robert S. Fulton[1,5], Paul T. Spellman[3], Elaine R. Mardis[1,5,7], Todd E. Druley[5,8], Timothy A. Graubert[7,9], Paul J. Goodfellow[10], Benjamin J. Raphael[4], Richard K. Wilson[1,5,7] & Li Ding[1,5,7,9]



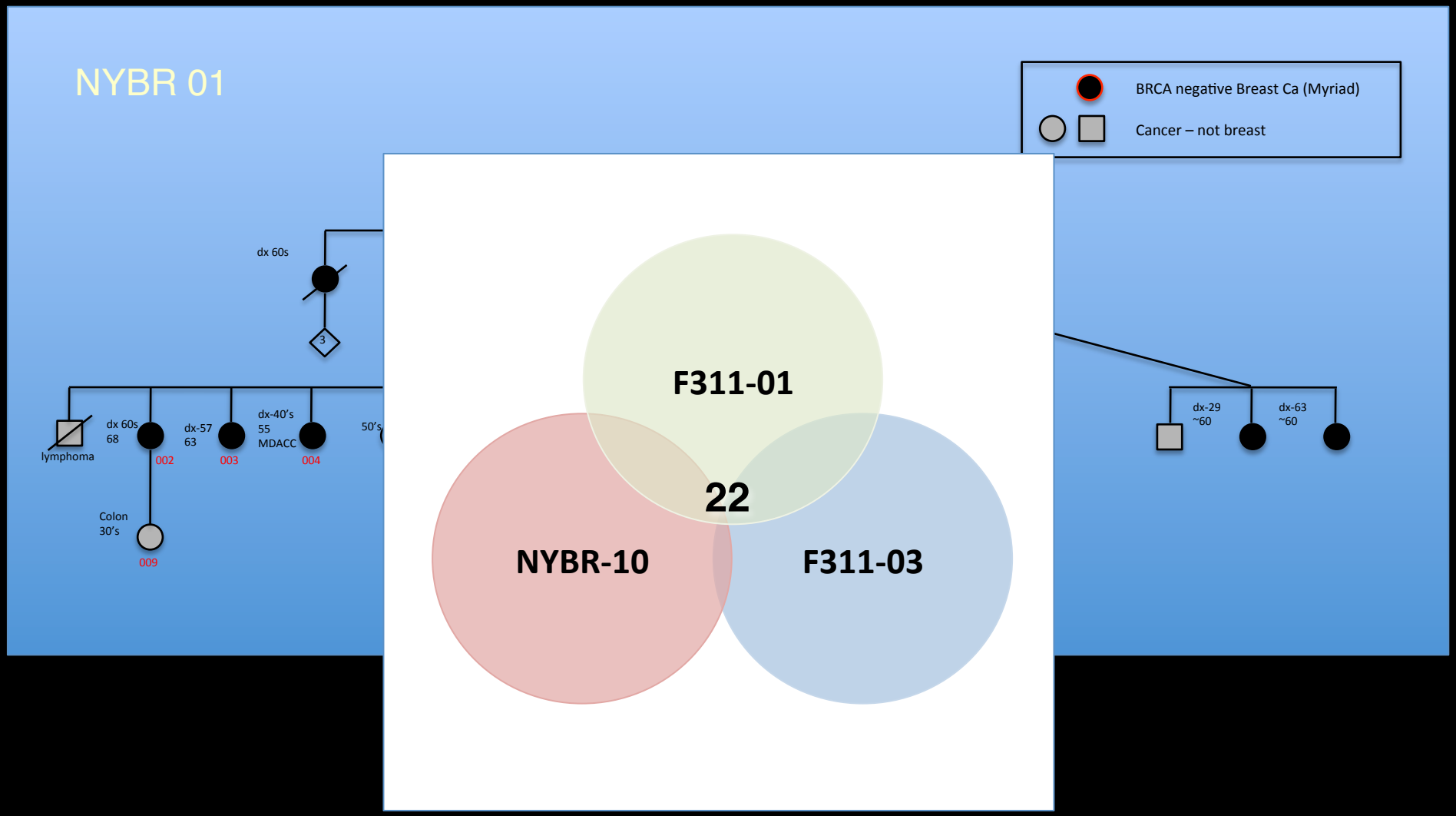
NYBR01: SEQUENCING STRATEGY



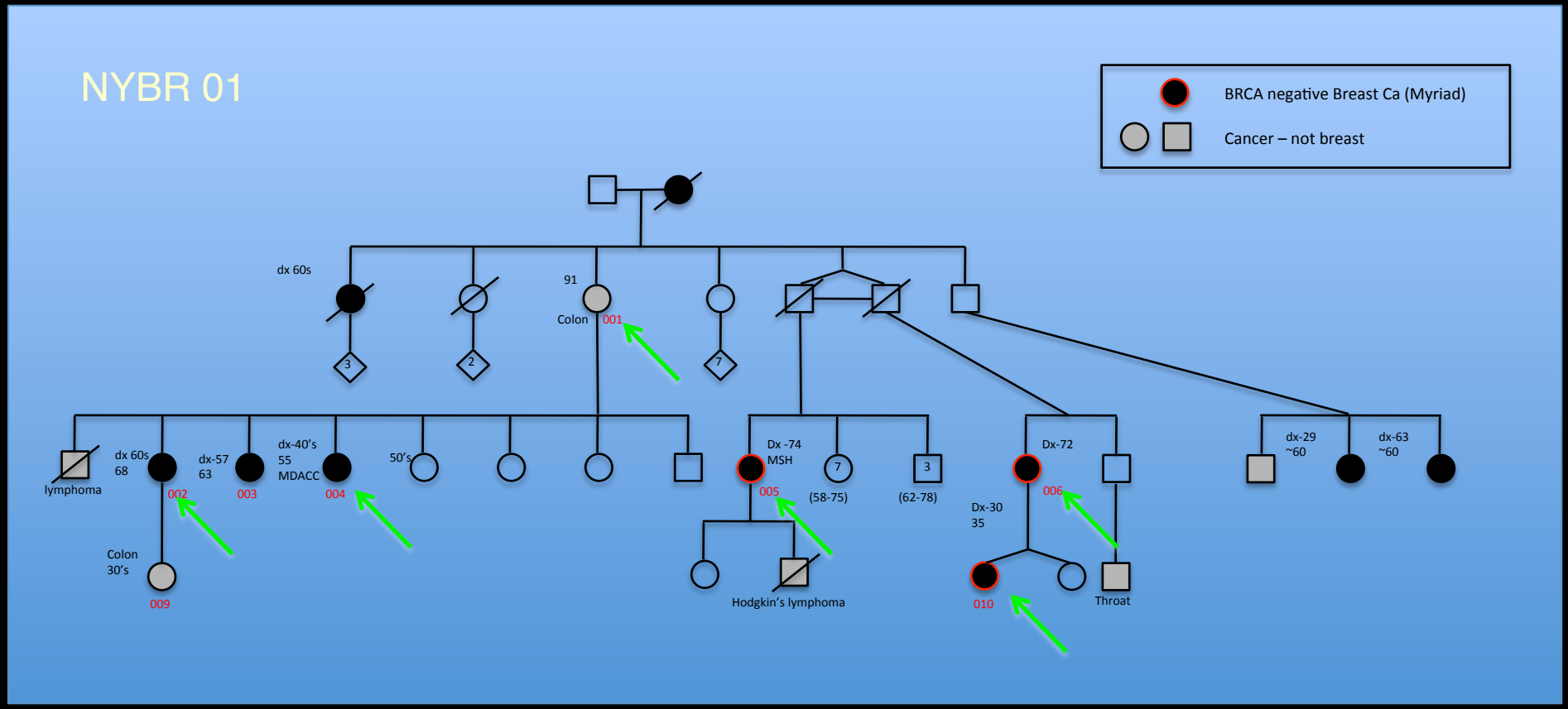
Selecting for Candidate Mutations from Germlines of Multiple Families with Breast Cancer

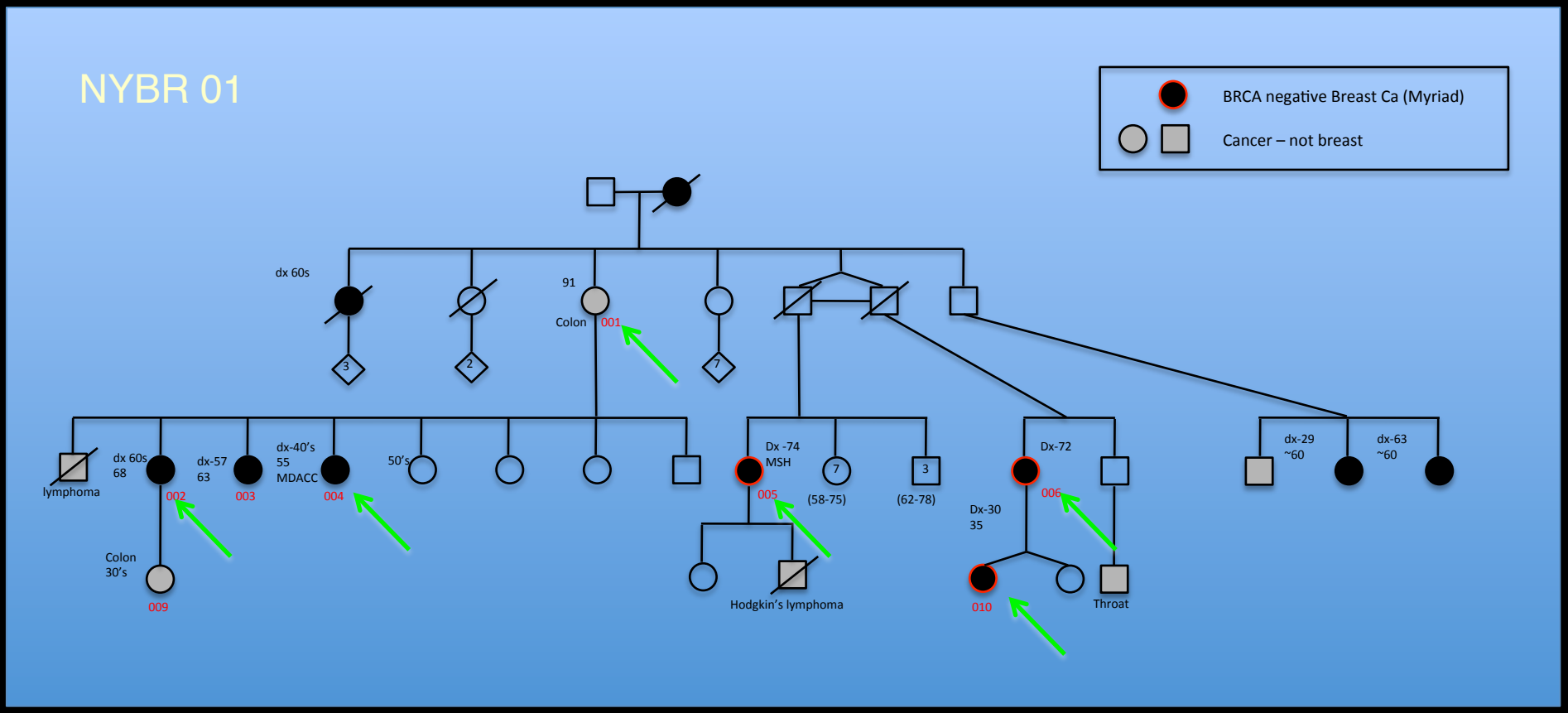


NYBR01: 22 Candidate Genes Shared by all Three Affecteds



NYBR01: Sanger Sequencing of 22 Candidate Genes in 6 Affecteds



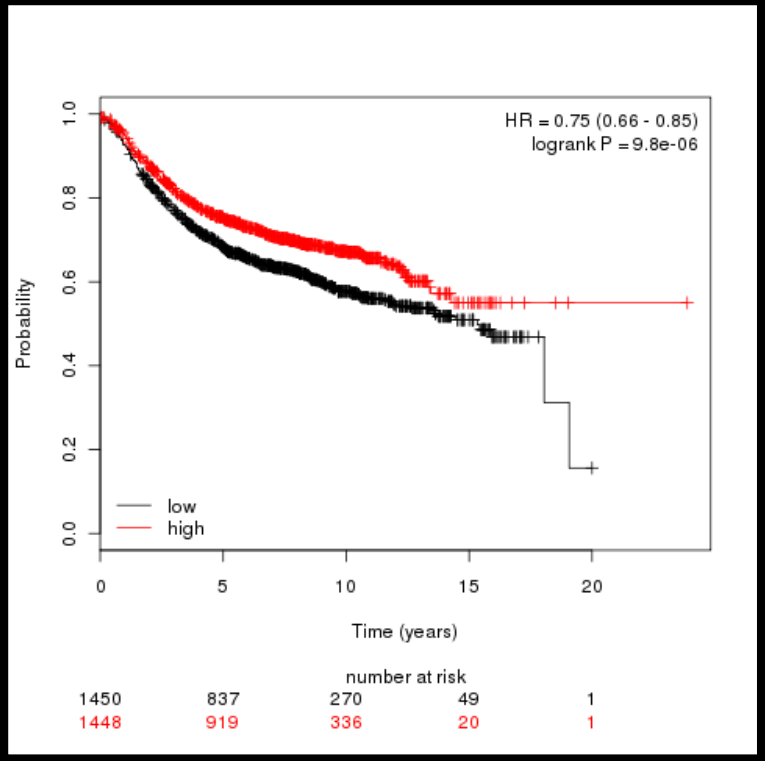


Candidate genes:

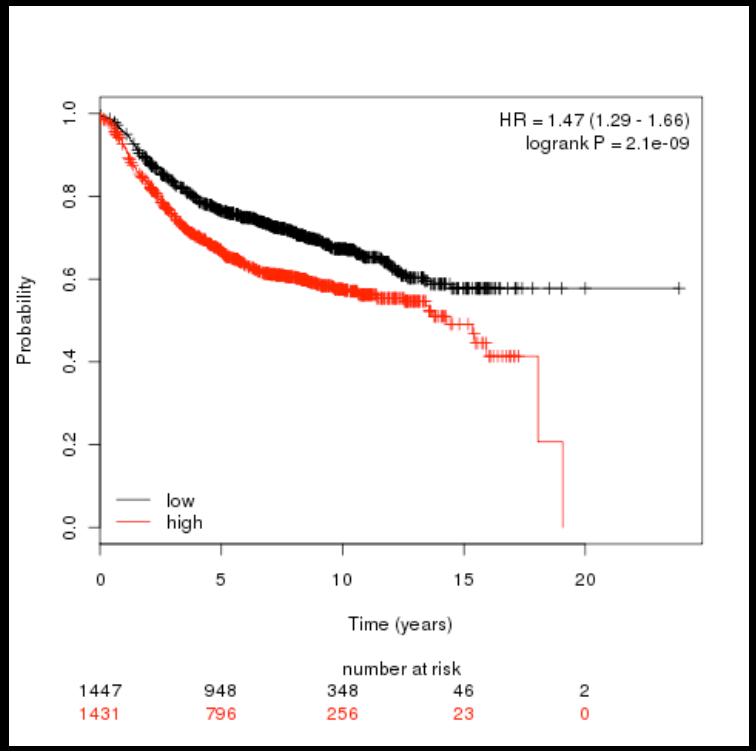
6 / 6 affecteds share 1 gene
5 / 6 affecteds share 5 genes

Identification of a Candidate Gene with Biologic Potency Affecting Long-term Survival in Breast Cancer and comparison with BRCA1

NYBR- 01 – Candidate #1



BRCA1



CONCLUSIONS

- A number of high interest “candidate” ovarian and breast cancer susceptibility mutations were discovered within (“personalized”) and between families.

- The use of germline TCGA data allowed refinement of this candidate list.

Validation studies are now planned.

- Generation of mutation-specific ovarian cancer cell lines and “humanized” mouse line for understanding the biology of OVCA#3.

- Sequencing of NYBR01 candidate in an independent breast cancer family cohort and functional studies in cultured cell lines.

Acknowledgements

- All the families that participated in these studies.

- Icahn School of Medicine at Mount Sinai (New York, NY)
  - Peter Dottino, M.D. - Sviatoslav Kendall
  - Melanie Babcock, Ph.D. - Boris Reva, PhD
  - Milind Mahajan
  - Hardik Shah

- University of Chile School of Medicine (Santiago, Chile)
  - Lilian Jara, Ph.D.

- Roswell Park Cancer Institute /
  Gilda Radner Familial Ovarian Cancer Registry (Buffalo, NY)
  - Kunle Odunsi, M.D., Ph.D.

- Station X, Inc. (San Francisco, CA) - POSTER #55
  - Sandeep Sanga, Ph.D.
  - Tod Klingler, Ph.D.
  - Richard Goold, Ph.D.
  - Antoaneta Vladimirova, Ph.D.