Prediction of breast cancer case/control status from rare CHEK2 sequence variants

Sean V. Tavtigian, PhD
Department of Oncological Sciences
Huntsman Cancer Institute
University of Utah School of Medicine

December 10, 2010

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

The University of Utah
Department of Oncological Sciences

Overview

Background

Case-control mutation screening

ATM and CHEK2

CAGI CHEK2

Topic progress

Background

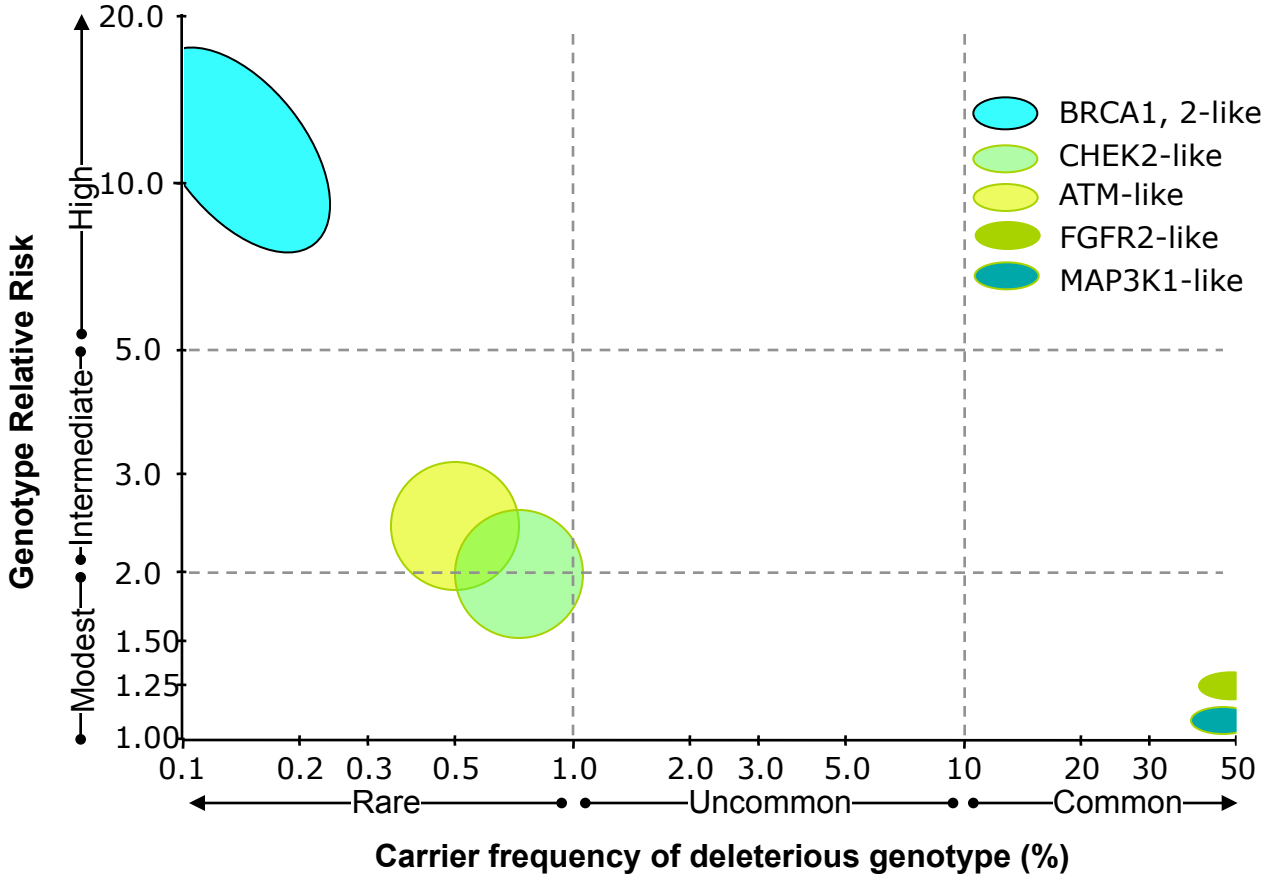
Case-control mutation screening

ATM and CHEK2

CAGI CHEK2

HUNTSMAN
CANCER INSTITUTE
UNIVERSITY OF UTAH

Gene classes: risk and frequency



HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

Topic progress

Background

Case-control mutation screening

ATM and CHEK2

CAGI CHEK2

HUNTSMAN
CANCER INSTITUTE
UNIVERSITY OF UTAH

Case-Control Mutation Screening

- Intended for identification/ validation/ characterization of intermediate-risk susceptibility genes.
- Requires epidemiologically sound case-control series
- All subjects are mutation screened across the coding exons (and proximal splice junctions) of target genes. *The lab is blind to the status of individual samples*.
- Analysis:

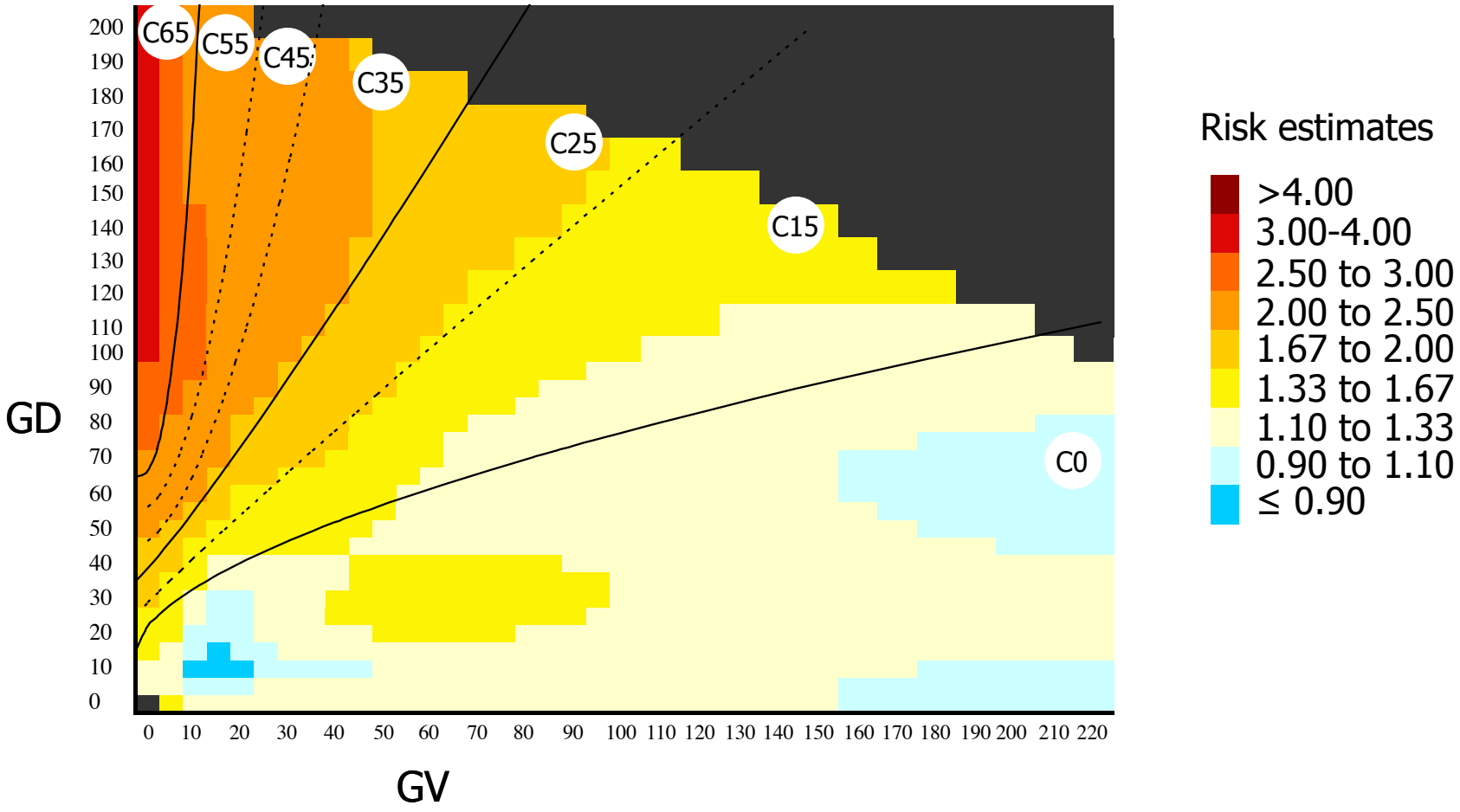
  Summed frequency of truncating variants in cases vs. controls.

  In silico grading of rare missense substitutions followed by a trend test to compare the *frequency distributions of the graded missense substitutions* in cases vs. controls.

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

Analysis of rare missense substitutions:

Distribution of risk in the GVGD plane



HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

Tavtigian et al. *Human Mutation.* **29**: 1342-1354, 2008

Topic progress

Background

Case-control mutation screening

ATM and CHEK2

CAGI CHEK2

HUNTSMAN
CANCER INSTITUTE
UNIVERSITY OF UTAH

Data collection for a meta-analysis of rare ATM variants

Study	Reference	Cases	Controls	Total	Cases selection	Geographic origin
1	Fitzgerald et al. 1997	401	202	604	Early onset	>90% Caucasian
2	Teraoka et al. 2001	142	81	223	Early onset	*Not collected*
3	Sommer et al. 2003	90	90	180	Unselected	89% Caucasian
4	Thorstenson et al. 2003	270	52	322	Family history	Austrian
5	Renwick et al. 2006	443	521	964	Family history	>97% Caucasian
6	Hirsch et al. 2007	37	95	132	Unselected	African American
7	Soukupova et al. 2008	161	183	344	Early onset + FH	Czech
8	Regensburg- kConFab	364	362	726	Family history	>95% Caucasian
9a	IARC- BCFR	392	413	805	Early onset	European
9b	IARC- East Asian	231	245	476	Early onset	Asian
	'True' case-control subtotal	**2,531**	**2,244**	**4,775**		

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

Tavtigian et al. *AJHG* **85**: 427-446, 2009

Analysis of ATM case-control mutation screening data

Rare variants only

Missense analysis limited to the FAT-kinase region

	Cases	Controls	OR*	[95% CI]	P-value
Non-carriers	2,505	2,235	1.00	[ref]	
Truncating variants	26	10	2.32	[1.12-4.83]	0.024

Tavtigian et al. *AJHG* **85**: 427-446, 2009

Towards replication of the ATM results: WECARE study

Overall analysis

ATM variant classification	Case subjects with dose estimates† (n = 606)	Control subjects with dose estimates† (n = 1200)	RR (95% CI)
All variants, broadly classified			
Wild type	223	418	1.0 (referent)
Silent	78	134	1.1 (0.8 to 1.6)
Missense	68	113	1.2 (0.8 to 1.8)
Splicing	4	14	0.7 (0.2 to 2.4)
Truncation	11	6	2.8 (0.9 to 8.9)
Common‡	308	655	0.8 (0.6 to 1.0)
Missense variants classified using SIFT§			
Wild type	223	418	1.0 (referent)
Tolerated	31	66	0.9 (0.5 to 1.5)
Deleterious	37	46	1.7 (0.9 to 2.9)

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

Bernstein et al. *JNCI* **102**: 475-483, 2010

Towards replication of the ATM results: WECARE study

Stratified by radiotherapy exposure

ATM variants	Radiation exposure, Gy†	Case subjects (n = 606)	Control subjects (n = 1200)	RR‡ (95% CI)	RR§ (95% CI)
All variants, broadly classified					
Wild type	0	112	72	1.0 (referent)	1.0 (referent)
	0.01–0.99	57	177	1.1 (0.7 to 1.6)	1.1 (0.7 to 1.6)
	≥1.0	54	169	1.1 (0.7 to 1.7)	1.1 (0.7 to 1.7)
Silent	0	38	29	1.1 (0.6 to 2.0)	1.0 (referent)
	0.01–0.99	25	59	1.3 (0.7 to 2.2)	1.1 (0.5 to 2.4)
	≥1.0	15	46	1.0 (0.5 to 2.0)	0.9 (0.4 to 2.2)
Splicing	0	0	2	—	1.0 (referent)
	0.01–0.99	3	6	1.5 (0.4 to 6.5)	—
	≥1.0	1	6	0.4 (0.0 to 3.6)	—
Truncation	0	6	3	1.6 (0.3 to 8.6)	1.0 (referent)
	0.01–0.99	3	3	2.9 (0.5 to 16.3)	1.7 (0.2 to 19.3)
	≥1.0	2	0	—	—
Rare missense, stratified by SIFT					
Tolerated	0	12	16	0.7 (0.3 to 1.7)	1.0 (referent)
	0.01–0.99	9	27	1.1 (0.4 to 2.7)	1.6 (0.5 to 5.2)
	≥1.0	10	23	1.3 (0.6 to 3.2)	1.8 (0.6 to 5.8)
Deleterious	0	14	14	0.6 (0.2 to 1.3)	1.0 (referent)
	0.01–0.99	12	17	2.8 (1.2 to 6.5)	5.3 (1.6 to 17.3)
	≥1.0	11	15	3.3 (1.4 to 8.0)	5.8 (1.8 to 19.0)

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

Bernstein et al. *JNCI* **102**: 475-483, 2010

Case-control mutation screening of CHEK2

Study characteristics

A. By Age

Age	Cases (%)		Controls (%)	
≤30	106	(8.1%)	66	(6.0%)
31-35	322	(24.7%)	171	(15.4%)
36-40	434	(33.3%)	231	(20.8%)
41-45	441	(33.8%)	199	(17.9%)
46-50	0	(0.0%)	230	(20.7%)
51-55	0	(0.0%)	212	(19.1%)
TOTAL	1,303	(100.0%)	1,109	(100.0%)

B. By race/ ethnicity

Race/ Ethnic group	Cases (%)		Controls (%)	
Caucasian	843	(64.7%)	956	(86.2%)
East Asian	204	(15.7%)	70	(6.3%)
Latina	158	(12.1%)	47	(4.2%)
Recent African Ancestry	98	(7.5%)	36	(3.2%)
TOTAL	1,303	(100.0%)	1,109	(100.0%)

HUNTSMAN CANCER INSTITUTE
UNIVERSITY OF UTAH

Case-control mutation screening of CHEK2

Results

	Cases	Controls	OR*	[95% CI]	P-value
Non-carriers	1,242	1,089	1.00	[ref]	
Any truncating variant	17	3	6.18	[1.76-21.8]	0.005
Any rare missense	44	17	2.20	[1.20-4.01]	0.010

*Adjusted for study center, age, and ethnicity

HUNTSMAN
CANCER INSTITUTE
UNIVERSITY OF UTAH

Case-control mutation screening of CHEK2

Results

	Cases	Controls	OR*	[95% CI]	P-value
Non-carriers	1,242	1,089	1.00	[ref]	
Any truncating variant	17	3	6.18	[1.76-21.8]	0.005
Any rare missense	44	17	2.20	[1.20-4.01]	0.010
Stratified rare missense					
C0	12	9	1.39	[0.55-3.56]	
C15	14	5	1.82	[0.62-5.34]	
C25	7	2	2.47	[0.45-13.5]	
C35	1	0	--	--	
C45	0	0	--	--	
C55	1	0	--	--	
C65	9	1	8.75	[1.06-72.2]	
					0.0055

*Adjusted for study center, age, and ethnicity

HUNTSMAN CANCER INSTITUTE
UNIVERSITY OF UTAH

Topic progress

Background

Case-control mutation screening

ATM and CHEK2

CAGI CHEK2

HUNTSMAN
CANCER INSTITUTE
UNIVERSITY OF UTAH

CHEK2 Prediction: Truncating variant data

Nucleotide	Amino acid	Cases	Controls
c.283C>T	p.R95X	1	0
c.405delA	p.K135fs	1	0
c.823G>T	p.E275X	1	0
c.1100delC	p.T367fs	13	3
c.1138delCT	p.L380fs	1	0
c.1263delT	p.L421fs	1	0
c.1528C>T	p.Q510X	1	0

From a total of 16 predictions:
Only 2 predictions included all of the truncating variants.
One additional prediction included the nonsense substitutions.

It's appropriate to include all of them as a matter of course.

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

CHEK2 Prediction: Missense substitution data

Number of variants called

All	2
Except T+SJVs	3
& except doubles	8
Fewer	2

Nucleotide	Amino acid	Cases	Controls	Race/ ethnicity
c.14C>T	p.S5L	2	1	
c.74T>C	p.V25A	1	0	
c.254C>T	p.P85L	1	3	
c.349A>G	p.R117G	3	1	1/4 Latina
c.410G>A	p.R137Q	1	0	
c.470T>C	p.I157T	2	1	
c.538C>T	p.R180C	3	0	East Asian, Af Am
c.539G>A	p.R180H	0	1	
c.575C>T	p.S192L	1	0	East Asian
c.663C>G	p.I221M	1	0	Latina
c.688G>T	p.A230S	0	1	
c.715G>A	p.E239K	2	0	
c.727T>C	p.C243R	0	1	
c.751A>T	p.I251F	0	1	
c.911T>C	p.M304T	1	0	
c.917G>C	p.G306A	1	0	Latina
c.931G>A	p.D311N	1	0	
c.967A>C	p.T323P	1	0	
c.1036C>T	p.R346C	3	0	1/3 East Asian
c.1054A>T	p.N352Y	1	0	
c.1111C>T	p.H371Y	2	1	East Asian
c.1216C>T	p.R406C	1	0	
c.1253T>G	p.F418C	1	0	
c.1276C>T	p.P426S	1	0	
c.1312G>T	p.D438Y	2	2	
c.1313A>G	p.D438G	1	0	
c.1336A>G	p.N446D	0	1	
c.1343T>G	p.I448S	7	2	Af Am
c.1427C>T	p.T476M	1	0	Latina
c.1451C>T	p.P484L	1	0	Latina
c.1534C>G	p.L512V	1	0	Latina
c.1556G>T	p.R519L	0	1	
c.715G>A + c.1037G>A	p.E239K + p.R346H	1	0	
c.1182A>T + c.1343T>G	p.E394D + p.I448S	1	0	Af Am

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

CHEK2 Prediction Results: Analysis

- Odds ratio:

  Predicted most pathogenic 3rd of missense substitutions vs most benign 3rd of missense substitutions.

OR ≤ 1.0	4 methods
1.0 < OR ≤ 2.0	2 methods
2.0 < OR < 4.5	5 methods
OR = 4.5	Align-GVGD
OR > 4.5	4 methods

- Linear regression:

  Regress predicted probability pathogenic for each variant against case/ control status of each subject. Report 1-sided P-value.

P = 1.0	4 methods
1.0 > P > 0.05	6 methods
0.05 > P > 0.02	2 methods
P = 0.021	Align-GVGD
P < 0.02	4 methods

CHEK2 Prediction Results: Correlation

- Five methods beat Align-GVGD in either the OR or LR tests. Here are their pairwise correlations.

	AGVGD	Method 1	Method2	Method3	Method 4	Method 5
Method 1	0.49					
Method 2	0.75	0.51				
Method 3	0.66	0.73	0.80			
Method 4	0.58	0.65	0.76	0.92		
Method 5	0.54	0.65	0.76	0.76	0.80	
Method 6	0.42	0.61	0.58	0.80	0.69	0.74

				AGVGD	M6	M1	M2	M3	M4	M5
HGVS_aa	case	con		0_tert	4_Tert	5_Tert	6_Tert	8_Tert	11_Tert	15_Tert
p.R137Q	1	0		0	0	1	0	0	0	2
p.R180H	0	1		0	2	1	0	1	0	1
p.S192L	1	0		1	1	2	0	1	0	1
p.C243R	0	1		0	1	2	0	2	1	1
p.G306A	1	0		0	2	2	2	2	2	2
p.T323P	1	0		0	2	1	2	2	2	2
p.D438Y	2	2		2	1	0	2	1	1	0
Ave tertile				0.43	1.29	1.29	0.86	1.29	0.86	1.29

CHEK2 Prediction Results: Interesting variants

- Maybe relatively pathogenic, concordant prediction:

HGVS_nt	HGVS_aa	case	con		Case>2x Con Severe 1/3	Case>2x Con benign 1/3
c.349A>G	p.R117G	3	1	P&con	12	1
c.917G>C	p.G306A	1	0	P&con	10	2
c.1054A>T	p.N352Y	1	0	P&con	15	1
c.1253T>G	p.F418C	1	0	P&con	13	1
c.1276C>T	p.P426S	1	0	P&con	13	1

- Maybe relatively pathogenic, lots of discord:

c.74T>C	p.V25A	1	0	P&dis	1	12
c.663C>G	p.I221M	1	0	P&dis	2	12
c.931G>A	p.D311N	1	0	P&dis	2	6
c.1534C>G	p.L512V	1	0	P&dis	1	11

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

CHEK2 Prediction Results: Interesting variants

- Maybe relatively neutral, concordant prediction:

					Con>Case Severe 1/3	Con>Case benign 1/3
HGVS_nt	HGVS_aa	case	con			
c.254C>T	p.P85L	1	3	N&con	2	7
c.539G>A	p.R180H	0	1	N&con	2	7
c.688G>T	p.A230S	0	1	N&con	0	7
c.1336A>G	p.N446D	0	1	N&con	1	15
c.1556G>T	p.R519L	0	1	N&con	2	7

- Maybe relatively neutral, lots of discord:

c.751A>T	p.I251F	0	1	N&dis	10	1

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

CHEK2 Prediction Results: Interesting variants

- Massive confusion:

HGVS_nt	HGVS_aa	case	con		Case>2x Con Severe 1/3	Con>Case Severe 1/3	Case>2x Con benign 1/3	Con>Case benign 1/3
c.715G>A	p.E239K	2	0	discord	5		7	
c.727T>C	p.C243R	0	1	discord		7		4
c.1427C>T	p.T476M	1	0	discord	6		4	
c.1451C>T	p.P484L	1	0	discord	7		4	

- Power calculation (from Align-GVGD)

Study scale	Single genes	Whole pathways*	Whole exome†
Alpha	0.05	0.0005	2.5×10^{-6}
Beta	0.80	0.80	0.80
rMSs alone	1,975	4,700	7,725
T+SJVs alone	1,425	3,400	5,600
rMSs plus T+SJVs	850	2,025	3,350

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH

Thanks to…

International Agency for Research on Cancer (IARC)

- Florence Le Calvez-Kelm
- Fabienne Lesueur
- Graham Byrnes
- Francesca Damiola
- Geoffroy Durand
- Nathalie Forey
- Sandrine McKay-Chopin
- Tu Nguyen-Dumont
- Nivonirina Robinot
- Maxime Vallée
- Catherine Voegele

Breast CFR PIs

- Irene Andrulis
- Esther John
- John Hopper
- Melissa Southey

University of Utah

- Ken Boucher
- Mia Hashibe
- Alun Thomas

HUNTSMAN CANCER INSTITUTE UNIVERSITY OF UTAH