US 20230250483A1

(19) **United States**

(12) **Patent Application Publication**

Kinde et al.

(10) **Pub. No.: US 2023/0250483 A1**

(43) **Pub. Date: Aug. 10, 2023**

(54) **PAPANICOLAOU TEST FOR OVARIAN AND ENDOMETRIAL CANCERS**

(71) Applicant: **The Johns Hopkins University**, Baltimore, MD (US)

(72) Inventors: **Isaac Kinde**, Beaumont, CA (US); **Kenneth W. Kinzler**, Frankford, DE (US); **Bert Vogelstein**, Baltimore, MD (US); **Nickolas Papadopoulos**, Towson, MD (US); **Luis Diaz**, Ellicott City, MD (US); **Chetan Bettegowda**, Lutherville, MD (US); **Yuxuan Wang**, Baltimore, MD (US)

(21) Appl. No.: **18/050,809**

(22) Filed: **Oct. 28, 2022**

Related U.S. Application Data

(63) Continuation of application No. 14/439,041, filed on Apr. 28, 2015, now Pat. No. 11,525,163, filed as application No. PCT/US13/65342 on Oct. 17, 2013.

(60) Provisional application No. 61/719,942, filed on Oct. 29, 2012.

Publication Classification

(51) **Int. Cl.**

C12Q 1/6886	(2006.01)
G01N 33/574	(2006.01)

(52) **U.S. Cl.**

CPC ***C12Q 1/6886*** (2013.01); ***G01N 33/57442*** (2013.01); ***G01N 33/57449*** (2013.01); *C12Q 2600/16* (2013.01); *C12Q 2600/154* (2013.01); *C12Q 1/6851* (2013.01)

(57) **ABSTRACT**

The recently developed liquid-based Papanicolaou (Pap) smear allows not only cytologic evaluation but also collection of DNA for detection of HPV, the causative agent of cervical cancer. We tested these samples to detect somatic mutations present in rare tumor cells that might accumulate in the cervix once shed from endometrial and ovarian cancers. A panel of commonly mutated genes in endometrial and ovarian cancers was assembled and used to identify mutations in all 46 endometrial or cervical cancer tissue samples. We were able also able to identify the same mutations in the DNA from liquid Pap smears in 100% of endometrial cancers (24 of 24) and in 41% of ovarian cancers (9 of 22). We developed a sequence-based method to query mutations in 12 genes in a single liquid Pap smear without prior knowledge of the tumor's genotype.

Specification includes a Sequence Listing.

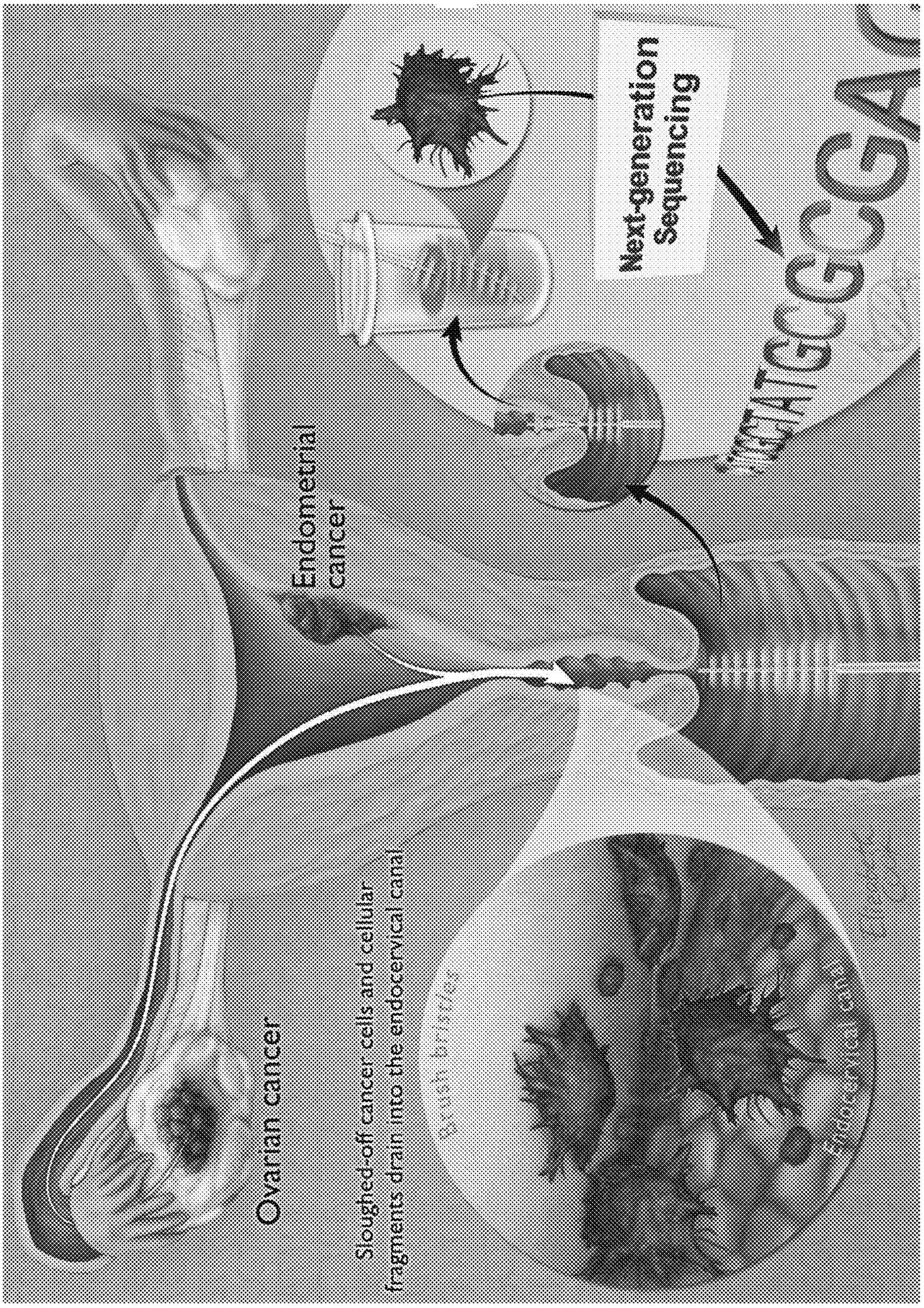
Next-generation Sequencing
TCGA
Endometrial cancer
Ovarian cancer
Sloughed-off cancer cells and cellular fragments drain into the endocervical canal

Fig. 1

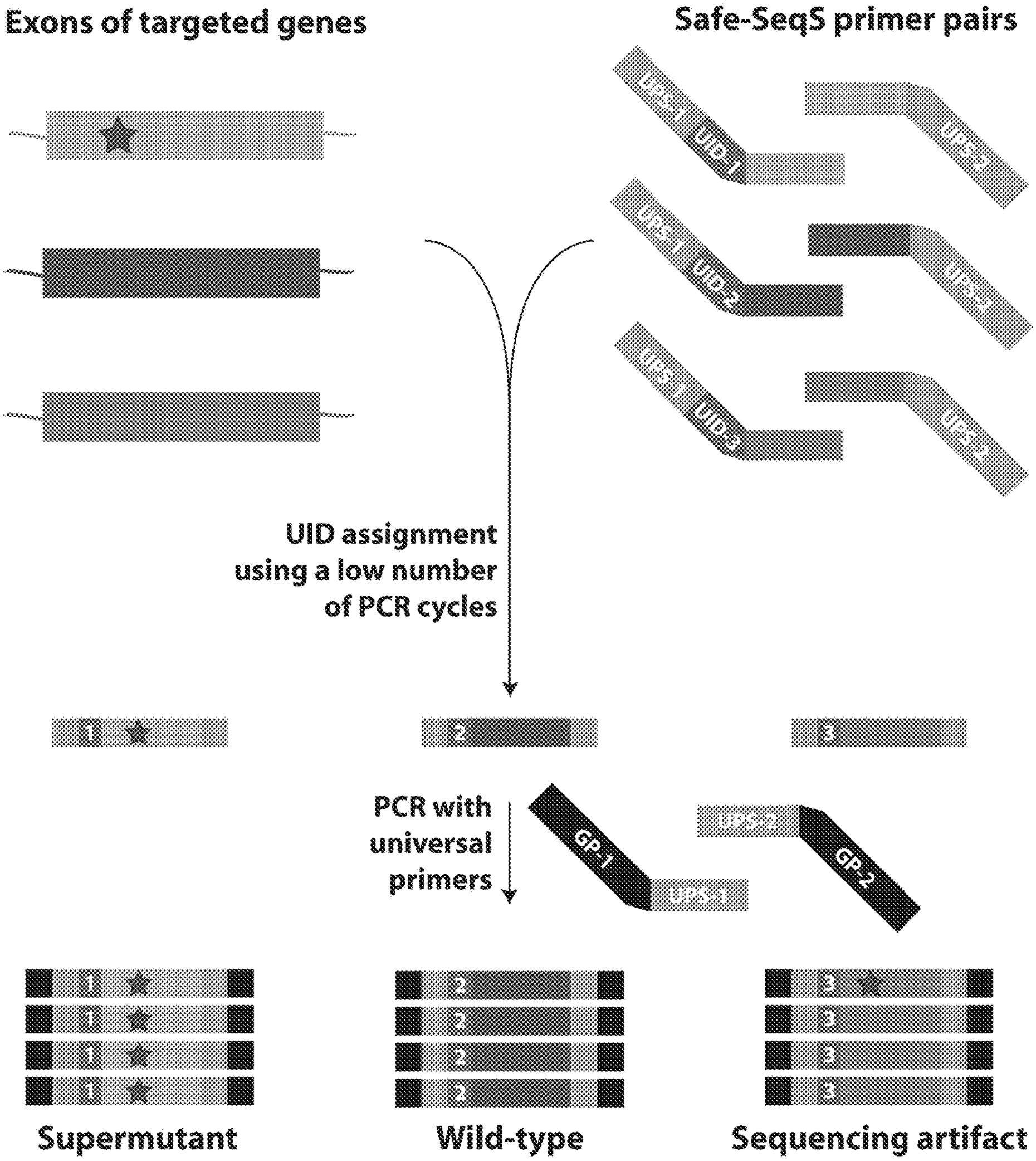
Exons of targeted genes
Safe-SeqS primer pairs
UID-1
UP-S-1
UP-S-2
UID-2
UID-3
UID assignment using a low number of PCR cycles
1
2
3
PCR with universal primers
GP-1
GP-2
Supermutant
Wild-type
Sequencing artifact
Next-generation sequencing

Fig. 2

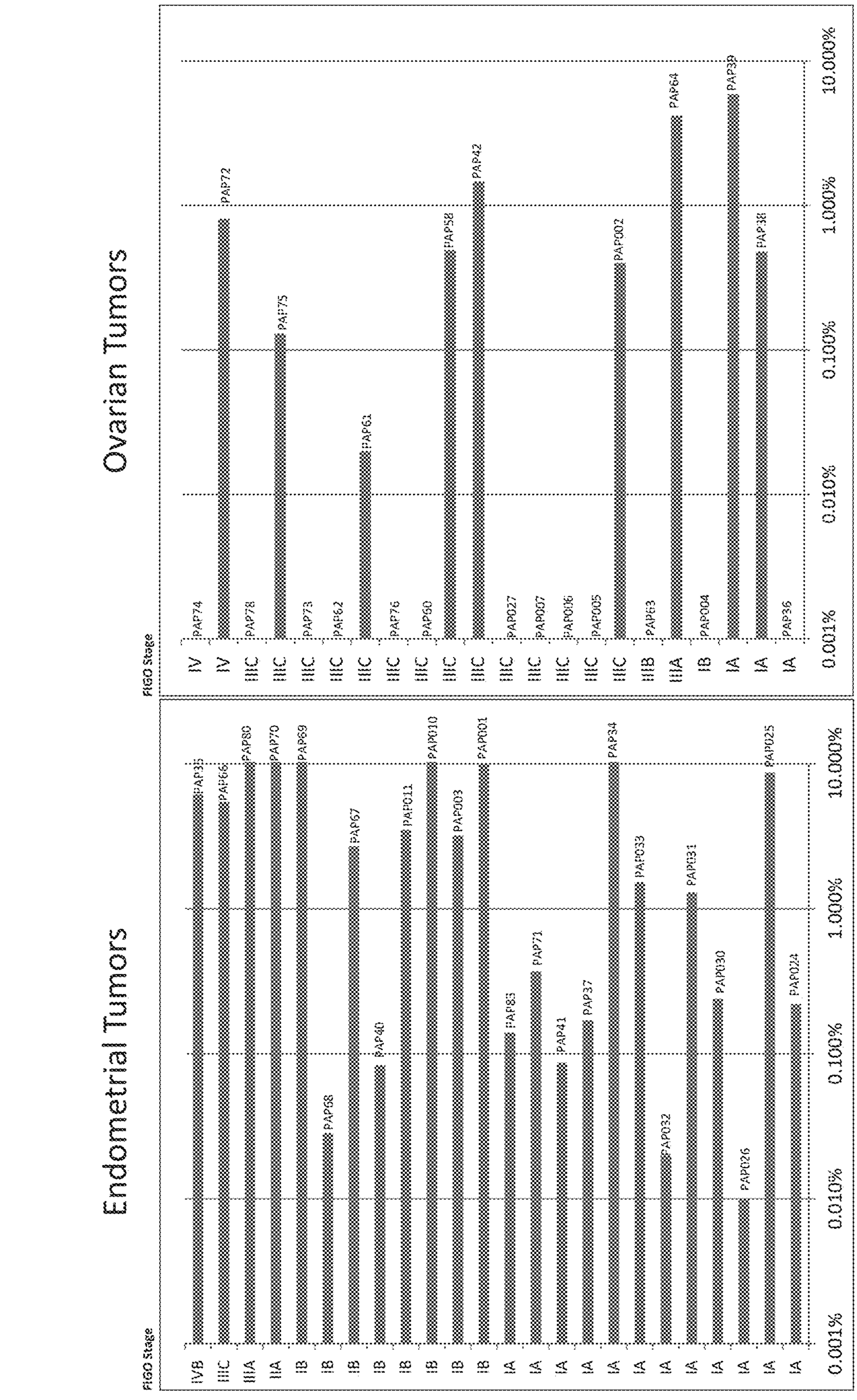
Ovarian Tumors
FIGO Stage
IV PAP74
IV PAP78
IIIC PAP73
IIIC PAP62
IIIC PAP61
IIIC PAP76
IIIC PAP60
IIIC PAP027
IIIC PAP007
IIIC PAP006
IIIC PAP005
IIIB PAP63
IIIA PAP004
IB PAP036
IA PAP035
IA PAP066
IA PAP080
IA PAP070
IA PAP069
IA PAP067
IA PAP011
IA PAP010
IA PAP003
IA PAP001
IA PAP034
IA PAP033
IA PAP031
IA PAP025
IA PAP036
0.001% 0.010% 0.100% 1.000% 10.000%
Endometrial Tumors
FIGO Stage
IVB PAP72
IIIC PAP75
IIIA PAP58
IIA PAP42
IB PAP002
IB PAP64
IB PAP39
IB PAP38
IB PAP68
IB PAP40
IB PAP83
IB PAP71
IB PAP41
IB PAP37
IA PAP032
IA PAP026
IA PAP030
IA PAP024
0.001% 0.010% 0.100% 1.000%
% mutant alleles in liquid Pap smear sample

Fig. 3

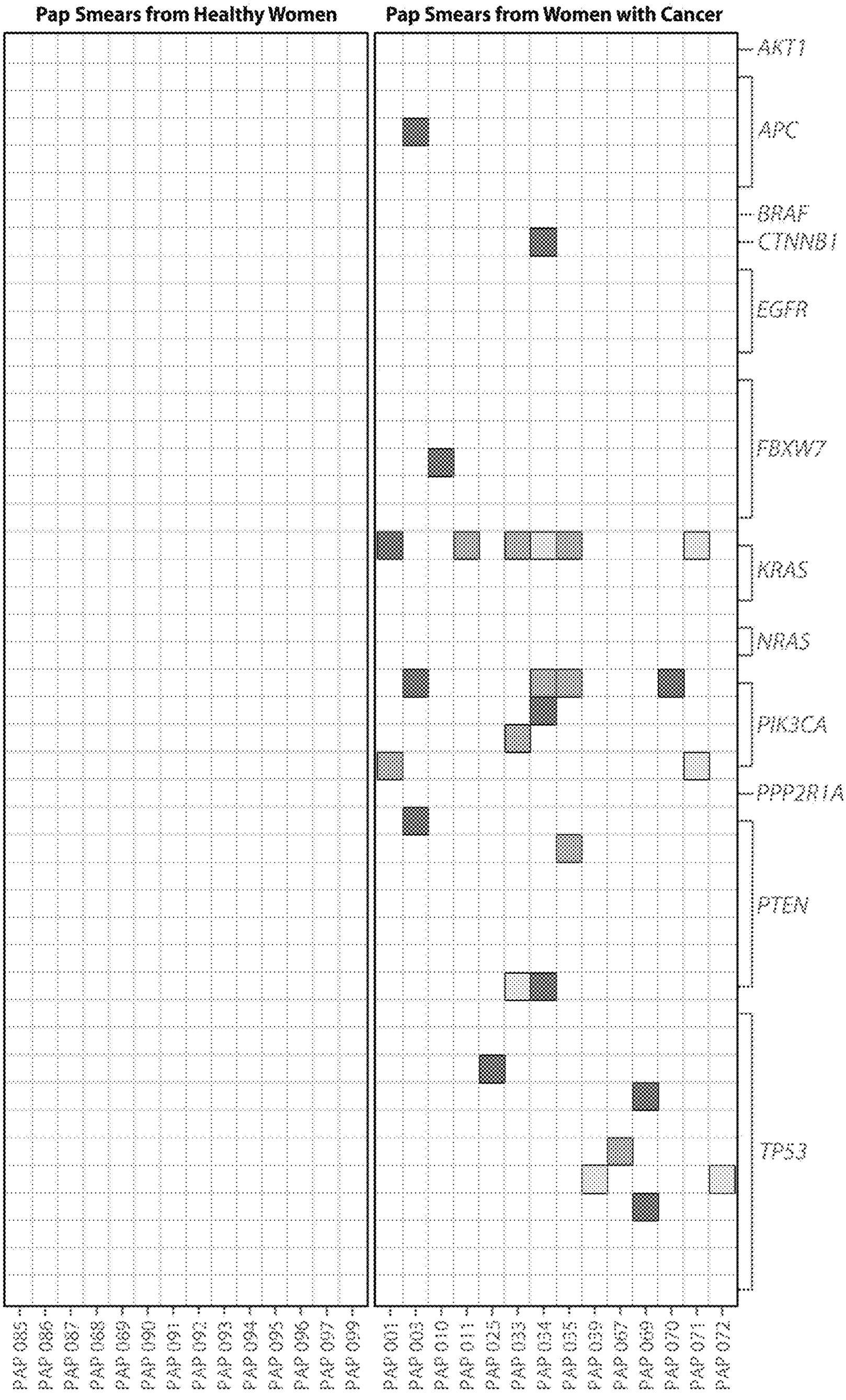
Pap Smears from Healthy Women
Pap Smears from Women with Cancer
AKT1
APC
BRAF
CTNNB1
EGFR
FBXW7
KRAS
NRAS
PIK3CA
PPP2R1A
PTEN
TP53
PAP 085
PAP 086
PAP 087
PAP 088
PAP 089
PAP 060
PAP 061
PAP 062
PAP 063
PAP 064
PAP 065
PAP 066
PAP 067
PAP 068
PAP 001
PAP 008
PAP 010
PAP 011
PAP 025
PAP 033
PAP 034
PAP 035
PAP 036
PAP 037
PAP 069
PAP 070
PAP 071
PAP 072

Fig. 4

Table 1 - Epidemiology of Ovarian and Endometrial Tumors

Tissue	Type	Subtype	Fraction of total	Estimated New Cases in U.S., 2012	Estimated Deaths in U.S., 2012	5-year Survival	Reference No.
Ovarian	Epithelial	High-grade Serous	60%	13,368	9,224	9%	43, 44
		Endometrioid	15%	3,342	2,373	71%	43, 45
		Clear Cell	10%	2,228	1,381	62%	43, 45
		Low-grade Serous	8%	1,782	1,105	40%	43, 44
		Mucinous	2%	446	348	65%	43, 45
		Other	5%	1,114	722	N/A	43, 44, 45
Endometrial	Type I: Endometrioid	Endometrioid	85%	40,060	5,180	91%	43, 45
	Type II: Non-Endometrioid	Papillary Serous	10%	4,713	2,194	45%	43, 45
		Clear Cell	5%	2,357	650	68%	43, 46

Fig. 5

Table 2 - Genetic Characteristics of Ovarian and Endometrial Cancers

Tissue	Type	Subtype	Common Mutations (Frequency)	Reference No.
Ovarian	Epithelial	High-grade Serous	TP53 (96%)	23
		Endometrioid	TP53 (68%)	26
			ARID1A (30%)	26
			CTNNB1 (26%)	26
			PTEN (17%)	26
			PIK3CA (15%)	26
			KRAS (10%)	26
			PPP2R1A (11%)	26
			CDKN2A (12%)	26
			BRAF (8%)	26
		Clear Cell	ARID1A (57%)	24
			PIK3CA (40%)	24
			PPP2R1A (7.1%)	24
			KRAS (4.7%)	24
		Low-grade Serous	BRAF (38%)	25
			KRAS (19%)	25
		Mucinous	TP53 (56%)	26
			KRAS (40%)	26
			PPP2R1A (33%)	26
			CDKN2A (16%)	26
			PTEN (11%)	26

Tissue	Type	Subtype	Common Mutations (Frequency)	Reference No.
Endometrial	Type I: Endometriod	Endometrioid	PTEN (64%)	Current study
			PIK3CA (59%)	Current study
			ARID1A (55%)	Current study
			CTNNB1 (32%)	Current study
			MLL2 (32%)	Current study
			FBXW7 (27%)	Current study
			RNF43 (27%)	Current study
			APC (23%)	Current study
			FGFR2 (18%)	Current study
			KRAS (9%)	Current study
			PIK3R1 (9%)	Current study
			EGFR (14%)	Current study
			AKT1 (5%)	Current study
			NRAS (5%)	Current study
			TP53 (5%)	Current study
	Type II: Non-Endometrioid	Papillary serous	TP53 (81.6%)	29
			PIK3CA (24%)	29
			FBXW7 (19.7%)	29
			PPP2R1A (18.4%)	29
		Clear Cell	TP53 (45%)	26
			PPP2R1A (33%)	26
			PIK3CA (29%)	26
			PTEN (13%)	26
			PIK3R1 (9%)	26
			KRAS (5%)	26

Fig. 6

Table S1: Endometrial Cancers (Endometrioid Subtype) Studied by Whole-exome Sequencing

Tumor ID	Age	Stage (FIGO)	Pathologic Stage (TNM class)	Number of Mutations	Microsatellite Stability Status*
PAP 003	53	IB	T1bN0M0	847	MSS
PAP 010	73	IB	T1bN0M1	29	MSS
PAP 011	58	IB	T1bN0M2	579	MSI-H
PAP 024	56	IA	T1aN0	7	MSS
PAP 026	86	IA	T1a	769	MSI-H
PAP 030	73	IA	T1aNx	49	MSS
PAP 031	61	IA	T1aNx	41	MSS
PAP 032	82	IA	T1aNx	9	MSS
PAP 033	68	IA	T1aNX	34	MSS
PAP 034	55	IA	T1aN0	454	MSI-H
PAP 043	55	IB	T1BN0MX	26	MSS
PAP 045	57	IB	T1BN0MX	4629	MSS
PAP 046	44	IIA	T2ANXMX	40	MSS
PAP 047	53	IA	T1AN0MX	1767	MSS
PAP 048	62	IIIC	T2AN1MX	394	MSI-H
PAP 049	45	IIB	T2BN0MX	20	MSS
PAP 050	39	IB	T1BN0MX	50	MSS
PAP 052	70	IVB	T1AN1M1	164	MSI-H
PAP 053	66	IB	T1BN0MX	1102	MSI-H
PAP 054	73	IA	T1ANXMX	413	MSI-H
PAP 055	61	IA	T1AN0MX	1195	MSS
PAP 057	59	IIB	T2BN0MX	176	MSI-H

* MSI-H: micosatellite unstable; MSS: microsatellite stable. See Materials and Methods

Fig. 7

Table S3. Mutations Assessed in Pap Smears

Case #	Age	Tissue	Subtype	Clinical Stage (FIGO)	Race/ Ethnicity	DNA recovered from Pap smear fluid (ug)	Mutated Gene Name	Mutated Gene ID	Nucleotide (genomic)*	Transcript	Nucleotide (transcript)	Amino Acid (protein)	Mutation Type	Fraction of mutant alleles in Pap smear fluid
PAP 001	45	Endometrial	Endometrioid	IB	White	30.2	PIK3CA	ENSG00000121879	g.chr3:180418479A>G	CCDS43171.1	c.3140A>G	p.K1047R	Missense	10.00%
PAP 002	39	Ovarian	Papillary serous	IIIC	White	3.65	TP53	ENSG00000141510	g.chr17:7519215A>T	CCDS11118.1	c.440T>A	p.V147D	Missense	0.40%
PAP 003	53	Endometrial	Endometrioid	IB	White	5.2	APC	ENSG00000134982	g.chr5:112809538C>T	CCDS4107.1	c.4348C>T	p.R1450X	Missense	3.20%
PAP 004	38	Ovarian	Papillary serous	IB	White	0.36	TP53	ENSG00000141510	g.chr17:7519154delC	CCDS11118.1	c.501delG	Frameshift	Indel	<0.005%
PAP 005	59	Ovarian	Borderline papillary serous cystadenocarcinoma	IIIC	White	7.6	NF1	ENSG00000196712	g.chr17:26552610insC	CCDS42292.1	c.1241insC	Frameshift	Indel	<0.005%
PAP 006	57	Ovarian	Papillary serous	IIIC	White	8.2	TP53	ENSG00000141510	g.chr17:7517824C>T	CCDS11118.1	c.839G>A	p.R280K	Missense	<0.005%
PAP 007	63	Ovarian	Papillary serous	IIIC	White	6.8	TP53	ENSG00000141510	g.chr17:7517783C>A	CCDS11118.1	c.859G>T	p.E286X	Nonsense	<0.005%
PAP 010	73	Endometrial	Endometrioid	IB	White	8.6	FBXW7	ENSG00000109670	g.chr4:153466817G>A	CCDS3777.1	c.1435C>T	p.R479X	Nonsense	15.00%
PAP 011	62	Endometrial	Endometrioid	IB	White	15.2	KRAS	ENSG00000133703	g.chr12:25289551C>G	CCDS8703.1	c.35G>C	p.G12A	Missense	3.50%
PAP 024	56	Endometrial	Endometrioid with squamous differentiation	IA	White	5.6	CTNNB1	ENSG00000168036	g.chr3:41241108G>T	CCDS2694.1	c.101G>T	p.G34V	Missense	0.22%
PAP 025	78	Endometrial	Serous carcinoma	IA	Black	0.74	TP53	ENSG00000141510	g.chr17:7520119_7520120delGG	CCDS11118.1	c.292_293delCC	Frameshift	Indel	8.70%
PAP 026	86	Endometrial	Endometrioid carcinoma	IA	White	17.64	PTEN	ENSG00000171862	g.chr10:89710752G>A	CCDS31238.1	c.923G>A	p.R308H	Missense	0.61%
PAP 027	65	Ovarian	Papillary serous	IIIC	White	18.6	SETD2	ENSG00000181555	g.chr3:47138125G>A	NM_014159	c.805C>T	p.Q269X	Nonsense	<0.005%
PAP 030	73	Endometrial	Endometrioid	IA	White	17.58	MSH6	ENSG00000116062	g.chr2:47800779G>A	CCDS1836.1	c.2153G>A	p.S718N	Missense	0.24%
PAP 031	61	Endometrial	Endometrioid	IA	Asian	301.4	CTNNB1	ENSG00000168036	g.chr3:41241127C>A	CCDS2694.1	c.110C>A	p.S37Y	Missense	1.90%
PAP 032	82	Endometrial	Endometrioid	IA	White	2.218	PTEN	ENSG00000171862	g.chr10:89682884C>G	CCDS31238.1	c.388C>G	p.R130G	Missense	0.02%
PAP 033	68	Endometrial	Endometrioid with squamous differentiation	IA	White	11.3	KRAS	ENSG00000133703	g.chr12:25289551C>A	CCDS8703.1	c.35G>T	p.G12V	Missense	1.50%
PAP 34	55	Endometrial	Endometrioid	IA	White	11.24	CTNNB1	ENSG00000168036	g.chr3:41241107G>A	CCDS2694.1	c.100G>A	p. G34R	Missense	20.12%
PAP 35	50	Endometrial	Endometrioid	IVB	White	0.48	KRAS	ENSG00000133703	g.chr12:25289551C>T	CCDS8703.1	c.35G>A	p.G12D	Missense	0.14%
PAP 36	57	Ovarian	Endometrioid with squamous differentiation	IA	White	2.47	PIK3CA	ENSG00000121879	g.chr3:180418789A>T	CCDS43171.1	c.1637A>T	p.Q546L	Missense	<0.005%
PAP 37	69	Endometrial	Endometrioid	IA	Asian	10.3	PIK3CA	ENSG00000121879	g.chr3:180434712A>G	CCDS43171.1	c.3079A>G	p.T1027A	Missense	0.17%
PAP 38	47	Ovarian	Papillary serous	IA	Asian	5.2	TP53	ENSG00000141510	g.chr17:7518191C>T	CCDS11118.1	c.659A>G	p.Y220C	Missense	0.48%
PAP 39	73	Ovarian	Papillary serous	IA	Asian	9.0	TP53	ENSG00000141510	g.chr17:7518248delG	CCDS11118.1	c.871delC	Frameshift	Indel	5.90%
PAP 40	57	Endometrial	Endometrioid	IB	Asian	7.7	CTNNB1	ENSG00000168036	g.chr3:41241141C>A	CCDS2694.1	c.134C>A	p.S45Y	Missense	0.08%
PAP 41	68	Endometrial	Papillary serous	IA	Asian	9.7	PGFR2	ENSG00000066468	g.chr10:123358886T>C	CCDS7620.2	c.759C>G	p.S253W	Missense	0.15%
PAP 91	68	Endometrial	Papillary serous	IA	Asian	9.7	PTEN	ENSG00000171862	g.chr10:89682902T>C	CCDS31238.1	c.406T>C	p.C136R	Missense	0.06%
PAP 4	45	Ovarian	Adenosquamous Carcinoma	IIIC	White	14.3	TP53	ENSG00000141510	g.chr17:7520074insA	CCDS11118.1	c.339insT	Frameshift	Indel	1.47%
PAP 58	62	Ovarian	Papillary serous	IIIC	White	14.468	PIK3CA	ENSG00000121879	g.chr3:180399630G>A	CCDS43171.1	c.323G>A	p.R108H	Missense	0.45%
PAP 60	69	Ovarian	Papillary serous	IIIC	White	9.312	TP53	ENSG00000141510	g.chr17:7518131C>T	CCDS11118.1	c.524G>A	p.R175H	Missense	<0.05%
PAP 61	70	Ovarian	Papillary serous	IIIC	White	5.872	TP53	ENSG00000141510	g.chr17:7517846G>A	CCDS11118.1	c.817C>T	p.R273C	Missense	0.02%
PAP 62	55	Ovarian	Papillary serous	IIIC	White	1.444	TP53	ENSG00000141510	g.chr17:7520213_7520216delGCAT	CCDS11118.1	c.196_199delATGC	Frameshift	Indel	<0.005%
PAP 63	66	Ovarian	Papillary serous	IIIB	White	7.376	TP53	ENSG00000141510	g.chr17:7519131C>T	CCDS11118.1	c.524G>A	p.R175H	Missense	<0.005%
PAP 64	50	Ovarian	Dysgerminoma	IIIA	Asian	5.6	TP53	ENSG00000141510	g.chr17:7518233T>G	CCDS11118.1	c.773A>C	p.E258A	Missense	4.20%
PAP 66	70	Endometrial	Endometrioid Adenocarcinoma, Villoglandular Type	IIIC	Asian	7.2	TP53	ENSG00000141510	g.chr17:7518960T>C	CCDS11118.1	c.634A>G	p.K [illegible]	Missense	5.40%
PAP 67	63	Endometrial	Endometrioid	IB	Asian	9.3	TP53	ENSG00000141510	g.chr17:7518937G>A	CCDS11118.1	c.637C>T	p.R213X	Nonsense	2.70%
PAP 68	63	Endometrial	Endometrioid Adenocarcinoma with Mucinous Differentiation	IB	Asian	8.6	PIK3CA	ENSG00000121879	g.chr3:180410067A>C	CCDS43171.1	c.1258T>C	p.C420R	Missense	0.03%
PAP 69	51	Endometrial	Endometrioid	IB	Asian	7.4	TP53	ENSG00000141510	g.chr17:7520084G>A	CCDS11118.1	c.328C>T	p.R110C	Missense	14.70%
PAP 70	48	Endometrial	Endometrioid	IIA	Asian	1.6	PIK3CA	ENSG00000121879	g.chr3:180399570G>A	CCDS43171.1	c.263G>A	p.R88Q	Missense	29.60%
PAP 71	57	Endometrial	Endometrioid	IA	White	4.42	PIK3CA	ENSG00000121879	g.chr3:180434779A>G	CCDS43171.1	C.3140A>G	p.H1047R	Missense	0.37%
PAP 72	64	Ovarian	Papillary serous	IV	White	4.04	TP53	ENSG00000141510	g.chr17:7518272C>T	CCDS11118.1	c.794G>A	p.G248D	Missense	0.83%
PAP 73	64	Ovarian	Papillary serous	IIIC	White	4.62	TP53	ENSG00000141510	g.chr17:7519167T>C	CCDS11118.1	c.488A>G	p.Y163C	Missense	<0.005%
PAP 74	79	Ovarian	Papillary serous	IV	White	6.26	TP53	ENSG00000141510	g.chr17:7519131C>T	CCDS11118.1	c.524G>A	p.R175H	Missense	<0.05%
PAP 75	54	Ovarian	Papillary serous	IIIC	White	7.42	PIK3CA	ENSG00000121879	g.chr3:180434768G>A	CCDS43171.1	c.3129G>A	p.M1043I		0.13%
PAP 76	53	Ovarian	Papillary serous	IIIC	White	9.16	TP53	ENSG00000141510	g.chr17:7518275G>A	CCDS11118.1	c.380C>T	p.S127F	Missense	<0.005%
PAP 78	53	Ovarian	Papillary serous	IIIC	White	1.36	TP53	ENSG00000141510	g.chr17:7519131C>T	CCDS11118.1	c.524G>A	p.R175H	Missense	<0.005%
PAP 80	54	Endometrial	Papillary serous	IIIA	Black	7.78	TP53	ENSG00000141510	g.chr17:7519008delG	CCDS11118.1	c.565delC	Frameshift	Indel	79.80%
PAP 83	53	Endometrial	Endometrioid	IA	Black	4.5	PIK3CA	ENSG00000121879	g.chr3:180434779A>G	CCDS43171.1	C.3140A>G	p.H1047R	Missense	0.14%
PAP 83	53	Endometrial	Endometrioid	IA	Black	4.5	CTNNB1	ENSG00000168036	g.chr3:41241105C>A	CCDS2694.1	c.98C>A	p.S33Y	Missense	0.07%

*Coordinates refer to the human reference genome hg18 release (NCBI 36.1, March 2006)

Fig. 8

Table S4. Primers Used to Assess Individual Mutations in Pap Smears

Sample#	Gene	Mutation	Amplified Genomic Fragment Size (bp)	Forward Primer Sequence	Reverse Primer Sequence
PAP 001	PIK3CA	H1047R	82	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNGATCCAATCCATTTTTGTTGTCCAG	cacacaggaaacagctatgaccatgTGAGCAAGAGGCTTTGGAGT
PAP 002	TP53	V147D	93	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNGGCCAAGACCTGCCCTG	cacacaggaaacagctatgaccatgTGCTGTGACTGCTTGTAGATGG
PAP 003	APC	R1450X	93	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCCACCTCCTCAAACAGCTCAA	cacacaggaaacagctatgaccatgTGCAGCTTGCTTAGGTCCACT
PAP 004	TP53	Q167fs*(1 base deletion of C)	85	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNTGGCCATCTACAAGCAGTCA	cacacaggaaacagctatgaccatgNNNNTCACCATCGCTATCTGAGCA
PAP 005	NF1	L414fs*(1 base insertion of C)	74	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCATTGGTGATGATTCGATGG	cacacaggaaacagctatgaccatgCTGCCTGGCTCAGAATTCAC
PAP 006	TP53	R280K	90	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCCCTTTCTTGCGGAGATTC	cacacaggaaacagctatgaccatgCTACTGGGACGGAACAGCTT
PAP 007	TP53	E294X	87	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNGGAAGAGAATCTCCGCAAGA	cacacaggaaacagctatgaccatgGCTTCTTGTCCTGCTTGCTT
PAP 010	FBXW7	R479X	82	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNGGCCTGTCTCAATATCCCAAA	cacacaggaaacagctatgaccatgTTGTTTTTCTGTTTCTCCCTCTG
PAP 011	KRAS	G12A	92	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCAAGGCACTCTTGCCTACG	cacacaggaaacagctatgaccatgCATTTTCATTATTTTTATTATAAGGCCTG
PAP 024	CTNNB1	G34V	72	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCTGTGGTAGTGGCACCAGAA	cacacaggaaacagctatgaccatgNNNNAAGCGGCTGTTAGTCACTGG
PAP 025	TP53	P98fs*(2 base deletion of CC)	91	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNGGCCCCTGTCATCTTCTGT	cacacaggaaacagctatgaccatgGACTTGGCTGTCCCAGAATG
PAP 026	PTEN	R308H	99	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCAAGAAATCGATAGCATTTGCAG	cacacaggaaacagctatgaccatgTTTATTTGCTTTGTCAAGATCATTTTT
PAP 027	SETD2	Q269X	85	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNAGGAAATATCTGCTTGCTCATTCAA	cacacaggaaacagctatgaccatgGAAGCAGATACTAAGCAGGACACTATATC
PAP 030	MSH6	S718N	76	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNTTCCCTTGGATTCTGACACA	cacacaggaaacagctatgaccatgAGCACCATTCGTTGATAGGC
PAP 031	CTNNB1	S37Y	85	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCACTGGCAGCAACAGTCTTACC	cacacaggaaacagctatgaccatgGATTGCCTTTACCACTCAGAGAAG
PAP 032	PTEN	R130G	71	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNtgcagcaattcactgtaaagc	cacacaggaaacagctatgaccatgccgatgtaataaatatgcacatatcattac
PAP 033	KRAS	G12V	92	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCAAGGCACTCTTGCCTACG	cacacaggaaacagctatgaccatgCATTTTCATTATTTTTATTATAAGGCCTG
PAP 034	CTNNB1	G34R	85	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCACTGGCAGCAACAGTCTTACC	cacacaggaaacagctatgaccatgGATTGCCTTTACCACTCAGAGAAG
PAP 035	KRAS	G12D	92	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCAAGGCACTCTTGCCTACG	cacacaggaaacagctatgaccatgCATTTTCATTATTTTTATTATAAGGCCTG
PAP 036	PIK3CA	Q546L	84	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNAGCTCAAAGCAATTTCTACACGA	cacacaggaaacagctatgaccatgNNNNGCACTTACCTGTGACTCCATAGAA
PAP 037	PIK3CA	T1025A	70	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNtcttttgatgacattgcatacattc	cacacaggaaacagctatgaccatgNNNNactccaaagcctcttGCTCA
PAP 038	TP53	Y220C	83	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNcagttgcaaaccagacctca	cacacaggaaacagctatgaccatgttggagtatttggatgacagaaa
PAP 039	TP53	I254fs*(1 base deletion of C)	76	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNgtggcaagtggcctctga	cacacaggaaacagctatgaccatgNNNNCATGGGCGGCATGAAC
PAP 040	CTNNB1	S45Y	90	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNactggcagcaacagtcttacct	cacacaggaaacagctatgaccatgNNNNCCTCAGGATTGCCTTTACCa
PAP 041	FGFR2	S252W	71	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNagtccggcttGGAGGATG	cacacaggaaacagctatgaccatgtcccacactcctccttcttc
PAP 042	PTEN	C136R	75	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNggaaagGGACGAACTGGTG	cacacaggaaacagctatgaccatgNNNNtaggcctcttgtgccttta
PAP 058	TP53	F113fs*(1 base insertion of T)	80	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNtcttctgtcccttcccagaa	cacacaggaaacagctatgaccatgNNNNgacttGGCTGTCCCAGAATG
PAP 059	PIK3CA	R108H	82	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNtcatACCAATTTCTCGATTGAGG	cacacaggaaacagctatgaccatgNNNNcggctttttcaaccctttt
PAP 060	TP53	R175H	85	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNTGGCCATCTACAAGCAGTCA	cacacaggaaacagctatgaccatgNNNNTCACCATCGCTATCTGAGCA
PAP 061	TP53	R273C	76	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNggGACGGAACAGCTTTGAG	cacacaggaaacagctatgaccatgNNNNgcggagattcictictctdgt
PAP 062	TP53	M66fs*(4 base deletion of GCAT	77	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCGGTGTAGGAGCTGCTGGT	cacacaggaaacagctatgaccatgNNNNACCCAGGTCCAGATGAAGC
PAP 063	TP53	R175H	85	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNTGGCCATCTACAAGCAGTCA	cacacaggaaacagctatgaccatgNNNNTCACCATCGCTATCTGAGCA
PAP 064	TP53	E258A	76	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNgtggcaagtggcctctga	cacacaggaaacagctatgaccatgNNNNCATGGGCGGCATGAAC
PAP 066	TP53	Y205C	74	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCGAAAAGTGTTTCTGTCATCCA	cacacaggaaacagctatgaccatgNNNNGCCCCTCCTCAGCATCTTAT
PAP 067	TP53	R213X	83	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNcagttgcaaaccagacctca	cacacaggaaacagctatgaccatgNNNNtgtggagtatttggatgacagaaa
PAP 068	PIK3CA	C420R	90	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNTTATATTTCCCCATGCCAATG	cacacaggaaacagctatgaccatgNNNNggtgttttgaaatgtgttttataatttaga
PAP 069	TP53	R110C	80	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNtcttctgtcccttcccagaa	cacacaggaaacagctatgaccatgNNNNgacttGGCTGTCCCAGAATG
PAP 070	PIK3CA	R88Q	66	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNGAAAAAGCCGAAGGTCACAA	cacacaggaaacagctatgaccatgNNNNCTCAAGAAGCAGAAAGGGAAGA
PAP 071	PIK3CA	H1047R	82	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNGATCCAATCCATTTTTGTTGTCCAG	cacacaggaaacagctatgaccatgNNNNTGAGCAAGAGGCTTTGGAGT
PAP 072	TP53	G245D	74	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNTGTGATGATGGTGAGGATGg	cacacaggaaacagctatgaccatgNNNNtccacTACAACTACATGTGTAACAGTTC
PAP 073	TP53	Y163C	84	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNCTCCGTCATGTGCTGTGACT	cacacaggaaacagctatgaccatgNNNNcagctgtgggttgattcca
PAP 074	TP53	R175H	85	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNTGGCCATCTACAAGCAGTCA	cacacaggaaacagctatgaccatgNNNNTCACCATCGCTATCTGAGCA
PAP 075	PIK3CA	M1043I	79	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNccaatccatttttgttgtcca	cacacaggaaacagctatgaccatgNNNNTGAGCAAGAGGCTTTGGAGT
PAP 076	TP53	S127F	82	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNctgcacagggcaggtctt	cacacaggaaacagctatgaccatgNNNNctctgtctcctcctcttcctaca
PAP 078	TP53	R175H	85	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNTGGCCATCTACAAGCAGTCA	cacacaggaaacagctatgaccatgNNNNTCACCATCGCTATCTGAGCA
PAP 080	PIK3CA	A189fs*(1 base deletion of C)	78	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNcacgcaaATTTCCTTCCACT	cacacaggaaacagctatgaccatgNNNNggcctctgattcctcactgat
PAP 083	PIK3CA	H1047R	82	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNGATCCAATCCATTTTTGTTGTCCAG	cacacaggaaacagctatgaccatgNNNNTGAGCAAGAGGCTTTGGAGT
	CTNNB1	S33Y	90	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNactggcagcaacagtcttacct	cacacaggaaacagctatgaccatgNNNNCCTCAGGATTGCCTTTACCa

Fig. 9

Table S5. Primers Used to Simultaneously Assess 12 Genes in Pap Smears

Set	Gene and Codons Covered by Amplicon	Amplicon Length (bp)	Forward Primer Sequence	Reverse Primer sequence
#1	CTNNB1 005 To 064	251	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNAAAGTAACATTTCCAATCTACTAATGCT	cacacaggaaacagctatgaccatgNNNNtgagaaaatcctgttcccac
	EGFR 697 to 728	264	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNTGGAGCCTCTTACACCCAGT	cacacaggaaacagctatgaccatgNNNNAAAAACACTGGAGTTTCCCAAA
	PIK3CA 0023 to 88	241	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNgaaaAAGCCGAAGGTCACAA	cacacaggaaacagctatgaccatgNNNNATGCCCCCAAGAATCCTAGT
	PTEN 091 to 026	257	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNcatccgtctactcccacgtt	cacacaggaaacagctatgaccatgNNNNatcagctaccgccaagtcc
	PTEN 085 to 123	261	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNTTCGTCCCTTTCCAGCTTTA	cacacaggaaacagctatgaccatgNNNNggaatccagtgtttctttttaaatacct
	TP53 126 to 186	259	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNNaccagccctgtcgtctctc	cacacaggaaacagctatgaccatgNNNNgccctgactttcaactctgtct
	TP53 308 To 331	259	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNgcctcagattcacttttatcacct	cacacaggaaacagctatgaccatgNNNNaccaggagccattgtctttg
	BRAF 439 to 475	251	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNTCACCACATTACATACTTACCATGC	cacacaggaaacagctatgaccatgNNNNaaggggatctctcctgtatcc
	EGFR 738 to 761	279	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNTCTGGATCCCAGAAGGTGAG	cacacaggaaacagctatgaccatgNNNNGGCCAGTGCTGTCTCTAAGG
#2	KRAS 001 to 021	281	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNGTCCACAAAATGATTCTGAATTAGC	cacacaggaaacagctatgaccatgNNNNACGATACACGTCTGCAGTCAAC
	PIK3CA 0062 to 0117	269	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNACAGAAATATTTTAGAAAGGGACAACA	cacacaggaaacagctatgaccatgNNNNAGAAAATACCCCCTCCATCAAC
	PIK3CA 0980 to 1049	272	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNGATCCAATCCATTTTTGTTGTCCAG	cacacaggaaacagctatgaccatgNNNNTCCAAACTGACCAAACTGTTCTTA
	PTEN 100 to 164	294	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNgaaacccaaaatctgttttcca	cacacaggaaacagctatgaccatgNNNNGACCATAACCCACCACAGCTA
	TP53 186 to 224	251	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNctcctcccagagaccccagt	cacacaggaaacagctatgaccatgNNNNcatgagcgctgctcagatag
	TP53 332 To 366	263	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNcctaggaaggcaggggagt	cacacaggaaacagctatgaccatgNNNNtgcatgttgctttgtaccg
	AKT1 017 to 058	279	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNaccaccccgcacgtctgtag	cacacaggaaacagctatgaccatgNNNNagccagtgctgttgcttg
	EGFR 762 to 823	251	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNCACACTGACGTGCCTCTCC	cacacaggaaacagctatgaccatgNNNNttatctcccctcccgtatc
	NRAS 001 to 014	275	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNGATTGTCAGTGCGCTTTTCC	cacacaggaaacagctatgaccatgNNNNgctaaggatgggggttgcta
#3	PIK3CA 0272 to 347	279	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNTGACTTTACCTTATCAATGTCTCGAA	cacacaggaaacagctatgaccatgNNNNGCTCGCCCCCTTAATCTCT
	PPP2R1A 176 to 217	277	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNGGTACTTCCGGAACCTGTGC	cacacaggaaacagctatgaccatgNNNNccgagtcctagggagaggag
	PTEN 056 to 070	250	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNttgttaatggtggcttttttgttt	cacacaggaaacagctatgaccatgNNNNaaatgatctaacaatgctctggac
	PTEN 165 to 211	265	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNcatggaaggatgagaatttcaag	cacacaggaaacagctatgaccatgNNNNtggctacgaccacgttacca
	TP53 033 to 102	256	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNAAACCGTAGCTGCCCTGGTA	cacacaggaaacagctatgaccatgNNNNtgactgctcttttcaccatc
	TP53 225 to 260	257	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNtcatcttgggcctgtgttatc	cacacaggaaacagctatgaccatgNNNNgatgagaggtggatgggtagtagt
#4	EGFR 830 to 875	258	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNTTCAGGGCATGAACTACTTGG	cacacaggaaacagctatgaccatgNNNNatcctcccctgcatgtgtta
	KRAS 038 to 061	260	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNTCCCTCATTGCACTGTACTCC	cacacaggaaacagctatgaccatgNNNNggtgcttagtggccatttgt
	NRAS 038 to 062	250	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNTGGTCTCTCATGGCACTGTACT	cacacaggaaacagctatgaccatgNNNNattagcaatttgagggacaaacc
	PTEN 071 to 084	250	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNtgtatctcactcgataaatctggatgac	cacacaggaaacagctatgaccatgNNNNgtcacattataaagattcaggcaat
	PTEN 213 to 267	271	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNagtttgacagttaaaggcatttcc	cacacaggaaacagctatgaccatgNNNNtgtccttatttggatatttctccc
#5	TP53 063 to 125	276	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNGAAGACCCAGGTCCAGATGA	cacacaggaaacagctatgaccatgNNNNagaaatgcaggggatacg
	TP53 262 To 306	280	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNgaggcataactgcaccctg	cacacaggaaacagctatgaccatgNNNNgggagtagatggagcctggt
	APC 1254 to 1328	269	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNtgctggatttggttctaggg	cacacaggaaacagctatgaccatgNNNNtgccacttgcaaagtttcttc
	APC 1403 to 1476	264	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNggaagaacctggaccctctg	cacacaggaaacagctatgaccatgNNNNttttgagagtcgttcgattgc
	APC 1537 to 1617	289	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNtgcaacctgttttgtgatgg	cacacaggaaacagctatgaccatgNNNNaggaaaatgacaatgggaatga
	FBXW7 474 to 548	296	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNtgattcatcaggagagcatttaag	cacacaggaaacagctatgaccatgNNNNttgttttctgtttctcccctctg
	FBXW7 558 to 618	289	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNGATGGTATCCATGTGGTGAGTG	cacacaggaaacagctatgaccatgNNNNtttgtgatgctaaggctccat
#6	TP53 001 to 24	254	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNactgccttccgggtcact	cacacaggaaacagctatgaccatgNNNNagcccaaccctgtccttac
	TP53 368 To 393	265	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNgaggctgtcagtggggaac	cacacaggaaacagctatgaccatgNNNNaacatatttgcatggggtgtg
	APC 1326 to 1405	283	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNccactgcatggttcactctg	cacacaggaaacagctatgaccatgNNNNatcctgtgagcgaagttcca
	APC 1475 to 1545	264	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNtgcctctttctcttggtttca	cacacaggaaacagctatgaccatgNNNNggacctaagcaagctgcagtaa
	FBXW7 243 to 287	257	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNacacccaatgaagaatgtaattgat	cacacaggaaacagctatgaccatgNNNNggttgtgtgtagatgtgagtttcc
	FBXW7 329 to 374	276	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNttctgttacattgtgcagagttca	cacacaggaaacagctatgaccatgNNNNtggttttgagcagagagatgg
	FBXW7 413 to 472	266	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNggaagaagtcccaaccatga	cacacaggaaacagctatgaccatgNNNNtcacttttcctttctacccaaaag
	FBXW7 549 to 593	259	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNGAATCTGCATTCCCAGAGACA	cacacaggaaacagctatgaccatgNNNNcctgtttcccatcctcttcc
	KRAS 134 to 150	267	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNGACACAAAACAGGCTCAGGAC	cacacaggaaacagctatgaccatgNNNNagaaaccaaagccaaaagca
	TP53 026 to 32	253	cgacgtaaaacgacggccagtNNNNNNNNNNNNNNcccatgggactgacttctgc	cacacaggaaacagctatgaccatgNNNNtcatctggacctgggtcttc

Fig. 10

Table S6. Mutations Identified in Pap Smears through Simultaneous Assessment of 12 Genes

Case #	Tumor Type	Mutated Gene Name	Mutated Gene ID	Nucleotide (genomic)*	Transcript	Nucleotide (transcript)*	Amino Acid (protein)	Mutation Type	Fraction of mutant alleles in Pap smear fluid
PAP 001	Endometrial	KRAS	ENSG00000133703	g.chr12:25289551C>G	CCDS8703.1	c.35G>C	p.G12A	Missense	12.51%
		PIK3CA	ENSG00000121879	g.chr3:180434779A>G	CCDS43171.1	c.3140A>G	p.H1047R	Missense	5.74%
PAP 003	Endometrial	PIK3CA	ENSG00000121879	g.chr3:180399570G>A	CCDS43171.1	c.263G>A	p.R88Q	Missense	11.60%
		APC	ENSG00000134982	g.chr5:112203538C>T	CCDS4107.1	c.4348C>T	p.R1450X	Missense	12.50%
		PTEN	ENSG00000171862	g.chr10:89614243A>C	CCDS31238.1	c.38A>C	p.K13T	Missense	12.38%
PAP 010	Endometrial	FBXW7	ENSG00000109670	g.chr4:153466817G>A	CCDS3777.1	c.1435C>T	p.R479X	Nonsense	20.00%
PAP 011	Endometrial	KRAS	ENSG00000133703	g.chr12:25289551C>G	CCDS8703.1	c.35G>C	p.G12A	Missense	3.23%
PAP 025	Endometrial	TP53	ENSG00000141510	g.chr17:7520119_7520120delGG	CCDS11118.1	c.292_293delCC	Frameshift	Indel	10.42%
PAP 033	Endometrial	PIK3CA	ENSG00000121879	g.chr3:180404243T>G	CCDS43171.1	c.1031T>G	p.V344G	Missense	1.22%
		KRAS	ENSG00000133703	g.chr12:25289551C>A	CCDS8703.1	c.35G>T	p.G12V	Missense	1.13%
		PTEN	ENSG00000171862	g.chr10:89707637insC	CCDS31238.1	c.682insC	Frameshift	Indel	0.87%
PAP 34	Endometrial	PIK3CA	ENSG00000121879	g.chr3:180399630G>A	CCDS43171.1	c.323G>A	p.R108H	Missense	22.78%
		CTNNB1	ENSG00000168036	g.chr3:41241107G>A	CCDS2694.1	c.100G>A	p. G34R	Missense	18.41%
		PTEN	ENSG00000171862	g.chr10:89707750delA	CCDS31238.1	c.795delA	Frameshift	Indel	13.28%
		PIK3CA	ENSG00000121879	g.chr3:180399554T>G	CCDS43171.1	c.247T>G	p.F83V	Missense	4.49%
		KRAS	ENSG00000133703	g.chr12:25289551C>T	CCDS8703.1	c.35G>A	p.G12D	Missense	0.92%
PAP 35	Endometrial	KRAS	ENSG00000133703	g.chr12:25289551C>T	CCDS8703.1	c.35G>A	p.G12D	Missense	5.83%
		PIK3CA	ENSG00000121879	g.chr3:180399548G>A	CCDS43171.1	c.241G>A	p.E81K	Missense	5.32%
		PTEN	ENSG00000171862	g.chr10:89675287T>C	CCDS31238.1	c.202T>C	p.Y68H	Missense	4.73%
PAP 39	Ovarian	TP53	ENSG00000141510	g.chr17:7518244delG	CCDS11118.1	c.871delC	Frameshift	Indel	0.73%
PAP 67	Endometrial	TP53	ENSG00000141510	g.chr17:7518937G>A	CCDS11118.1	c.637C>T	p.R213X	Nonsense	2.31%
PAP 69	Endometrial	TP53	ENSG00000141510	g.chr17:7520084G>A	CCDS11118.1	c.328C>T	p.R110C	Missense	19.23%
		TP53	ENSG00000141510	g.chr17:7517747G>A	CCDS11118.1	c.916C>T	p.R306X	Nonsense	13.60%
PAP 70	Endometrial	PIK3CA	ENSG00000121879	g.chr3:180399570G>A	CCDS43171.1	c.263G>A	p.R88Q	Missense	28.05%
PAP 71	Endometrial	KRAS	ENSG00000133703	g.chr12:25289551C>T	CCDS8703.1	c.35G>A	p.G12D	Missense	0.39%
		PIK3CA	ENSG00000121879	g.chr3:180434779A>G	CCDS43171.1	c.3140A>G	p.H1047R	Missense	0.31%
PAP 72	Ovarian	TP53	ENSG00000141510	g.chr17:7518272C>T	CCDS11118.1	c.734G>A	p.G245D	Missense	0.54%

*Coordinates refer to the human reference genome hg18 release (NCBI 36.1, March 2006).

Fig. 11

PAPANICOLAOU TEST FOR OVARIAN AND ENDOMETRIAL CANCERS

[0001] This invention was made with government support under grant no. CA043460, CA062924, CN043309, and CA129825 awarded by the National Institutes of Health. The government has certain rights in the invention.

TECHNICAL FIELD OF THE INVENTION

[0002] This invention is related to the area of cancer screening. In particular, it relates to ovarian and endometrial cancers.

BACKGROUND OF THE INVENTION

[0003] Since the introduction of the Papanicolaou test, the incidence and mortality of cervical cancer in screened populations has been reduced by more than 75% (1, 2). In contrast, deaths from ovarian and endometrial cancers have not substantially decreased during that same time period. As a result, more than 69,000 women in the U.S. will be diagnosed with ovarian and endometrial cancer in 2012. Although endometrial cancer is more common than ovarian cancer, the latter is more lethal. In the U.S., approximately 15,000 and 8,000 women are expected to die each year from ovarian and endometrial cancers, respectively (Table 1). World-wide, over 200,000 deaths from these tumors are expected this year alone (3, 4).

[0004] In an effort to replicate the success of cervical cancer screening, several approaches for the early detection of endometrial and ovarian cancers have been devised. For endometrial cancers, efforts have focused on cytology and transvaginal ultrasound (TVS). Cytology can indeed indicate a neoplasm within the uterus in some cases, albeit with low specificity (5). TVS is a noninvasive technique to measure the thickness of the endometrium based on the fact that endometria harboring a cancer are thicker than normal endometria (6). As with cytology, screening measurement of the endometrial thickness using TVS lacks sufficient specificity because benign lesions, such as polyps, can also result in a thickened endometrium. Accordingly, neither cytology nor TVS fulfills the requirements for a screening test (5, 7).

[0005] Even greater efforts have been made to develop a screening test for ovarian cancer, using serum CA-125 levels and TVS. CA-125 is a high molecular weight transmembrane glycoprotein expressed by coelomic- and Müllerian-derived epithelia that is elevated in a subset of ovarian cancer patients with early stage disease (8). The specificity of CA-125 is limited by the fact that it is also elevated in a variety of benign conditions, such as pelvic inflammatory disease, endometriosis and ovarian cysts (9). TVS can visualize the ovary but can only detect large tumors and cannot definitively distinguish benign from malignant tumors. Several clinical screening trials using serum CA-125 and TVS have been conducted but none has shown a survival benefit. In fact, some have shown an increase in morbidity compared to controls because false positive tests elicit further evaluation by laparoscopy or exploratory laparotomy (10-12).

[0006] Accordingly, the U.S. Preventative Services Task Force, the American Cancer Society, the American Congress of Obstetricians and Gynecologists, as well as the National Comprehensive Cancer Network, do not recommend routine screening for endometrial or ovarian cancers in the general population. In fact, these organizations warn that "the potential harms outweigh the potential benefits" (13-16). An exception to this recommendation has been made for patients with a hereditary predisposition to ovarian cancer, such as those with germline mutations in a BRCA gene or those with Lynch syndrome. It is recommended that BRCA mutation carriers be screened every 6 months with TVS and serum CA-125, starting at a relatively early age. Screening guidelines for women with Lynch syndrome include annual endometrial sampling and TVS beginning between age 30 and 35 (15, 17).

[0007] The mortality associated with undetected gynecologic malignancies has made the development of an effective screening tool a high priority. An important observation that inspired the current study is that asymptomatic women occasionally present with abnormal glandular cells (AGCs) detected in a cytology specimen as part of their routine cervical cancer screening procedure. Although AGCs are associated with premalignant or malignant disease in some cases (18-22), it is often difficult to distinguish the AGCs arising from endocervical, endometrial or ovarian cancer from one another or from more benign conditions. There is a continuing need in the art to detect these cancers at an earlier stage than done currently.

SUMMARY OF THE INVENTION

[0008] According to one aspect of the invention a method is provided for detecting or monitoring endometrial or ovarian cancer. A liquid Pap smear of a patient is tested for agenetic or epigenetic change in one or more genes, mRNAs, or proteins mutated in endometrial or ovarian cancer. Detection of the change indicates the presence of such a cancer in the patient.

[0009] According to another aspect of the invention a method is provided for screening for endometrial and ovarian cancers. A liquid Pap smear is tested for one or more mutations in a gene, mRNA, or protein selected from the group consisting of CTNNB1, EGFR, PI3KCA, PTEN, TP53, BRAF, KRAS, AKT1, NRAS, PPP2R1A, APC, FBXW7, ARID1A, CDKN2A, MLL2, RFF43, and FGFR2. Detection of the mutation indicates the presence of such a cancer in the patient.

[0010] Another aspect of the invention is a kit for testing a panel of genes in Pap smear samples for ovarian or endometrial cancers. The kit comprises at least 10 probes or at least 10 primer pairs. Each probe or primer comprises at least 15 nt of complementary sequence to one of the panel of genes. At least 10 different genes are interrogated. The panel is selected from the group consisting of CTNNB1, EGFR, PI3KCA, PTEN, TP53, BRAF, KRAS, AKT1, NRAS, PPP2R1A, APC, FBXW7, ARID1A, CDKN2A, MLL2, RFF43, and FGFR2.

[0011] Still another aspect of the invention is a solid support comprising at least 10 attached probes. Each probe comprises at least 15 nt of complementary sequence to one of a panel of genes, wherein the panel is selected from the group consisting of CTNNB1, EGFR, PI3KCA, PTEN, TP53, BRAF, KRAS, AKT1, NRAS, PPP2R1A, APC, FBXW7, ARID1A, CDKN2A, MLL2, RFF43, and FGFR2.

[0012] Another aspect of the invention is a solid support comprising at least 10 primers attached thereto. Each primer comprises at least 15 nt of complementary sequence to one of a panel of genes. The panel is selected from the group consisting of CTNNB1, EGFR, PI3KCA, PTEN, TP53,

BRAF, KRAS, AKT1, NRAS, PPP2R1A, APC, FBXW7, ARID1A, CDKN2A, MLL2, RFF43, and FGFR2.

[0013] These and other embodiments which will be apparent to those of skill in the art upon reading the specification provide the art with methods for assessing ovarian and endometrial cancers in a screening environment using samples that are already routinely collected.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] FIG. **1**. Schematic demonstrating the principle steps of the procedure described in this study. Tumors cells shed from ovarian or endometrial cancers are carried into the endocervical canal. These cells can be captured by the brush used for performing a routine Pap smear. The brush contents are transferred into a liquid fixative, from which DNA is isolated. Using next-generation sequencing, this DNA is queried for mutations that indicate the presence of a malignancy in the female reproductive tract.

[0015] FIG. **2**. Diagram of the assay used to simultaneously detect mutations in 12 different genes. A modification of the Safe-SeqS (Safe-Sequencing System) protocol, for simultaneous interrogation of multiple mutations in a single sample, is depicted. In the standard Safe-SeqS procedure, only one amplicon is assessed, while the new system is used to assess multiple amplicons from multiple genes at once.

[0016] FIG. **3**. Mutant allele fractions in Pap smear fluids. The fraction of mutant alleles from each of 46 pap smear fluids is depicted. The stage of each tumor is listed on the Y-axis. The X-axis demonstrates the % mutant allele fraction as determined by Safe-SeqS.

[0017] FIG. **4**. Heat map depicting the results of multiplex testing of 12 genes in Pap smear fluids. Each block on the y-axis represents a 30-bp block of sequence from the indicated gene. The 28 samples assessed (14 from women with cancer, 14 from normal women without cancer) are indicated on the x-axis. Mutations are indicated as colored blocks, with white indicating no mutation, yellow indicating a mutant fraction of 0.1% to 1%, orange indicate a mutant fraction of 1% to 10%, and red indicating a mutant fraction of >10%.

[0018] FIG. **5**. Table 1. Epidemiology of Ovarian and Endometrial Tumors. The estimated numbers of new cases and deaths in the U.S. from the major subtypes of ovarian and endometrial cancers are listed.

[0019] FIG. **6**. Table 2. Genetic Characteristics of Ovarian and Endometrial Cancers. The frequencies of the commonly mutated genes in ovarian and endometrial cancers are listed.

[0020] FIG. **7**. Table S1. Endometrial Cancers (Endometrioid Subtype) Studied by Whole-exome Sequencing. The summary characteristics of the 22 cancers used for exome sequencing are listed.

[0021] FIG. **8**. Table S3. Mutations Assessed in Pap Smears. Clinical characteristics of the 46 tumor samples are listed, along with the mutation identified in each case and the fraction of mutant alleles identified in the Pap smears.

[0022] FIG. **9**. Table S4. Primers Used to Assess Individual Mutations in Pap Smears. The sequences of the forward and reverse primers used to test each mutation via Safe-SeqS are listed in pairs (SEQ ID NO: 4-99, respectively).

[0023] FIG. **10**. Table S5. Primers Used to Simultaneously Assess 12 Genes in Pap Smears. The sequences of the forward and reverse primers for each tested region are listed in pairs (SEQ ID NO: 100-191, respectively).

[0024] FIG. **11**. Table S6. Mutations Identified in Pap Smears through Simultaneous Assessment of 12 Genes. The fraction of mutant alleles identified in the Pap smears using this approach is listed, along with the precise mutations identified.

DETAILED DESCRIPTION OF THE INVENTION

[0025] The inventors have developed a test for detecting different cancers using samples that are already routinely collected for diagnosing uterine cancer and HPV (human papilloma virus) infection. Using a panel of genes, a high level of detection of both endometrial and ovarian cancers was achieved.

[0026] Certain genes have been identified as mutated in a high proportion of endometrial and ovarian cancers. These include CTNNB1, EGFR, PI3KCA, PTEN, TP53, BRAF, KRAS, AKT1, NRAS, PPP2R1A, APC, FBXW7, ARID1A, CDKN2A, MLL2, RFF43, and FGFR2. The test can be performed on at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 of these genes. In addition, other genes can be added or substituted into the panel to achieve a higher rate of detection.

[0027] Testing for a mutation may be done by analysis of nucleic acids, such as DNA or mRNA or cDNA. The nucleic acid analytes are isolated from cells or cell fragments found in the liquid PAP smear sample. Suitable tests may include any hybridization or sequencing based assay. Analysis may also be performed on protein encoded by the genes in the panel. Any suitable test may be used including but not limited to mass spectrometry. Other suitable assays may include immunological assays, such as, immunoblotting, immunocytochemistry, immunoprecipitation, enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), immunoradiometric assays (IRMA) and immunoenzymatic assays (IEMA), including sandwich assays using monoclonal or polyclonal antibodies.

[0028] Genetic changes which can be detected are typically mutations such as deletions, insertions, duplications, substitutions (missense or nonsense mutations), rearrangements, etc. Such mutations can be detected inter alia by comparing to a wild type in another (non-tumor) tissue or fluid of an individual or by comparing to reference sequences, for example in databases. Mutations that are found in all tissues of an individual are germline mutations, whereas those that occur only in a single tissue are somatic mutations. Epigenetic changes can also be detected. These may be loss or gain of methylation at specific locations in specific genes, as well as histone modifications, including acetylation, ubiquitylation, phosphorylation and sumoylation.

[0029] Tests may be done in a multiplex format, in which a single reaction pot is used to detect multiple analytes. Examples of such tests include amplifications using multiple primer sets, amplifications using universal primers, array based hybridization or amplification, emulsion based amplification.

[0030] While probes and primers may be designed to interrogate particular mutations or particular portions of a gene, mRNA, or cDNA, these may not be separate entities. For example, probes and primers may be linked together to form a concatamer, or they may be linked to one or more solid supports, such as a bead or an array.

[0031] Kits for use in the disclosed methods may include a carrier for the various components. The carrier can be a container or support, in the form of, e.g., bag, box, tube, rack, and is optionally compartmentalized. The kit also includes various components useful in detecting mutations, using the above-discussed detection techniques. For example, the detection kit may include one or more oligonucleotides useful as primers for amplifying all or a portion of the target nucleic acids. The detection kit may also include one or more oligonucleotide probes for hybridization to the target nucleic acids. Optionally the oligonucleotides are affixed to a solid support, e.g., incorporated in a microarray included in the kit or supplied separately.

[0032] Solid supports may contain one single primer or probe or antibody for detecting a single gene, protein, mRNA, or portion of a gene. A solid support may contain multiple primers, probes, or antibodies. They may be provided as a group which will interrogate mutations at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 of the genes of the desired panel. The panel may be selected from or comprise CTNNB1, EGFR, PI3KCA, PTEN, TP53, BRAF, KRAS, AKT1, NRAS, PPP2R1A, APC, FBXW7, ARID1A, CDKN2A, MLL2, RFF43, and FGFR2.

[0033] Primer pairs may be used to synthesize amplicons of various sizes. Amplicons may be for example from 50, 60, 75, 100, 125, 150, 200, 140, 180 bp in length. Amplicons may run up to 200, 250, 300, 400, 500, 750, 1000 bp in length, as examples. The size of the amplicon may be limited by the size and/or quality of the template retrieved from the liquid PAP smear. Probes and primers for use in the invention may contain a wild-type sequence or may contain a sequence of a particular mutant.

[0034] In one embodiment, the test can be performed using samples that are collected over time. The test results can be compared for quantitative or qualitative changes. Such analysis can be used after a potentially curative therapy, such as surgery.

[0035] Georgios Papanicolaou published his seminal work, entitled "Diagnosis of Uterine Cancer by the Vaginal Smear," in 1943 (31). At that time, he suggested that endocervical sampling could, in theory, be used to detect not only cervical cancers but also other cancers arising in the female reproductive tract, including endometrial carcinomas. The research reported here moves us much closer to that goal. In honor of Papanicolaou's pioneering contribution to the field of early cancer detection, we have named the approach described herein as the "PapGene" test.

[0036] One of the most important developments over the last several years is the recognition that all human cancers are the result of mutations in a limited set of genes and an even more limited set of pathways through which these genes act (32). The whole-exome sequencing data we present, combined with previous genome-wide studies, provide a striking example of the common genetic features of cancer (FIG. **6**, Table 2). Through the analysis of particular regions of only 12 genes (FIG. **11**, table S5), we could detect at least one driver mutation in the vast majority of nine different gynecologic cancers (FIG. **5**, Table 1). Though several of these 12 genes were tumor suppressors, and therefore difficult to therapeutically target, knowledge of their mutational patterns provides unprecedented opportunities for cancer diagnostics.

[0037] The most important finding in this paper is that significant amounts of cells or cell fragments from endometrial and ovarian cancers are present in the cervix and can be detected through molecular genetic approaches. Detection of malignant cells from endometrial and ovarian carcinomas in cervical cytology specimens is relatively uncommon. Microscopic examination cannot always distinguish them from one another, from cervical carcinomas, or from more benign conditions. Our study showed that 100% of endometrial cancers (n=24), even those of low grade, and 41% of ovarian cancers (n=22), shed cells into the cervix that could be detected from specimens collected as part of routine Pap smears. This finding, in conjunction with technical advances allowing the reliable detection of mutations present in only a very small fraction of DNA templates, provided the foundation for the PapGene test.

[0038] This study demonstrates the value of sensitive endocervical DNA testing but there are many issues that need to be addressed before optimal clinical use is achieved. The test, even in its current format, appears to be promising for screening endometrial cancer, as the data in FIG. **3** show that even the lowest stage endometrial cancers could be detected through the analysis of DNA in Pap smear fluid through Safe-SeqS. However, only 41% of ovarian cancers could be detected in Pap smears even when the mutations in their tumors were known. In eight of the nine Pap smears from ovarian cancer patients that contained detectable mutations, the mutant allele fractions were >0.1% and therefore within the range currently detectable by PapGene testing (FIG. **9**, table S3). Further improvements in the technology could increase the technical sensitivity of the PapGene test and allow it to detect more ovarian cancers. Other strategies to increase this sensitivity involve physical maneuvers, such as massaging the adnexal region during the pelvic examination or by performing the PapGene test at specified times during the menstrual cycle. Development of an improved method of collection may also be able to improve sensitivity. The current liquid specimen is designed for the detection of cervical cancer and as such utilizes a brush that collects cells from the ectocervix and only minimally penetrates the endocervical canal. A small cannula that can be introduced into the endometrial cavity similar to the pipelle endometrial biopsy instrument could theoretically obtain a more enriched sample of cells coming from the endometrium, fallopian tube and ovary (33).

[0039] The high sensitivity and the quantitative nature of the PapGene test also opens the possibility of utilizing it to monitor the response to hormonal agents (e.g., progestins) used to treat young women with low risk endometrial cancers. Some of these women choose to preserve fertility, undergoing medical therapy rather than hysterectomy (34). The detection of pre-symptomatic ovarian cancers, even if advanced, could also be of benefit. Although not entirely analogous, it has been demonstrated that one of the most important prognostic indicators for ovarian cancer is the amount of residual disease after surgical debulking. Initially, debulking was considered optimal if the residual tumor was less than 2 cm. Subsequently, the threshold was reduced to 1 cm and surgeons now attempt to remove any visible tumor. With each improvement in surgical debulking, survival has lengthened (35). A small volume of tumor is likely to be more sensitive to cytotoxic chemotherapy than the large, bulky disease typical of symptomatic high-grade serous carcinoma.

[0040] An essential aspect of the screening approach described here is that it should be relatively inexpensive and

easily incorporated into the pelvic examination. Evaluation of HPV DNA is already part of routine Pap smear testing because HPV analysis increases the test's sensitivity (36, 37). The DNA purification component of the PapGene test is identical to that used for HPV, so this component is clearly feasible. The preparation of DNA, multiplex amplification, and the retail cost of the sequencing component of the PapGene test can also be performed at a cost comparable to a routine HPV test in the U.S. today. Note that the increased sensitivity provided by the Safe-SeqS component of the PapGene test (see Example 6) can be implemented on any massively parallel sequencing instrument, not just those manufactured by Illumina. With the reduction in the cost of massively parallel sequencing expected in the future, PapGene testing should become even less expensive.

[0041] There are millions of Pap smear tests performed annually in the U.S. Could PapGene testing be performed on such a large number of specimens? We believe so, because the entire DNA purification and amplification process can be automated, just as it is for HPV testing. Though it may now seem unrealistic to have millions of these sophisticated sequence-based tests performed every year, it would undoubtedly have seemed unrealistic to have widespread, conventional Pap smear testing performed when Papanicolaou published his original paper (31). Even today, when many cervical cytology specimens are screened using automated technologies, a significant percentage require evaluation by a skilled cyotpathologist. In contrast, the analysis of PapGene testing is done completely in silico and the readout of the test is objective and quantitative.

[0042] In sum, PapGene testing has the capacity to increase the utility of conventional cytology screening through the unambiguous detection of endometrial and ovarian carcinomas. In addition to the analysis of much larger numbers of patients with and without various types of endometrial, ovarian, and fallopian tube cancers, the next step in this line of research is to include genes altered in cervical cancer as well as HPV amplicons in the multiplexed Safe-SeqS assay (FIG. **11**, table S5). These additions will provide information that could be valuable for the management of patients with the early stages of cervical neoplasia, as HPV positivity alone is not specific for the detection of cervical cancer and its precursor lesions, particularly in young, sexually active women who frequently harbor HPV infections in the absence of neoplasia.

[0043] The above disclosure generally describes the present invention. All references disclosed herein are expressly incorporated by reference. A more complete understanding can be obtained by reference to the following specific examples which are provided herein for purposes of illustration only, and are not intended to limit the scope of the invention.

EXAMPLE 1

[0044] We reasoned that more sophisticated molecular methods might be able to detect the presence of cancer cells in endocervical specimens at higher sensitivities and specificities than possible with conventional methods. In particular, we hypothesized that somatic mutations characteristic of endometrial and ovarian cancers would be found in the DNA purified from routine liquid-based Pap smears (henceforth denoted as "Pap smears"; FIG. **1**). Unlike cytologically abnormal cells, such oncogenic DNA mutations are specific, clonal markers of neoplasia that should be absent in non-neoplastic cells. However, we did not know if such DNA would indeed be present in endocervical specimens, and we did not know if they would be present in a sufficient amount to detect them. The experiments described here were carried out to test our hypothesis.

[0045] There were four components to this study: I. Determination of the somatic mutations typically present in endometrial and ovarian cancers; II. Identification of at least one mutation in the tumors of 46 patients with these cancers; III. Determination of whether the mutations identified in these tumors could also be detected in Pap smears from the same patients; and IV. Development of a technology that could directly assess cells from Pap smears for mutations commonly found in endometrial or ovarian cancers.

EXAMPLE 2

Prevalence of Somatically Mutated Genes in Endometrial and Ovarian Cancers

[0046] There are five major histopathologic subtypes of ovarian cancers. The most prevalent subtype is high grade serous (60% of total), followed by endometrioid (15%), clear cell (10%), and low-grade serous carcinoma (8%) (Table 1). Genome-wide studies have identified the most commonly mutated genes among the most prevalent ovarian cancer subtypes (Table 2) (23-25).

[0047] Such comprehensive studies have not yet been reported for the endometrioid and mucinous subtypes, collectively representing ~20% of ovarian cancer cases (Table 1). However, commonly mutated genes in the endometrioid and mucinous subtypes have been reported (26). In aggregate, the most commonly mutated gene in epithelial ovarian cancers was TP53, which was mutated in 69% of these cancers (Table 2). Other highly mutated genes included ARID1A, BRAF, CTNNB1, KRAS, PIK3CA, and PPP 2R1A (Table 2).

[0048] Among endometrial cancers, the endometrioid subtype is by far the most common, representing 85% of the total (Table 1). Because cancers of this subtype are so frequent and have not been analyzed at a genome-wide level, we evaluated them through whole-exome sequencing. The DNA purified from 22 sporadic endometrioid carcinomas, as well as from matched non-neoplastic tissues, was used to generate 44 libraries suitable for massively parallel sequencing. The clinical aspects of the patients and histopathologic features of the tumors are listed in table 51. Though the examination of 22 cancers cannot provide a comprehensive genome landscape of a tumor type, it is adequate for diagnostic purposes—as these only require the identification of the most frequently mutated genes.

[0049] Among the 44 libraries, the average coverage of each base in the targeted region was 149.1 with 88.4% of targeted bases represented by at least ten reads. Using stringent criteria for the identification of somatic mutations (as described in Materials and Methods), the sequencing data clearly demarcated the tumors into two groups: ten cancers (termed the N Group, for non-highly mutated) harbored <100 somatic mutations per tumor (median 32, range 7 to 50), while 12 cancers (termed the H Group, for highly mutated) harbored >100 somatic mutations per tumor (median 674, range 164 to 4,629) (FIG. **7**, table S1).

[0050] The high number of mutations in the Group H tumors was consistent with a deficiency in DNA repair. Eight of the 12 Group H tumors had microsatellite instability

(MSI-H, table S1), supporting this conjecture. Moreover, six of the Group H tumors contained somatic mutations in the mismatch repair genes MSH2 or MSH6, while none of the Group N cancers contained mutations in mismatch repair genes. Mismatch repair deficiency is known to be common among endometrial cancers and these tumors occur in 19-71% of women with inherited mutations of mismatch repair genes (i.e., patients with the Hereditary Nonpolyposis Colorectal Cancer) (27).

[0051] 12,795 somatic mutations were identified in the 22 cancers. The most commonly mutated genes included the PIK3 pathway genes PTEN and PIK3CA (28), the APC pathway genes APC and CTNNB1, the fibroblast growth factor receptor FGFR2, the adapter protein FBXW7, and the chromatin-modifying genes ARID1A and MLL2 (Table 2). Genes in these pathways were mutated in both Group N and H tumors. Our results are consistent with prior studies of endometrioid endometrial cancer that had evaluated small numbers of genes, though mutations in FBXW7, MLL2 and APC had not been appreciated to occur as frequently as we found them. It was also interesting that few TP53 mutations (5%) were found in these endometrial cancers (Table 2), a finding also consistent with prior studies.

[0052] Papillary serous carcinomas of the endometrium account for 10-15% of endometrial cancers, and a recent genome-wide sequencing study of this tumor subtype has been published (29). The most common mutations in this subtype are listed in Table 2. The least common subtype of endometrial cancers is clear cell carcinomas, which occur in <5%. Genes reported to be mutated in these cancers were garnered from the literature (Table 2).

EXAMPLE 3

Identification of Mutations in Tumor Tissues

[0053] We acquired tumors from 46 cancer patients in whom Pap smears were available. These included 24 patients with endometrial cancers and 22 with ovarian cancers; clinical and histopathologic features are listed in table S3.

[0054] Somatic mutations in the 46 tumors were identified through whole-exome sequencing as described above or through targeted sequencing of genes frequently mutated in the most common subtypes of ovarian or endometrial cancer (Table 2). Enrichment for these genes was achieved using a custom solid phase capture assay comprised of oligonucleotides ("capture probes") complementary to a panel of gene regions of interest. For the oncogenes, we only targeted their commonly mutated exons, whereas we targeted the entire coding regions of the tumor suppressor genes.

[0055] Illumina DNA sequencing libraries were generated from tumors and their matched non-neoplastic tissues, then captured with the assay described above. Following amplification by PCR, four to eight captured DNA libraries were sequenced per lane on an Illumina GA IIx instrument. In each of the 46 cases, we identified at least one somatic mutation (table S3) that was confirmed by an independent assay, as described below.

EXAMPLE 4

Identification of Somatic Mutations in Pap Smears

[0056] In the liquid-based Pap smear technique in routine use today, the clinician inserts a small brush into the endocervical canal during a pelvic exam and rotates the brush so that it dislodges and adheres to loosely attached cells or cell fragments. The brush is then placed in a vial of fixative solution (e.g., ThinPrep). Some of the liquid from the vial is used to prepare a slide for cytological analysis or for purification of HPV DNA. In our study, an aliquot of the DNA purified from the liquid was used to assess for the presence of DNA from the cancers of the 46 patients described above. Preliminary studies showed that the fixed cells or cell fragments in the liquid, pelleted by centrifugation at 1,000 g for five minutes, contained >95% of the total DNA in the vial. We therefore purified DNA from the cell pellets when the amount of available liquid was greater than 3 mL (as occurs with some liquid-based Pap smear kits) and, for convenience, purified DNA from both the liquid and cells when smaller amounts of liquid were in the kit. In all cases, the purified DNA was of relatively high molecular weight (95%>5 kb). The average amount of DNA recovered from the 46 Pap smears was 49.3±74.4 ng/ml (table S3).

[0057] We anticipated that, if present at all, the amount of DNA derived from neoplastic cells in the Pap smear fluid would be relatively small compared to the DNA derived from normal cells brushed from the endocervical canal. This necessitated the use of an analytic technique that could reliably identify a rare population of mutant alleles among a great excess of wild-type alleles. A modification of one of the Safe-SeqS (Safe-Sequencing System) procedures described in (30) was designed for this purpose (FIG. **2**).

[0058] In brief, a limited number of PCR cycles was performed with a set of gene-specific primers. One of the primers contained 14 degenerate N bases (equal probability of being an A, C, G, or T) located 5' to its gene-specific sequence, and both primers contained sequences that permitted universal amplification in the next step. The 14 N's formed unique identifiers (UID) for each original template molecule. Subsequent PCR products generated with universal primers were purified and sequenced on an Illumina MiSeq instrument. If a mutation preexisted in a template molecule, that mutation should be present in every daughter molecule containing that UID, and such mutations are called "supermutants" (30). Mutations not occurring in the original templates, such as those occurring during the amplification steps or through errors in base calling, should not give rise to supermutants. The Safe-SeqS approach used here is capable of detecting 1 mutant template among 5,000 to 1,000,000 wild-type templates, depending on the amplicon and the position within the amplicon that is queried (30).

[0059] We designed Safe-SeqS primers (table S4) to detect at least one mutation from each of the 46 patients described in table S3. In the 24 Pap smears from patients with endometrial cancers, the mutation present in the tumor was identified in every case (100%). The median fraction of mutant alleles was 2.7%, and ranged from 0.01% to 78% (FIG. **3** and table S3). Amplifications of DNA from non-neoplastic tissues were used as negative controls in these experiments to define the detection limits of each queried mutation. In all cases, the fraction of mutant alleles was significantly different from the background mutation levels determined from the negative controls ($P<0.001$, binomial test). There was no obvious correlation between the fraction of mutant alleles and the histopathologic subtype or the stage of the cancer (FIG. **3** and table S3).

[0060] In two endometrial cancer cases, two mutations found in the tumor DNA were evaluated in the Pap smears

(table S3). In both cases, the mutations were identified in DNA from the Pap smear (table S3). Moreover, the ratios between the mutant allele fractions of the two mutations in the Pap smears were correlated with those of the corresponding tumor samples. For example, in the Pap smear of case PAP 083 the mutant allele fractions for the CTNNB1 and PIK3CA mutations were 0.143% and 0.064%, respectively—a ratio of 2.2 (=0.14% to 0.064%). In the primary tumor from PAP 083, the corresponding ratio was 2.0 (79.5% to 39.5%).

[0061] Similar analysis of Pap smear DNA from ovarian cancer patients revealed detectable mutations in nine of the 22 patients (41%). The fraction of mutant alleles was smaller than in endometrial cancers (median of 0.49%, range 0.021% to 5.9%; see FIG. **3** and table S3). All but one of the cases with detectable mutations were epithelial tumors; the exception was a dysgerminoma, a malignant germ cell tumor of the ovary (table S3). As with endometrial cancers, there was no statistically significant correlation between the fraction of mutant alleles and histopathologic criteria. However, most ovarian cancers are detected only at an advanced stage, and this was reflected in the patients available in our cohort.

EXAMPLE 5

A Genetic Test for Screening Purposes

[0062] The results described above document that mutant DNA molecules from most endometrial cancers and some ovarian cancers can be found in routinely collected Pap smears. However, in all 46 cases depicted in FIG. **3**, a specific mutation was known to occur in the tumor, and an assay was subsequently designed to determine whether that mutation was also present in the corresponding Pap smears. In a screening setting, there obviously would be no known tumor prior to the test. We therefore designed a prototype test based on Safe-SeqS that could be used in a screening setting (FIG. **2**).

[0063] This multiplexed approach included 50 primer pairs that amplified segments of 241 to 296 bp containing frequently mutated regions of DNA. The regions to be amplified were chosen from the results described in Section I and included exons from APC, AKT1, BRAF, CTNNB1, EGFR, FBXW7, KRAS, PIK3CA, PPP2R1A, PTEN, and TP53. In control experiments, 46 of the 50 amplicons were shown to provide information on a minimum of 2,500 templates; the number of templates sequenced can be determined directly from SafeSeqS-based sequencing (FIG. **2**). Given the accuracy of SafeSeqS, this number was adequate to comfortably detect mutations existing in >0.1% of template molecules (30). The regions covered by these 46 amplicons (table S5), encompassing 10,257 bp, were predicted to be able to detect at least one mutation in >90% of either endometrial or ovarian cancers.

[0064] This test was applied to Pap smears of 14 cases—twelve endometrial and two ovarian—as well as 14 Pap smears collected from normal women. The 14 cancer cases were arbitrarily chosen from those which had mutant allele fractions >0.1% (table S3) and therefore above the detection limit of the multiplexed assay. In all 14 Pap smears from women with cancer, the mutation expected to be present (table S3) was identified (FIG. **4** and table S6). The fraction of mutant alleles in the multiplexed test was similar to that observed in the original analysis of the same samples using only one Safe-SeqS primer pair per amplicon (table S3 and table S6). Importantly, no mutations were detected in the 14 Pap smears from women without cancer (FIG. **4**; see Materials and Methods).

EXAMPLE 6

Materials and Methods

Patient Samples

[0065] All samples for this study were obtained using protocols approved by the Institutional Review Boards of The Johns Hopkins Medical Institutions (Baltimore, Md.), Memorial Sloan Kettering Cancer Center (New York, N.Y.), University of Sao Paulo (Sao Paulo, Brazil), and ILSbio, LLC (Chestertown, Md.). Demographic, clinical and pathologic staging data was collected for each case. All histopathology was centrally re-reviewed by board-certified pathologists. Staging was based on 2009 FIGO criteria (38).

[0066] Fresh-frozen tissue specimens of surgically resected neoplasms of the ovary and endometrium were analyzed by frozen section to assess neoplastic cellularity by a board-certified pathologist. Serial frozen sections were used to guide the trimming of Optimal Cutting Temperature (OCT) compound embedded frozen tissue blocks to enrich the fraction of neoplastic cells for DNA extraction.

[0067] Formalin-fixed paraffin embedded (FFPE) tissue samples were assessed by a board-certified pathologist (Propath LLC, Dallas, Tex.) for tumor cellularity and to demarcate area of high tumor cellularity. Tumor tissue from serial 10 micron sections on slides from the original tumor block were macrodissected with a razorblade to enrich the fraction of neoplastic cells for DNA extraction.

[0068] The source of normal DNA was matched whole blood or non-neoplastic normal adjacent tissue.

[0069] Liquid-based Pap smears were collected using cervical brushes and transport medium from Digene HC2 DNA Collection Device (Qiagen) or ThinPrep 2000 System (Hologic) and stored using the manufacturer's recommendations.

[0070] Unless otherwise indicated, all patient-related values are reported as mean ±1 standard deviation.

DNA Extraction

[0071] DNA was purified from tumor and normal tissue as well as liquid-based Pap Smears using an AllPrep kit (Qiagen) according to the manufacturer's instructions. DNA was purified from tumor tissue by adding 3 mL RLTM buffer (Qiagen) and then binding to an AllPrep DNA column (Qiagen) following the manufacturer's protocol. DNA was purified from Pap smear liquids by adding five volumes of RLTM buffer when the amount of liquid was less than 3 mL. When the amount of liquid was >3 mL, the cells and cell fragments were pelleted at 1,000×g for five minutes and the pellets were dissolved in 3 mL RLTM buffer. DNA was quantified in all cases with qPCR, employing the primers and conditions previously described (39).

Microsatellite Instability Testing

[0072] Microsatellite instability was detected using the MSI Analysis System (Promega), containing five mononucleotide repeats (BAT-25, BAT-26, NR-21, NR-24 and MONO-27) and two pentanucleotide repeat loci, per the

manufacturer's instructions. Following amplification, the fluorescent PCR products were sized on an Applied Biosystems 3130 capillary electrophoresis instrument (Invitrogen). Tumor samples were designated as follows: MSI-high if two or more mononucleotides varied in length compared to the germline DNA; MSI-low if only one locus varied; and microsatellite stable (MSS) if there was no variation compared to the germline. Pentanucleotide loci confirmed identity in all cases.

Preparation of Illumina DNA Libraries and Capture for Exomic Sequencing

[0073] Preparation of Illumina genomic DNA libraries for exomic and targeted DNA captures was performed according to the manufacturer's recommendations. Briefly, 1-3 μg of genomic DNA was used for library preparation using the TruSeqDNA Sample Preparation Kit (Illumina). The DNA was acoustically sheared (Covaris) to a target size of ~200 bp. The fragments were subsequently end-repaired to convert overhangs into blunt ends. A single "A" nucleotide was then added to the 3' ends of blunt fragments to prevent them from later self-ligation; a corresponding "T" on the 3' end of adaptor molecules provided the complementary overhang. Following ligation to adaptors, the library was amplified with 8-14 cycles of PCR to ensure yields of 0.5 and 4 μg for exomic and targeted gene captures, respectively.

[0074] Exomic capture was performed with the SureSelect Human Exome Kit V 4.0 (Agilent) according to the manufacturer's protocol, with the addition of TruSeq index-specific blocks in the hybridization mixture (AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC-XXXXX-ATCTCGTATGCCGTCTTCTGCTTGT (SEQ ID NO: 1), where the six base pair "XXXXXX" denotes one of 12 sample-specific indexes).

Targeted Gene Enrichment

[0075] Targeted gene enrichment was performed by modifications of previously described methods (40, 41). In brief, targeted regions of selected oncogenes and tumor suppressor genes were synthesized as oligonucleotide probes by Agilent Technologies. Probes of 36 bases were designed to capture both the plus and the minus strand of the DNA and had a 33-base overlap. The oligonucleotides were cleaved from the chip by incubating with 3 mL of 35% ammonium hydroxide at room temperate for five hours. The solution was transferred to two 2-ml tubes, dried under vacuum, and redissolved in 400 μL of ribonuclease (RNase)—and deoxyribonuclease (DNase)—free water. Five microliters of the solution was used for PCR amplification with primers complementary to the 12-base sequence common to all probes: 5'-TGATCCCGCGACGA*C-3' (SEQ ID NO: 2) and 5'-GACCGCGACTCCAG*C-3' (SEQ ID NO: 3), with * indicating a phosphorothioate bond. The PCR products were purified with a MinElute Purification Column (Qiagen), end-repaired with End-IT DNA End-Repair Kit (Epicentre), and then purified with a MinElute Purification Column (Qiagen). The PCR products were ligated to form concatamers as described (40).

[0076] The major difference between the protocol described in (40, 41) and the one used in the present study involved the amplification of the ligated PCR products and the solid phase capture method. The modifications were as follows: 50 ng of ligated PCR product was amplified using the REPLI-g Midi Kit (Qiagen) with the addition of 2.5 nmol Biotin-dUTP (Roche) in a 27.5 μL reaction. The reaction was incubated at 30° C. for 16 hours, the polymerase was inactivated at 65° C. for 3 mins. The amplified probes were purified with QiaQuick PCR Purification Columns (Qiagen). For capture, 4-5 μg of library DNA was incubated with 1 μg of the prepared probes in a hybridization mixture as previously described (40). The biotinylated probes and captured library sequences were subsequently purified using 500 μg Dynabeads® MyOne Streptavidin (Invitrogen). After washing as per the manufacturer's recommendations, the captured sequences were eluted with 0.1 M NaOH and then neutralized with 1M Tris-HCl (pH 7.5). Neutralized DNA was desalted and concentrated using a QIAquick MinElute Column (Qiagen) in 20 μL. The elute was amplified in a 100 μL Phusion Hot Start II (Thermo Scientific) reaction containing 1×Phusion HF buffer, 0.25 mM dNTPs, 0.5 μM each forward and reverse TruSeq primers, and 2 U polymerase with the following cycling conditions: 98° C. for 30 s; 14 cycles of 98° C. for 10 s, 60° C. for 30 s, 72° C. for 30 s; and 72° C. for 5 min. The amplified pool containing enriched target sequences was purified using an Agencourt AMPure XP system (Beckman) and quantified using a 2100 Bioanalyzer (Agilent).

Next-Generation Sequencing and Somatic Mutation Identification

[0077] After capture of targeted sequences, paired-end sequencing using an Illumina GA IIx Genome Analyzer provided 2×75 base reads from each fragment. The sequence tags that passed filtering were aligned to the human genome reference sequence (hg18) and subsequent variant-calling analysis was performed using the ELANDv2 algorithm in the CASAVA 1.6 software (Illumina). Known polymorphisms recorded in dbSNP were removed from the analysis. Identification of high confidence mutations was performed as described previously (24).

Assessment of Low-Frequency Mutations

[0078] Primer Design. We attempted to design primer pairs to detect mutations in the 46 cancers described in the text. Primers were designed as described (30), using Primer3. (42) Sixty percent of the primers amplified the expected fragments; in the other 40%, a second or third set of primers had to be designed to reduce primer dimers or non-specific amplification.

[0079] Sequencing Library Preparation. Templates were amplified as described previously (30), with modifications that will be described in full elsewhere. In brief, each strand of each template molecule was encoded with a 14 base unique identifier (UID)—comprised of degenerate N bases (equal probability of being an A, C, G, or T)—using two to four cycles of amplicon-specific PCR (UID assignment PCR cycles, see FIG. **2**). While both forward and reverse gene-specific primers contained universal tag sequences at their 5' ends—providing the primer binding sites for the second-round amplification—only the forward primer contained the UID, positioned between the 5' universal tag and the 3' gene-specific sequences (four N's were included in the reverse primer to facilitate sequencing done on paired-end libraries) (table S4). The UID assignment PCR cycles included Phusion Hot Start II (Thermo Scientific) in a 50 μL reaction containing 1×Phusion HF buffer, 0.25 mM dNTPs,

0.5 μM each of forward (containing 14 N's) and reverse primers, and 2 U of polymerase. Carryover of residual UID-containing primers to the second-round amplification, which can complicate template quantification (30), was minimized through exonuclease digestion at 370 C to degrade unincorporated primers and subsequent purification with AMPure XP beads (Beckman) and elution in 10 μL TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

[0080] The eluted templates were amplified in a second-round PCR using primers containing the grafting sequences necessary for hybridization to the Illumina GA IIx flow cell at their 5' ends (FIG. **2**) and two terminal 3' phosphorothioates to protect them from residual exonuclease activity (30). The reverse amplification primer additionally contained an index sequence between the 5' grafting and 3' universal tag sequences to enable the PCR products from multiple individuals to be simultaneously analyzed in the same flow cell compartment of the sequencer (30). The second-round amplification reactions contained 1×Phusion HF buffer, 0.25 mM dNTPs, 0.5 μM each of forward and reverse primers, and 2 U of polymerase in a total of 50 4. After an initial heat activation step at 980 C for 2 minutes, twenty-three cycles of PCR were performed using the following cycling conditions: 980 C for 10 s, 650 C for 15 s, and 720 C for 15 s. The multiplexed assay was performed in similar fashion utilizing six independent amplifications per sample with the primers described in table S5. The PCR products were purified using AMPure XP beads and used directly for sequencing on either the Illumina MiSeq or GA IIx instruments, with equivalent results.

[0081] Data Analysis. High quality sequence reads were analyzed as previously described. (30) Briefly, reads in which each of the 14 bases comprising the UID (representing one original template strand; see FIG. **2**) had a quality score ≥15 were grouped by their UID. Only the UIDs supported by more than one read were retained for further analysis. The template-specific portion of the reads that contained the sequence of an expected amplification primer was matched to a reference sequence set using a custom script (available from the authors upon request). Artifactual mutations—introduced during the sample preparation and/or sequencing steps—were eliminated by requiring that >50% of reads sharing the same UID contained the identical mutation (a "supermutant;" see FIG. **2**). For the 46 assays querying a single amplicon, we required that the fraction of mutant alleles was significantly different from the background mutation levels determined from a negative control ($P<0.001$, binomial test). As mutations are not known a priori in a screening environment, we used a more agnostic metric to detect mutations in the multiplexed assay. A threshold supermutant frequency was defined for each sample as equaling the mean frequency of all supermutants plus six standard deviations of the mean. Only supermutants exceeding this threshold were designated as mutations and reported in FIG. **4** and table S6.

REFERENCES

[0082] The disclosure of each reference cited is expressly incorporated herein.

[0083] 1. M. Arbyn, A. Anttila, J. Jordan, G. Ronco, U. Schenck, N. Segnan, H. Wiener, A. Herbert, L. von Karsa, European Guidelines for Quality Assurance in Cervical Cancer Screening. Second edition—summary document. *Ann Oncol* 21, 448-458 (2010).

[0084] 2. R. M. DeMay, *Practical principles of cytopathology*. (ASCP Press, Chicago, 2010), pp. xi, 402 p.

[0085] 3. F. Bray, J. S. Ren, E. Masuyer, J. Ferlay, Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *Int J Cancer*, (2012).

[0086] 4. J. Ferlay, H. R. Shin, F. Bray, D. Forman, C. Mathers, D. M. Parkin, *GLOBOCAN* 2008 *v*2.0, *Cancer Incidence and Mortality Worldwide* (IARC CancerBase No. 10, Lyon, France, 2010).

[0087] 5. S. B. Sams, H. S. Currens, S. S. Raab, Liquid-based Papanicolaou tests in endometrial carcinoma diagnosis. Performance, error root cause analysis, and quality improvement. *Am J Clin Pathol* 137, 248-254 (2012).

[0088] 6. P. Smith, O. Bakos, G. Heimer, U. Ulmsten, Transvaginal ultrasound for identifying endometrial abnormality. *Acta Obstet Gynecol Scand* 70, 591-594 (1991).

[0089] 7. H. Mitchell, G. Giles, G. Medley, Accuracy and survival benefit of cytological prediction of endometrial carcinoma on routine cervical smears. *Int J Gynecol Pathol* 12, 34-40 (1993).

[0090] 8. K. J. Carlson, S. J. Skates, D. E. Singer, Screening for ovarian cancer. *Ann Intern Med* 121, 124-132 (1994).

[0091] 9. H. Meden, A. Fattahi-Meibodi, CA 125 in benign gynecological conditions. *Int J Biol Markers* 13, 231-237 (1998).

[0092] 10. S. S. Buys, E. Partridge, A. Black, C. C. Johnson, L. Lamerato, C. Isaacs, D. J. Reding, R. T. Greenlee, L. A. Yokochi, B. Kessel, E. D. Crawford, T. R. Church, G. L. Andriole, J. L. Weissfeld, M. N. Fouad, D. Chia, B. O'Brien, L. R. Ragard, J. D. Clapp, J. M. Rathmell, T. L. Riley, P. Hartge, P. F. Pinsky, C. S. Zhu, G. Izmirlian, B. S. Kramer, A. B. Miller, J. L. Xu, P. C. Prorok, J. K. Gohagan, C. D. Berg, Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA* 305, 2295-2303 (2011).

[0093] 11. ACOG Practice Bulletin. Clinical Management Guidelines for Obstetrician-Gynecologists. Number 60, March 2005. Pregestational diabetes mellitus. *Obstet Gynecol* 105, 675-685 (2005).

[0094] 12. E. Partridge, A. R. Kreimer, R. T. Greenlee, C. Williams, J. L. Xu, T. R. Church, B. Kessel, C. C. Johnson, J. L. Weissfeld, C. Isaacs, G. L. Andriole, S. Ogden, L. R. Ragard, S. S. Buys, Results from four rounds of ovarian cancer screening in a randomized trial. *Obstet Gynecol* 113, 775-782 (2009).

[0095] 13. American Cancer Society. Detailed guide: ovarian cancer—can ovarian cancer be found early? (Available at http://www.cancer.org/Cancer/OvarianCancer/DetailedGuide/ovarian-cancer-detection).

[0096] 14. Screening for ovarian cancer: recommendation statement. U.S. Preventive Services Task Force. *Am Fam Physician* 71, 759-762 (2005).

[0097] 15. ACOG Committee Opinion: number 280, December 2002. The role of the generalist obstetrician-gynecologist in the early detection of ovarian cancer. *Obstet Gynecol* 100, 1413-1416 (2002).

[0098] 16. National Comprehensive Cancer Network Practice Guidelines in Oncology: ovarian cancer and genetic screening. (Available at http://www.nccn.org/professionals/physician_gls/PDF/genetics_screening.pdf).

[0099] 17. N. M. Lindor, G. M. Petersen, D. W. Hadley, A. Y. Kinney, S. Miesfeldt, K. H. Lu, P. Lynch, W. Burke, N. Press, Recommendations for the care of individuals with an inherited predisposition to Lynch syndrome: a systematic review. *JAMA* 296, 1507-1517 (2006).

[0100] 18. J. P. Marques, L. B. Costa, A. P. Pinto, A. F. Lima, M. E. Duarte, A. P. Barbosa, P. L. Medeiros, Atypical glandular cells and cervical cancer: systematic review. *Rev Assoc Med Bras* 57, 234-238 (2011).

[0101] 19. R. P. Insinga, A. G. Glass, B. B. Rush, Diagnoses and outcomes in cervical cancer screening: a population-based study. *Am J Obstet Gynecol* 191, 105-113 (2004).

[0102] 20. K. E. Sharpless, P. F. Schnatz, S. Mandavilli, J. F. Greene, J. I. Sorosky, Dysplasia associated with atypical glandular cells on cervical cytology. *Obstet Gynecol* 105, 494-500 (2005).

[0103] 21. C. P. DeSimone, M. E. Day, M. M. Tovar, C. S. Dietrich, 3rd, M. L. Eastham, S. C. Modesitt, Rate of pathology from atypical glandular cell Pap tests classified by the Bethesda 2001 nomenclature. *Obstet Gynecol* 107, 1285-1291 (2006).

[0104] 22. C. S. Geier, M. Wilson, W. Creasman, Clinical evaluation of atypical glandular cells of undetermined significance. *Am J Obstet Gynecol* 184, 64-69 (2001).

[0105] 23. Integrated genomic analyses of ovarian carcinoma. *Nature* 474, 609-615 (2011).

[0106] 24. S. Jones, T. L. Wang, M. Shih le, T. L. Mao, K. Nakayama, R. Roden, R. Glas, D. Slamon, L. A. Diaz, Jr., B. Vogelstein, K. W. Kinzler, V. E. Velculescu, N. Papadopoulos, Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. *Science* 330, 228-231 (2010).

[0107] 25. S. Jones, T. L. Wang, R. J. Kurman, K. Nakayama, V. E. Velculescu, B. Vogelstein, K. W. Kinzler, N. Papadopoulos, M. Shih le, Low-grade serous carcinomas of the ovary contain very few point mutations. *J Pathol* 226, 413-420 (2012).

[0108] 26. S. A. Forbes, N. Bindal, S. Bamford, C. Cole, C. Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, J. W. Teague, P. J. Campbell, M. R. Stratton, P. A. Futreal, COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic *Acids Res* 39, D945-950 (2011).

[0109] 27. E. Barrow, L. Robinson, W. Alduaij, A. Shenton, T. Clancy, F. Lalloo, J. Hill, D. G. Evans, Cumulative lifetime incidence of extracolonic cancers in Lynch syndrome: a report of 121 families with proven mutations. *Clin Genet* 75, 141-149 (2009).

[0110] 28. K. Oda, D. Stokoe, Y. Taketani, F. McCormick, High frequency of coexistent mutations of PIK3CA and PTEN genes in endometrial carcinoma. *Cancer Res* 65, 10669-10673 (2005).

[0111] 29. E. Kuhn, R. C. Wu, B. Guan, G. Wu, J. Zhang, Y. Wang, L. Song, X. Yuan, L. Wei, R. B. Roden, K. T. Kuo, K. Nakayama, B. Clarke, P. Shaw, N. Olvera, R. J. Kurman, D. A. Levine, T. L. Wang, I. M. Shih, Identification of Molecular Pathway Aberrations in Uterine Serous Carcinoma by Genome-wide Analyses. *J Natl Cancer Inst*, (2012).

[0112] 30. I. Kinde, J. Wu, N. Papadopoulos, K. W. Kinzler, B. Vogelstein, Detection and quantification of rare mutations with massively parallel sequencing. *Proc Natl Acad Sci USA* 108, 9530-9535 (2011).

[0113] 31. H. F. Traut, G. N. Papanicolaou, Cancer of the Uterus: The Vaginal Smear in Its Diagnosis. *Cal West Med* 59, 121-122 (1943).

[0114] 32. B. Vogelstein, K. W. Kinzler, Cancer genes and the pathways they control. *Nat Med* 10, 789-799 (2004).

[0115] 33. J. M. Cooper, M. L. Erickson, Endometrial sampling techniques in the diagnosis of abnormal uterine bleeding. *Obstet Gynecol Clin North Am* 27, 235-244 (2000).

[0116] 34. C. C. Gunderson, A. N. Fader, K. A. Carson, R. E. Bristow, Oncologic and reproductive outcomes with progestin therapy in women with endometrial hyperplasia and grade 1 adenocarcinoma: a systematic review. *Gynecol Oncol* 125, 477-482 (2012).

[0117] 35. R. E. Bristow, R. S. Tomacruz, D. K. Armstrong, E. L. Trimble, F. J. Montz, Survival effect of maximal cytoreductive surgery for advanced ovarian carcinoma during the platinum era: a meta-analysis. *J Clin Oncol* 20, 1248-1259 (2002).

[0118] 36. M. H. Mayrand, E. Duarte-Franco, I. Rodrigues, S. D. Walter, J. Hanley, A. Ferenczy, S. Ratnam, F. Coutlee, E. L. Franco, Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. *N Engl J Med* 357, 1579-1588 (2007).

[0119] 37. P. Naucler, W. Ryd, S. Tornberg, A. Strand, G. Wadell, K. Elfgren, T. Radberg, B. Strander, B. Johansson, O. Forslund, B. G. Hansson, E. Rylander, J. Dillner, Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med* 357, 1589-1597 (2007).

[0120] 38. S. Pecorelli, Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. *Int J Gynaecol Obstet* 105, 103-104 (2009).

[0121] 39. C. Rago, D. L. Huso, F. Diehl, B. Karim, G. Liu, N. Papadopoulos, Y. Samuels, V. E. Velculescu, B. Vogelstein, K. W. Kinzler, L. A. Diaz, Jr., Serial assessment of human tumor burdens in mice by the analysis of circulating DNA. *Cancer Res* 67, 9364-9370 (2007).

[0122] 40. J. Wu, H. Matthaei, A. Maitra, M. Dal Molin, L. D. Wood, J. R. Eshleman, M. Goggins, M. I. Canto, R. D. Schulick, B. H. Edil, C. L. Wolfgang, A. P. Klein, L. A. Diaz, Jr., P. J. Allen, C. M. Schmidt, K. W. Kinzler, N. Papadopoulos, R. H. Hruban, B. Vogelstein, Recurrent GNAS mutations define an unexpected pathway for pancreatic cyst development. *Sci Transl Med* 3, 92ra66 (2011).

[0123] 41. J. He, J. Wu, Y. Jiao, N. Wagner-Johnston, R. F. Ambinder, L. A. Diaz, Jr., K. W. Kinzler, B. Vogelstein, N. Papadopoulos, IgH gene rearrangements as plasma biomarkers in Non-Hodgkin's lymphoma patients. *Oncotarget* 2, 178-185 (2011).

[0124] 42. S. Rozen, H. Skaletsky, Primer3 on the WWW for general users and for biologist programmers. Methods *Mol Biol* 132, 365-386 (2000).

[0125] 43. N. Howlader, A. M. Noone, M. Krapcho, N. Neyman, R. Aminou, S. F. Altekruse, C. L. Kosary, J. Ruhl, Z. Tatalovich, H. Cho, A. Mariotto, M. P. Eisner, D. R. Lewis, H. S. Chen, E. J. Feuer, K. A. Cronin, *SEER Cancer Statistics Review,* 1975-2009 (National Cancer Institute. Bethesda, Md., 2012).

[0126] 44. A. Malpica, M. T. Deavers, K. Lu, D. C. Bodurka, E. N. Atkinson, D. M. Gershenson, E. G. Silva, Grading ovarian serous carcinoma using a two-tier system. *Am J Surg Pathol* 28, 496-504 (2004).

[0127] 45. L. A. G. Ries, J. L. Young, G. E. Keel, M. P. Eisner, Y. D. Lin, M-J. Homer, *SEER Survival Monograph: Cancer Survival Among Adults: US SEER Program,* 1988-2001, *Patient and Tumor Characteristics* (NIH Pub. No. 07-6215. National Cancer Institute, Bethesda, Md., 2007).

[0128] 46. C. A. Hamilton, M. K. Cheung, K. Osann, L. Chen, N. N. Teng, T. A. Longacre, M. A. Powell, M. R. Hendrickson, D. S. Kapp, J. K. Chan, Uterine papillary serous and clear cell carcinomas predict for poorer survival compared to grade 3 endometrioid corpus cancers. *Br J Cancer* 94, 642-646 (2006).

SEQUENCE LISTING

Sequence total quantity: 191

SEQ ID NO: 1 moltype = DNA length = 65
FEATURE Location/Qualifiers
source 1..65
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 1
agatcggaag agcacacgtc tgaactccag tcacnnnnnn atctcgtatg ccgtcttctg 60
cttgt 65

SEQ ID NO: 2 moltype = DNA length = 15
FEATURE Location/Qualifiers
source 1..15
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 2
tgatcccgcg acgac 15

SEQ ID NO: 3 moltype = DNA length = 15
FEATURE Location/Qualifiers
source 1..15
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 3
gaccgcgact ccagc 15

SEQ ID NO: 4 moltype = DNA length = 60
FEATURE Location/Qualifiers
source 1..60
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 4
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngatcc aatccatttt tgttgtccag 60

SEQ ID NO: 5 moltype = DNA length = 45
FEATURE Location/Qualifiers
source 1..45
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 5
cacacaggaa acagctatga ccatgtgagc aagaggcttt ggagt 45

SEQ ID NO: 6 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 6
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnggcca agacctgccc tg 52

SEQ ID NO: 7 moltype = DNA length = 47
FEATURE Location/Qualifiers
source 1..47
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 7
cacacaggaa acagctatga ccatgtgctg tgactgcttg tagatgg 47

SEQ ID NO: 8 moltype = DNA length = 56
FEATURE Location/Qualifiers
source 1..56
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 8
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnccacc tcctcaaaca gctcaa 56

SEQ ID NO: 9 moltype = DNA length = 46

-continued

FEATURE	Location/Qualifiers	
source	1..46	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 9		
cacacaggaa acagctatga ccatgtgcag cttgcttagg tccact		46
SEQ ID NO: 10	moltype = DNA length = 55	
FEATURE	Location/Qualifiers	
source	1..55	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 10		
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntggcc atctacaagc agtca		55
SEQ ID NO: 11	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 11		
cacacaggaa acagctatga ccatgnnnnt caccatcgct atctgagca		49
SEQ ID NO: 12	moltype = DNA length = 55	
FEATURE	Location/Qualifiers	
source	1..55	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 12		
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncattg gtgatgattc gatgg		55
SEQ ID NO: 13	moltype = DNA length = 45	
FEATURE	Location/Qualifiers	
source	1..45	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 13		
cacacaggaa acagctatga ccatgctgcc tggctcagaa ttcac		45
SEQ ID NO: 14	moltype = DNA length = 54	
FEATURE	Location/Qualifiers	
source	1..54	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 14		
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnccctt tcttgcggag attc		54
SEQ ID NO: 15	moltype = DNA length = 45	
FEATURE	Location/Qualifiers	
source	1..45	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 15		
cacacaggaa acagctatga ccatgctact gggacggaac agctt		45
SEQ ID NO: 16	moltype = DNA length = 55	
FEATURE	Location/Qualifiers	
source	1..55	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 16		
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnggaag agaatctccg caaga		55
SEQ ID NO: 17	moltype = DNA length = 45	
FEATURE	Location/Qualifiers	
source	1..45	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 17		
cacacaggaa acagctatga ccatggcttc ttgtcctgct tgctt		45
SEQ ID NO: 18	moltype = DNA length = 56	
FEATURE	Location/Qualifiers	
source	1..56	
	mol_type = other DNA	
	organism = synthetic construct	

-continued

SEQUENCE: 18
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnggcct gtctcaatat cccaaa 56

SEQ ID NO: 19 moltype = DNA length = 48
FEATURE Location/Qualifiers
source 1..48
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 19
cacacaggaa acagctatga ccatgttgtt tttctgtttc tccctctg 48

SEQ ID NO: 20 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 20
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncaagg cactcttgcc tacg 54

SEQ ID NO: 21 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 21
cacacaggaa acagctatga ccatgcattt tcattatttt tattataagg cctg 54

SEQ ID NO: 22 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 22
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnctgtg gtagtggcac cagaa 55

SEQ ID NO: 23 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 23
cacacaggaa acagctatga ccatgnnnna agcggctgtt agtcactgg 49

SEQ ID NO: 24 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 24
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnggccc ctgtcatctt ctgt 54

SEQ ID NO: 25 moltype = DNA length = 45
FEATURE Location/Qualifiers
source 1..45
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 25
cacacaggaa acagctatga ccatggactt ggctgtccca gaatg 45

SEQ ID NO: 26 moltype = DNA length = 58
FEATURE Location/Qualifiers
source 1..58
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 26
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncaaga aatcgatagc atttgcag 58

SEQ ID NO: 27 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 27
cacacaggaa acagctatga ccatgtttat ttgctttgtc aagatcattt tt 52

SEQ ID NO: 28 moltype = DNA length = 60

-continued

FEATURE Location/Qualifiers
source 1..60
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 28
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnaggaa atatctgctt gctcattcaa 60

SEQ ID NO: 29 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 29
cacacaggaa acagctatga ccatggaagc agatactaag caggacacta tatc 54

SEQ ID NO: 30 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 30
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnttccc ttggattctg acaca 55

SEQ ID NO: 31 moltype = DNA length = 45
FEATURE Location/Qualifiers
source 1..45
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 31
cacacaggaa acagctatga ccatgagcac cattcgttga taggc 45

SEQ ID NO: 32 moltype = DNA length = 57
FEATURE Location/Qualifiers
source 1..57
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 32
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncactg gcagcaacag tcttacc 57

SEQ ID NO: 33 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 33
cacacaggaa acagctatga ccatggattg cctttaccac tcagagaag 49

SEQ ID NO: 34 moltype = DNA length = 57
FEATURE Location/Qualifiers
source 1..57
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 34
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnttgca gcaattcact gtaaagc 57

SEQ ID NO: 35 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 35
cacacaggaa acagctatga ccatgccgat gtaataaata tgcacatatc attac 55

SEQ ID NO: 36 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 36
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncaagg cactcttgcc tacg 54

SEQ ID NO: 37 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct

-continued

SEQUENCE: 37
cacacaggaa acagctatga ccatgcattt tcattatttt tattataagg cctg 54

SEQ ID NO: 38 moltype = DNA length = 57
FEATURE Location/Qualifiers
source 1..57
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 38
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncactg gcagcaacag tcttacc 57

SEQ ID NO: 39 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 39
cacacaggaa acagctatga ccatggattg cctttaccac tcagagaag 49

SEQ ID NO: 40 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 40
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncaagg cactcttgcc tacg 54

SEQ ID NO: 41 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 41
cacacaggaa acagctatga ccatgcattt tcattatttt tattataagg cctg 54

SEQ ID NO: 42 moltype = DNA length = 58
FEATURE Location/Qualifiers
source 1..58
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 42
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnagctc aaagcaattt ctacacga 58

SEQ ID NO: 43 moltype = DNA length = 53
FEATURE Location/Qualifiers
source 1..53
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 43
cacacaggaa acagctatga ccatgnnnng cacttacctg tgactccata gaa 53

SEQ ID NO: 44 moltype = DNA length = 60
FEATURE Location/Qualifiers
source 1..60
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 44
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntcttt tgatgacatt gcatacattc 60

SEQ ID NO: 45 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 45
cacacaggaa acagctatga ccatgnnnna ctccaaagcc tcttgctca 49

SEQ ID NO: 46 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 46
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncagtt gcaaaccaga cctca 55

SEQ ID NO: 47 moltype = DNA length = 49

-continued

FEATURE	Location/Qualifiers	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 47		
cacacaggaa acagctatga ccatgtgtgg agtatttgga tgacagaaa		49
SEQ ID NO: 48	moltype = DNA length = 53	
FEATURE	Location/Qualifiers	
source	1..53	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 48		
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngtggc aagtggctcc tga		53
SEQ ID NO: 49	moltype = DNA length = 45	
FEATURE	Location/Qualifiers	
source	1..45	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 49		
cacacaggaa acagctatga ccatgnnnnc atgggcggca tgaac		45
SEQ ID NO: 50	moltype = DNA length = 57	
FEATURE	Location/Qualifiers	
source	1..57	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 50		
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnactgg cagcaacagt cttacct		57
SEQ ID NO: 51	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 51		
cacacaggaa acagctatga ccatgnnnnc ctcaggattg cctttacca		49
SEQ ID NO: 52	moltype = DNA length = 53	
FEATURE	Location/Qualifiers	
source	1..53	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 52		
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnagtcc ggcttggagg atg		53
SEQ ID NO: 53	moltype = DNA length = 45	
FEATURE	Location/Qualifiers	
source	1..45	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 53		
cacacaggaa acagctatga ccatgtcccc actcctcctt tcttc		45
SEQ ID NO: 54	moltype = DNA length = 54	
FEATURE	Location/Qualifiers	
source	1..54	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 54		
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnggaaa gggacgaact ggtg		54
SEQ ID NO: 55	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 55		
cacacaggaa acagctatga ccatgnnnnt agggcctctt gtgccttta		49
SEQ ID NO: 56	moltype = DNA length = 55	
FEATURE	Location/Qualifiers	
source	1..55	
	mol_type = other DNA	
	organism = synthetic construct	

-continued

SEQUENCE: 56
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntcttc tgtcccttcc cagaa 55

SEQ ID NO: 57 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 57
cacacaggaa acagctatga ccatgnnnng acttggctgt cccagaatg 49

SEQ ID NO: 58 moltype = DNA length = 58
FEATURE Location/Qualifiers
source 1..58
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 58
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntcata ccaatttctc gattgagg 58

SEQ ID NO: 59 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 59
cacacaggaa acagctatga ccatgnnnnc ggctttttca acccttttt 49

SEQ ID NO: 60 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 60
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntggcc atctacaagc agtca 55

SEQ ID NO: 61 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 61
cacacaggaa acagctatga ccatgnnnnt caccatcgct atctgagca 49

SEQ ID NO: 62 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 62
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngggac ggaacagctt tgag 54

SEQ ID NO: 63 moltype = DNA length = 50
FEATURE Location/Qualifiers
source 1..50
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 63
cacacaggaa acagctatga ccatgnnnng cggagattct cttcctctgt 50

SEQ ID NO: 64 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 64
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncggtg taggagctgc tggt 54

SEQ ID NO: 65 moltype = DNA length = 48
FEATURE Location/Qualifiers
source 1..48
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 65
cacacaggaa acagctatga ccatgnnnna cccaggtcca gatgaagc 48

SEQ ID NO: 66 moltype = DNA length = 55

-continued

FEATURE	Location/Qualifiers
source	1..55
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 66	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntggcc atctacaagc agtca	55
SEQ ID NO: 67	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 67	
cacacaggaa acagctatga ccatgnnnnt caccatcgct atctgagca	49
SEQ ID NO: 68	moltype = DNA length = 53
FEATURE	Location/Qualifiers
source	1..53
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 68	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngtggc aagtggctcc tga	53
SEQ ID NO: 69	moltype = DNA length = 45
FEATURE	Location/Qualifiers
source	1..45
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 69	
cacacaggaa acagctatga ccatgnnnnc atgggcggca tgaac	45
SEQ ID NO: 70	moltype = DNA length = 57
FEATURE	Location/Qualifiers
source	1..57
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 70	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncgaaa agtgtttctg tcatcca	57
SEQ ID NO: 71	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 71	
cacacaggaa acagctatga ccatgnnnng cccctcctca gcatcttat	49
SEQ ID NO: 72	moltype = DNA length = 55
FEATURE	Location/Qualifiers
source	1..55
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 72	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncagtt gcaaaccaga cctca	55
SEQ ID NO: 73	moltype = DNA length = 53
FEATURE	Location/Qualifiers
source	1..53
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 73	
cacacaggaa acagctatga ccatgnnnnt gtggagtatt tggatgacag aaa	53
SEQ ID NO: 74	moltype = DNA length = 56
FEATURE	Location/Qualifiers
source	1..56
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 74	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnttata tttccccatg ccaatg	56
SEQ ID NO: 75	moltype = DNA length = 59
FEATURE	Location/Qualifiers
source	1..59
	mol_type = other DNA
	organism = synthetic construct

-continued

SEQUENCE: 75
cacacaggaa acagctatga ccatgnnnng gtgttttgaa atgtgtttta taatttaga 59

SEQ ID NO: 76 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 76
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntcttc tgtcccttcc cagaa 55

SEQ ID NO: 77 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 77
cacacaggaa acagctatga ccatgnnnng acttggctgt cccagaatg 49

SEQ ID NO: 78 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 78
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngaaaa agccgaaggt cacaa 55

SEQ ID NO: 79 moltype = DNA length = 51
FEATURE Location/Qualifiers
source 1..51
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 79
cacacaggaa acagctatga ccatgnnnnc tcaagaagca gaaagggaag a 51

SEQ ID NO: 80 moltype = DNA length = 60
FEATURE Location/Qualifiers
source 1..60
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 80
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngatcc aatccatttt tgttgtccag 60

SEQ ID NO: 81 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 81
cacacaggaa acagctatga ccatgnnnnt gagcaagagg ctttggagt 49

SEQ ID NO: 82 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 82
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntgtga tgatggtgag gatgg 55

SEQ ID NO: 83 moltype = DNA length = 57
FEATURE Location/Qualifiers
source 1..57
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 83
cacacaggaa acagctatga ccatgnnnnt ccactacaac tacatgtgta acagttc 57

SEQ ID NO: 84 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 84
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnctccg tcatgtgctg tgact 55

SEQ ID NO: 85 moltype = DNA length = 48

-continued

FEATURE	Location/Qualifiers
source	1..48
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 85	
cacacaggaa acagctatga ccatgnnnnc agctgtgggt tgattcca	48
SEQ ID NO: 86	moltype = DNA length = 55
FEATURE	Location/Qualifiers
source	1..55
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 86	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntggcc atctacaagc agtca	55
SEQ ID NO: 87	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 87	
cacacaggaa acagctatga ccatgnnnnt caccatcgct atctgagca	49
SEQ ID NO: 88	moltype = DNA length = 56
FEATURE	Location/Qualifiers
source	1..56
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 88	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnccaat ccatttttgt tgtcca	56
SEQ ID NO: 89	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 89	
cacacaggaa acagctatga ccatgnnnnt gagcaagagg ctttggagt	49
SEQ ID NO: 90	moltype = DNA length = 53
FEATURE	Location/Qualifiers
source	1..53
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 90	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnctgca cagggcaggt ctt	53
SEQ ID NO: 91	moltype = DNA length = 53
FEATURE	Location/Qualifiers
source	1..53
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 91	
cacacaggaa acagctatga ccatgnnnnc tctgtctcct tcctcttcct aca	53
SEQ ID NO: 92	moltype = DNA length = 55
FEATURE	Location/Qualifiers
source	1..55
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 92	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntggcc atctacaagc agtca	55
SEQ ID NO: 93	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 93	
cacacaggaa acagctatga ccatgnnnnt caccatcgct atctgagca	49
SEQ ID NO: 94	moltype = DNA length = 55
FEATURE	Location/Qualifiers
source	1..55
	mol_type = other DNA
	organism = synthetic construct

-continued

SEQUENCE: 94
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncacgc aaatttcctt ccact 55

SEQ ID NO: 95 moltype = DNA length = 50
FEATURE Location/Qualifiers
source 1..50
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 95
cacacaggaa acagctatga ccatgnnnng gcctctgatt cctcactgat 50

SEQ ID NO: 96 moltype = DNA length = 60
FEATURE Location/Qualifiers
source 1..60
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 96
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngatcc aatccatttt tgttgtccag 60

SEQ ID NO: 97 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 97
cacacaggaa acagctatga ccatgnnnnt gagcaagagg ctttggagt 49

SEQ ID NO: 98 moltype = DNA length = 57
FEATURE Location/Qualifiers
source 1..57
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 98
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnactgg cagcaacagt cttacct 57

SEQ ID NO: 99 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 99
cacacaggaa acagctatga ccatgnnnnc ctcaggattg cctttacca 49

SEQ ID NO: 100 moltype = DNA length = 62
FEATURE Location/Qualifiers
source 1..62
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 100
gacgtaaaac gacggccagt nnnnnnnnnn nnnnaaagta acatttccaa tctactaatg 60
ct 62

SEQ ID NO: 101 moltype = DNA length = 50
FEATURE Location/Qualifiers
source 1..50
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 101
cacacaggaa acagctatga ccatgnnnnt gagaaaatcc ctgttcccac 50

SEQ ID NO: 102 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 102
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntggag cctcttacac ccagt 55

SEQ ID NO: 103 moltype = DNA length = 51
FEATURE Location/Qualifiers
source 1..51
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 103
cacacaggaa acagctatga ccatgnnnna aaaacactgg agtttcccaa a 51

-continued

SEQ ID NO: 104 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 104
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngaaaa agccgaaggt cacaa 55

SEQ ID NO: 105 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 105
cacacaggaa acagctatga ccatgnnnna tgcccccaag aatcctagt 49

SEQ ID NO: 106 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 106
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncatcc gtctactccc acgtt 55

SEQ ID NO: 107 moltype = DNA length = 48
FEATURE Location/Qualifiers
source 1..48
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 107
cacacaggaa acagctatga ccatgnnnna tcagctaccg ccaagtcc 48

SEQ ID NO: 108 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 108
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnttcgt ccctttccag cttta 55

SEQ ID NO: 109 moltype = DNA length = 56
FEATURE Location/Qualifiers
source 1..56
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 109
cacacaggaa acagctatga ccatgnnnng gaatccagtg tttcttttaa atacct 56

SEQ ID NO: 110 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 110
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnaccag ccctgtcgtc tctc 54

SEQ ID NO: 111 moltype = DNA length = 51
FEATURE Location/Qualifiers
source 1..51
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 111
cacacaggaa acagctatga ccatgnnnng ccctgacttt caactctgtc t 51

SEQ ID NO: 112 moltype = DNA length = 59
FEATURE Location/Qualifiers
source 1..59
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 112
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngcctc agattcactt ttatcacct 59

SEQ ID NO: 113 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA

-continued

organism = synthetic construct
SEQUENCE: 113
cacacaggaa acagctatga ccatgnnnna ccaggagcca ttgtctttg 49

SEQ ID NO: 114 moltype = DNA length = 60
FEATURE Location/Qualifiers
source 1..60
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 114
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntcacc acattacata cttaccatgc 60

SEQ ID NO: 115 moltype = DNA length = 51
FEATURE Location/Qualifiers
source 1..51
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 115
cacacaggaa acagctatga ccatgnnnna aggggatctc ttcctgtatc c 51

SEQ ID NO: 116 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 116
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntctgg atcccagaag gtgag 55

SEQ ID NO: 117 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 117
cacacaggaa acagctatga ccatgnnnng gccagtgctg tctctaagg 49

SEQ ID NO: 118 moltype = DNA length = 60
FEATURE Location/Qualifiers
source 1..60
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 118
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngtcca caaaatgatt ctgaattagc 60

SEQ ID NO: 119 moltype = DNA length = 51
FEATURE Location/Qualifiers
source 1..51
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 119
cacacaggaa acagctatga ccatgnnnna cgatacacgt ctgcagtcaa c 51

SEQ ID NO: 120 moltype = DNA length = 62
FEATURE Location/Qualifiers
source 1..62
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 120
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnacaga aatattttag aaagggacaa 60
ca 62

SEQ ID NO: 121 moltype = DNA length = 51
FEATURE Location/Qualifiers
source 1..51
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 121
cacacaggaa acagctatga ccatgnnnna gaaaataccc cctccatcaa c 51

SEQ ID NO: 122 moltype = DNA length = 60
FEATURE Location/Qualifiers
source 1..60
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 122
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngatcc aatccatttt tgttgtccag 60

-continued

SEQ ID NO: 123 moltype = DNA length = 53
FEATURE Location/Qualifiers
source 1..53
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 123
cacacaggaa acagctatga ccatgnnnnt ccaaactgac caaactgttc tta 53

SEQ ID NO: 124 moltype = DNA length = 57
FEATURE Location/Qualifiers
source 1..57
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 124
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngaaac ccaaaatctg ttttcca 57

SEQ ID NO: 125 moltype = DNA length = 50
FEATURE Location/Qualifiers
source 1..50
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 125
cacacaggaa acagctatga ccatgnnnng accataaccc accacagcta 50

SEQ ID NO: 126 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 126
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnctcct cccagagacc ccagt 55

SEQ ID NO: 127 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 127
cacacaggaa acagctatga ccatgnnnnc atgagcgctg ctcagatag 49

SEQ ID NO: 128 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 128
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncctag gaaggcaggg gagt 54

SEQ ID NO: 129 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 129
cacacaggaa acagctatga ccatgnnnnt gcatgttgct tttgtaccg 49

SEQ ID NO: 130 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 130
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnaccac ccgcacgtct gtag 54

SEQ ID NO: 131 moltype = DNA length = 48
FEATURE Location/Qualifiers
source 1..48
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 131
cacacaggaa acagctatga ccatgnnnna gccagtgctt gttgcttg 48

SEQ ID NO: 132 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54

-continued

	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 132	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncacac tgacgtgcct ctcc	54
SEQ ID NO: 133	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 133	
cacacaggaa acagctatga ccatgnnnnt tatctcccct ccccgtatc	49
SEQ ID NO: 134	moltype = DNA length = 55
FEATURE	Location/Qualifiers
source	1..55
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 134	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngattg tcagtgcgct tttcc	55
SEQ ID NO: 135	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 135	
cacacaggaa acagctatga ccatgnnnng ctaaggatgg gggttgcta	49
SEQ ID NO: 136	moltype = DNA length = 61
FEATURE	Location/Qualifiers
source	1..61
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 136	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntgact ttaccttatc aatgtctcga	60
a	61
SEQ ID NO: 137	moltype = DNA length = 48
FEATURE	Location/Qualifiers
source	1..48
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 137	
cacacaggaa acagctatga ccatgnnnng ctcgcccccc taatctct	48
SEQ ID NO: 138	moltype = DNA length = 55
FEATURE	Location/Qualifiers
source	1..55
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 138	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnggtac ttccggaacc tgtgc	55
SEQ ID NO: 139	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 139	
cacacaggaa acagctatga ccatgnnnnc cgagtcctag ggagaggag	49
SEQ ID NO: 140	moltype = DNA length = 58
FEATURE	Location/Qualifiers
source	1..58
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 140	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnttgtt aatggtggct ttttgttt	58
SEQ ID NO: 141	moltype = DNA length = 54
FEATURE	Location/Qualifiers
source	1..54
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 141	

-continued

cacacaggaa acagctatga ccatgnnnna aatgatctaa caatgctctt ggac 54

SEQ ID NO: 142 moltype = DNA length = 58
FEATURE Location/Qualifiers
source 1..58
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 142
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnncatgg aaggatgaga atttcaag 58

SEQ ID NO: 143 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 143
cacacaggaa acagctatga ccatgnnnnt ggctacgacc cagttacca 49

SEQ ID NO: 144 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 144
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnaaacc gtagctgccc tggta 55

SEQ ID NO: 145 moltype = DNA length = 50
FEATURE Location/Qualifiers
source 1..50
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 145
cacacaggaa acagctatga ccatgnnnnt gactgctctt ttcacccatc 50

SEQ ID NO: 146 moltype = DNA length = 56
FEATURE Location/Qualifiers
source 1..56
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 146
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntcatc ttgggcctgt gttatc 56

SEQ ID NO: 147 moltype = DNA length = 53
FEATURE Location/Qualifiers
source 1..53
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 147
cacacaggaa acagctatga ccatgnnnng atgagaggtg gatgggtagt agt 53

SEQ ID NO: 148 moltype = DNA length = 56
FEATURE Location/Qualifiers
source 1..56
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 148
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnttcag ggcatgaact acttgg 56

SEQ ID NO: 149 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 149
cacacaggaa acagctatga ccatgnnnna tcctcccctg catgtgtta 49

SEQ ID NO: 150 moltype = DNA length = 56
FEATURE Location/Qualifiers
source 1..56
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 150
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntccct cattgcactg tactcc 56

SEQ ID NO: 151 moltype = DNA length = 49
FEATURE Location/Qualifiers

-continued

source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 151
cacacaggaa acagctatga ccatgnnnng gtgcttagtg gccatttgt 49

SEQ ID NO: 152 moltype = DNA length = 57
FEATURE Location/Qualifiers
source 1..57
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 152
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntggtc tctcatggca ctgtact 57

SEQ ID NO: 153 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 153
cacacaggaa acagctatga ccatgnnnna ttagcaattt gagggacaaa cc 52

SEQ ID NO: 154 moltype = DNA length = 62
FEATURE Location/Qualifiers
source 1..62
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 154
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntgtat ctcactcgat aatctggatg 60
ac 62

SEQ ID NO: 155 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 155
cacacaggaa acagctatga ccatgnnnnt gtcacattat aaagattcag gcaat 55

SEQ ID NO: 156 moltype = DNA length = 59
FEATURE Location/Qualifiers
source 1..59
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 156
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnagttt gacagttaaa ggcatttcc 59

SEQ ID NO: 157 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 157
cacacaggaa acagctatga ccatgnnnnt gtccttattt tggatatttc tccc 54

SEQ ID NO: 158 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 158
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngaaga cccaggtcca gatga 55

SEQ ID NO: 159 moltype = DNA length = 48
FEATURE Location/Qualifiers
source 1..48
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 159
cacacaggaa acagctatga ccatgnnnna gaaatgcagg gggatacg 48

SEQ ID NO: 160 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct

-continued

SEQUENCE: 160
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngaggc ataactgcac ccttg 55

SEQ ID NO: 161 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 161
cacacaggaa acagctatga ccatgnnnng ggagtagatg gagcctggt 49

SEQ ID NO: 162 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 162
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntgctg gatttggttc taggg 55

SEQ ID NO: 163 moltype = DNA length = 50
FEATURE Location/Qualifiers
source 1..50
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 163
cacacaggaa acagctatga ccatgnnnnt gccacttgca aagtttcttc 50

SEQ ID NO: 164 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 164
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnggaag aacctggacc ctctg 55

SEQ ID NO: 165 moltype = DNA length = 50
FEATURE Location/Qualifiers
source 1..50
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 165
cacacaggaa acagctatga ccatgnnnnt tttgagagtc gttcgattgc 50

SEQ ID NO: 166 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 166
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntgcaa cctgttttgt gatgg 55

SEQ ID NO: 167 moltype = DNA length = 51
FEATURE Location/Qualifiers
source 1..51
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 167
cacacaggaa acagctatga ccatgnnnna ggaaaatgac aatgggaatg a 51

SEQ ID NO: 168 moltype = DNA length = 59
FEATURE Location/Qualifiers
source 1..59
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 168
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntgatt catcaggaga gcatttaag 59

SEQ ID NO: 169 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 169
cacacaggaa acagctatga ccatgnnnnt tgtttttctg tttctccctc tg 52

SEQ ID NO: 170 moltype = DNA length = 57

-continued

FEATURE	Location/Qualifiers
source	1..57
	mol_type = other DNA
	organism = synthetic construct

SEQUENCE: 170
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngatgg tatccatgtg gtgagtg 57

SEQ ID NO: 171	moltype = DNA length = 50
FEATURE	Location/Qualifiers
source	1..50
	mol_type = other DNA
	organism = synthetic construct

SEQUENCE: 171
cacacaggaa acagctatga ccatgnnnnt ttgtgatgct aaggctccat 50

SEQ ID NO: 172	moltype = DNA length = 53
FEATURE	Location/Qualifiers
source	1..53
	mol_type = other DNA
	organism = synthetic construct

SEQUENCE: 172
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnactgc cttccgggtc act 53

SEQ ID NO: 173	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct

SEQUENCE: 173
cacacaggaa acagctatga ccatgnnnna gcccaaccct tgtccttac 49

SEQ ID NO: 174	moltype = DNA length = 54
FEATURE	Location/Qualifiers
source	1..54
	mol_type = other DNA
	organism = synthetic construct

SEQUENCE: 174
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngaggc tgtcagtggg gaac 54

SEQ ID NO: 175	moltype = DNA length = 50
FEATURE	Location/Qualifiers
source	1..50
	mol_type = other DNA
	organism = synthetic construct

SEQUENCE: 175
cacacaggaa acagctatga ccatgnnnna acatatttgc atggggtgtg 50

SEQ ID NO: 176	moltype = DNA length = 55
FEATURE	Location/Qualifiers
source	1..55
	mol_type = other DNA
	organism = synthetic construct

SEQUENCE: 176
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnccact gcatggttca ctctg 55

SEQ ID NO: 177	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct

SEQUENCE: 177
cacacaggaa acagctatga ccatgnnnna tcctgtgagc gaagttcca 49

SEQ ID NO: 178	moltype = DNA length = 57
FEATURE	Location/Qualifiers
source	1..57
	mol_type = other DNA
	organism = synthetic construct

SEQUENCE: 178
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnntgcct ctttctcttg gttttca 57

SEQ ID NO: 179	moltype = DNA length = 51
FEATURE	Location/Qualifiers
source	1..51
	mol_type = other DNA
	organism = synthetic construct

-continued

SEQUENCE: 179
cacacaggaa acagctatga ccatgnnnng gacctaagca agctgcagta a 51

SEQ ID NO: 180 moltype = DNA length = 60
FEATURE Location/Qualifiers
source 1..60
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 180
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnacacc caatgaagaa tgtaattgat 60

SEQ ID NO: 181 moltype = DNA length = 54
FEATURE Location/Qualifiers
source 1..54
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 181
cacacaggaa acagctatga ccatgnnnng gttgtgtgta gatgtgagtt ttcc 54

SEQ ID NO: 182 moltype = DNA length = 59
FEATURE Location/Qualifiers
source 1..59
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 182
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnttctg ttacattgtg cagagttca 59

SEQ ID NO: 183 moltype = DNA length = 50
FEATURE Location/Qualifiers
source 1..50
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 183
cacacaggaa acagctatga ccatgnnnnt ggttttgagc agagagatgg 50

SEQ ID NO: 184 moltype = DNA length = 55
FEATURE Location/Qualifiers
source 1..55
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 184
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnggaag aagtcccaac catga 55

SEQ ID NO: 185 moltype = DNA length = 53
FEATURE Location/Qualifiers
source 1..53
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 185
cacacaggaa acagctatga ccatgnnnnt cacttttcct ttctacccaa aag 53

SEQ ID NO: 186 moltype = DNA length = 56
FEATURE Location/Qualifiers
source 1..56
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 186
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngaatc tgcattccca gagaca 56

SEQ ID NO: 187 moltype = DNA length = 49
FEATURE Location/Qualifiers
source 1..49
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 187
cacacaggaa acagctatga ccatgnnnnc ctgtttccca tcctcttcc 49

SEQ ID NO: 188 moltype = DNA length = 56
FEATURE Location/Qualifiers
source 1..56
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 188
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnngacac aaaacaggct caggac 56

SEQ ID NO: 189 moltype = DNA length = 49

-continued

FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 189	
cacacaggaa acagctatga ccatgnnnna gaaaccaaag ccaaaagca	49
SEQ ID NO: 190	moltype = DNA length = 55
FEATURE	Location/Qualifiers
source	1..55
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 190	
cgacgtaaaa cgacggccag tnnnnnnnnn nnnnnccatg ggactgactt tctgc	55
SEQ ID NO: 191	moltype = DNA length = 49
FEATURE	Location/Qualifiers
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 191	
cacacaggaa acagctatga ccatgnnnnt catctggacc tgggtcttc	49

1-25. (canceled)

26. A method comprising:

a) testing a liquid Pap specimen obtained from a human subject for at least one of the following nucleic acid mutations:

i) FBXW7 1435C>T,

ii) PIK3CA 323G>A,

iii) TP53 328C>T,

iv) PIK3CA 263G>A, and

v) CTNNB1 100G>A; and

b) detecting the presence of at least one of the following:

i) the T allele at position 1435 of FBXW7 gene;

ii) the A allele at position 323 of PIK3CA gene;

iii) the T allele at position 328 of TP53 gene;

iv) the A allele at position 263 of PIK3CA gene; and

v) the A allele at position 100 of CTNNB1 gene.

27. The method of claim **26**, wherein said human subject has an endometrial tumor.

28. The method of claim **26**, wherein said human subject has endometrial cancer.

29. The method of claim **26**, wherein the liquid Pap specimen is collected from the cervix.

30. The method of claim **26**, wherein the liquid Pap specimen comprises cells from the ectocervix.

31. The method of claim **26**, wherein the presence of the T allele at position 1435 of FBXW7 gene is detected.

32. The method of claim **26**, wherein the presence of the A allele at position 323 of PIK3CA gene is detected.

33. The method of claim **26**, wherein the presence of the T allele at position 328 of TP53 gene is detected.

34. The method of claim **26**, wherein the presence of the A allele at position 263 of PIK3CA gene is detected.

35. The method of claim **26**, wherein the presence of the A allele at position 100 of CTNNB1 gene is detected.

* * * * *