Published in final edited form as:
Sci Transl Med. 2013 January 9; 5(167): 167ra4. doi:10.1126/scitranslmed.3004952.

Evaluation of DNA from the Papanicolaou Test to Detect Ovarian and Endometrial Cancers

Isaac Kinde[1,*], **Chetan Bettegowda**[1,2,*], **Yuxuan Wang**[1,*], **Jian Wu**[1,3], **Nishant Agrawal**[1,4], **Ie-Ming Shih**[5], **Robert Kurman**[5], **Fanny Dao**[6], **Douglas A. Levine**[6], **Robert Giuntoli**[7], **Richard Roden**[5], **James R. Eshleman**[5], **Jesus Paula Carvalho**[8], **Suely Kazue Nagahashi Marie**[8,9], **Nickolas Papadopoulos**[1], **Kenneth W. Kinzler**[1], **Bert Vogelstein**[1,†], and **Luis A. Diaz Jr.**[1,†]

[1]The Ludwig Center for Cancer Genetics and Therapeutics, The Swim Across America Laboratory at Johns Hopkins, and The Howard Hughes Medical Institute at Johns Hopkins Kimmel Cancer Center, Baltimore, MD 21287, USA

[2]Department of Neurosurgery, The Johns Hopkins Medical Institutes, Baltimore, MD 21231, USA

[3]MyGenostics Inc., 801 West Baltimore Street, Suite 502L, Baltimore, MD 21201, USA

[4]Department of Otolaryngology–Head and Neck Surgery, The Johns Hopkins Medical Institutes, Baltimore, MD 21231, USA

[5]Department of Pathology, The Johns Hopkins Medical Institutes, Baltimore, MD 21231, USA

[6]Gynecology Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA

[7]Department of Gynecology and Obstetrics, The Johns Hopkins Medical Institutes, Baltimore, MD 21231, USA

[8]Institute of Cancer of Sao Paulo State, School of Medicine, University of Sao Paulo, Sao Paulo 01246903, Brazil

[†]To whom correspondence should be addressed: bertvog@gmail.com (B.V.); ldiaz1@jhmi.edu (L.A.D.).
[*]These authors contributed equally to this work.

SUPPLEMENTARY MATERIALS
www.sciencetranslationalmedicine.org/cgi/content/full/5/167/167ra4/DC1
Materials and Methods
Table S1. Summary characteristics of endometrial cancers (endometrioid subtype) studied by whole-exome sequencing.
Table S2. Mutations identified by whole-exome sequencing in 22 endometrioid endometrial cancers.
Table S3. Clinical characteristics and mutations assessed in Pap specimens.
Table S4. Primers used to assess individual mutations in Pap specimens.
Table S5. Primers used to simultaneously assess 12 genes in Pap specimens with the multiplexed Safe-SeqS strategy.
Table S6. Mutations identified in Pap specimens through simultaneous assessment of 12 genes with the multiplexed Safe-SeqS strategy.

Author contributions: I.K., C.B., Y.W., R.R., K.W.K., B.V., and L.A.D. designed the study; I.K., C.B., Y.W., J.W., N.A., J.R.E., N.P., K.W.K., B.V., and L.A.D. generated and analyzed the data; I.-M.S., F.D., D.A.L., R.G., J.P.C., S.K.N.M., and L.A.D. contributed cases; and I.K., C.B., Y.W., R.K., D.A.L., B.V., and L.A.D. wrote the manuscript.

Competing interests: N.P., K.W.K., B.V., and L.A.D. are co-founders of Inostics and Personal Genome Diagnostics (PGDx), own stock, and are members of their Scientific Advisory Boards. Inostics and PGDx have licensed several patent applications from Johns Hopkins, on which I.K., N.P., K.W.K., B.V., and L.A.D. are inventors. The patents related to the current manuscript are patent WO2012/142,213 titled “Safe Sequencing System,” and provisional patent application 61/719,942 titled “Papanicolaou test for ovarian and endometrial cancers.” These relationships are subject to certain restrictions under The Johns Hopkins University policy, and the terms of these arrangements are managed by the University in accordance with its conflict-of-interest policies. J.W. was an advisor for MyGenostics Inc. in 2012 and owns company stock.

Data and materials availability: All data and described custom-written scripts are available from the authors upon request.

[9]Departments of Pathology and Neurology, School of Medicine, University of Sao Paulo, Sao Paulo 01246903, Brazil

Abstract

Papanicolaou (Pap) smears have revolutionized the management of patients with cervical cancers by permitting the detection of early, surgically curable tumors and their precursors. In recent years, the traditional Pap smear has been replaced by a liquid-based method, which allows not only cytologic evaluation but also collection of DNA for detection of human papillomavirus, the causative agent of cervical cancer. We reasoned that this routinely collected DNA could be exploited to detect somatic mutations present in rare tumor cells that accumulate in the cervix once shed from endometrial or ovarian cancers. A panel of genes that are commonly mutated in endometrial and ovarian cancers was assembled with new whole-exome sequencing data from 22 endometrial cancers and previously published data on other tumor types. We used this panel to search for mutations in 24 endometrial and 22 ovarian cancers and identified mutations in all 46 samples. With a sensitive massively parallel sequencing method, we were able to identify the same mutations in the DNA from liquid Pap smear specimens in 100% of endometrial cancers (24 of 24) and in 41% of ovarian cancers (9 of 22). Prompted by these findings, we developed a sequence-based method to query mutations in 12 genes in a single liquid Pap smear specimen without previous knowledge of the tumor's genotype. When applied to 14 samples selected from the positive cases described above, the expected tumor-specific mutations were identified. These results demonstrate that DNA from most endometrial and a fraction of ovarian cancers can be detected in a standard liquid-based Pap smear specimen obtained during routine pelvic examination. Although improvements need to be made before applying this test in a routine clinical manner, it represents a promising step toward a broadly applicable screening methodology for the early detection of gynecologic malignancies.

INTRODUCTION

Since the introduction of the Papanicolaou (Pap) test, the incidence and mortality of cervical cancer in screened populations have 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 United States were estimated to be diagnosed with ovarian or endometrial cancer in 2012 (3). Although endometrial cancer is more common than ovarian cancer, the latter is more lethal. In the United States, about 15,000 and 8,000 women are expected to die each year from ovarian and endometrial cancers, respectively (3). Worldwide, more than 200,000 deaths from these tumors are expected this year alone (4, 5).

In an effort to replicate the success of cervical cancer screening, several approaches for the early detection of endometrial and ovarian cancers have been proposed. 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 (6). TVS is used to measure the thickness of the endometrium, because it is known that endometria harboring a cancer are thicker than normal endometria (7). As with cytology, screening measurement of the endometrial thickness with 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 (6, 8).

Even greater efforts have been made to develop a screening test for ovarian cancer, including the assessment of serum CA-125 levels in conjunction with 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 and in some cases before clinical diagnosis (9, 10). 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 (11). Although TVS can visualize the ovary, it can only detect large tumors and cannot definitively distinguish benign from malignant tumors. Several clinical screening trials with serum CA-125 and TVS have been conducted, but none have shown a survival benefit. Some have shown an increase in morbidity compared to controls because false-positive tests elicit further evaluation by laparoscopy or exploratory laparotomy (12–14).

Accordingly, the U.S. Preventative Services Task Force, the American Cancer Society, the American Congress of Obstetricians and Gynecologists, and the National Comprehensive Cancer Network do not recommend routine screening for endometrial or ovarian cancers in the general population. These organizations warn that "the potential harms outweigh the potential benefits" (15–18). 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 ages 30 and 35 years (17, 19).

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 pre-malignant or malignant disease in some cases (20–24), it is often difficult to distinguish the AGCs arising from endocervical, endometrial, or ovarian cancer from one another and from more benign conditions.

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 specimens"; Fig. 1). Unlike cytologically abnormal cells, such oncogenic DNA mutations are specific clonal markers of neoplasia that should be absent in non-neoplastic cells. The experiments described here were carried out to test this hypothesis.

RESULTS

There were four components to this study: (i) establishing the somatic mutations typically present in endometrial and ovarian cancers, (ii) identifying at least one mutation in each tumor from 46 patients with these cancers, (iii) determining whether the mutations identified in these tumors could also be detected in Pap specimens from the same patients, and (iv) developing a technology that could directly assess cells from Pap specimens for mutations commonly found in endometrial or ovarian cancers.

Prevalence of somatically mutated genes in endometrial and ovarian cancers

There are six major histopathologic subtypes of epithelial ovarian cancers (Table 1). The most prevalent subtype is high-grade serous (60% of total), followed by endometrioid (15%), clear cell (10%), low-grade serous carcinoma (8%), mucinous (2%), and transitional cell carcinoma (2%) (25–27). Most of these cases are found at an advanced stage, and the combined 5-year survival (3, 28) for these malignancies is about 27% (Table 1). Genome-

wide studies (29–31) have identified commonly mutated genes among the most prevalent ovarian cancer subtypes (Table 2).

Such comprehensive studies have not yet been reported for the endometrioid and mucinous subtypes, which collectively represent ~20% of ovarian cancer cases (Table 1). However, commonly mutated genes in the endometrioid and mucinous subtypes have been reported (32). 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 *PPP2R1A* (Table 2).

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 nonneoplastic 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 S1. Although 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.

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 10 reads. After applying stringent criteria for the identification of somatic mutations (as described in Materials and Methods), the sequencing data demarcated the tumors into two groups: 10 cancers (termed group N, for non–highly mutated) harbored <100 somatic mutations per tumor (median, 32; range, 7 to 50), and 12 cancers (termed group H, for highly mutated) harbored >100 somatic mutations per tumor (median, 674; range, 164 to 4629) (table S1).

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*, whereas none of the group N cancers contained mutations in mismatch repair genes (table S2). Mismatch repair deficiency is common among endometrial cancers, and these tumors occur in 19 to 71% of women with inherited mutations in mismatch repair genes (that is, patients with hereditary nonpolyposis colorectal cancer) (33).

A complete list of the 12,795 somatic mutations identified in the 22 cancers is provided in table S2. The most commonly mutated genes included the phosphatidylinositol 3-kinase pathway genes *PTEN* and *PIK3CA* (34), the adenomatous polyposis coli 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 previous studies of endometrioid endometrial cancer that had evaluated small numbers of genes, although 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 and table S2), a finding also consistent with previous studies.

Papillary serous carcinomas of the endometrium account for 10 to 15% of endometrial cancers, and a recent genome-wide sequencing study of this tumor subtype has been published (35). The most common mutations in this subtype are listed in Table 2. The least common subtype of endometrial cancers is clear cell carcinoma (36), which occurs in <5%

of cases. Genes reported to be mutated in these cancers were garnered from the literature (Table 2).

Identification of mutations in tumor tissues

We acquired tumors from 46 cancer patients for whom Pap specimens were available. These included 24 patients with endometrial cancers and 22 with ovarian cancers; their clinical, demographic, and histopathologic features are listed in table S3.

Somatic mutations in the 46 tumors were identified through whole-exome sequencing (table S2) 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 with a custom solid-phase capture assay composed 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.

DNA sequencing libraries were generated from tumors and their matched nonneoplastic tissues, and then captured with the assay described above. 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.

Identification of somatic mutations in Pap specimens

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 (for example, ThinPrep). Some of the liquid from the vial is used to prepare a slide for cytological analysis or for purification of human papillomavirus (HPV) DNA. In our study, an aliquot of the DNA purified from the liquid was assessed for the presence of DNA from the cancers of the 46 patients. Preliminary studies showed that the fixed cells or cell fragments in the liquid 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 specimens was 9.9 ± 14.8 μg (table S3).

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 (37), in which DNA templates are amplified with modified gene-specific primers, was designed for this purpose (Fig. 2).

In brief, a limited number of polymerase chain reaction (PCR) cycles were performed with a set of gene-specific primers. One of the primers contained 14 degenerate "N" bases (that is, 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" bases formed unique identifiers (UIDs) for each original template molecule. Subsequent PCR products generated with universal primers were purified and sequenced. 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" (37). 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 super-mutants. The Safe-SeqS approach used here is capable of detecting 1 mutant template among 5000 to 1,000,000 wild-type templates, depending on the amplicon and the position within the amplicon that is queried (37).

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 specimens from patients with endometrial cancers, a mutation present in the tumor was identified in every case (100%). The median fraction of mutant alleles was 3% and ranged from 0.01 to 80% (Fig. 3 and table S3). Amplifications of DNA from nonneoplastic 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 in 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).

In endometrial cancer cases PAP 041 and PAP 083, two mutations found in the tumor DNA were evaluated in the Pap specimens (table S3). In both cases, the mutations were identified in DNA from the Pap specimen (table S3). Moreover, the ratios between the mutant allele fractions of the two mutations in the Pap specimens were correlated with those of the corresponding tumor samples. For example, in the Pap specimen of case PAP 083, the mutant allele fractions for the *CTNNB1* and *PIK3CA* mutations were 0.14 and 0.064%, respectively—a ratio of 2.2. In the primary tumor from PAP 083, the corresponding ratio was 2.0 (79.5 to 39.5%).

Similar analysis of Pap specimen DNA from ovarian cancer patients revealed detectable mutations in 9 of the 22 patients (41%). The fraction of mutant alleles was smaller than in endometrial cancers (median, 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 the 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 assessed in our cohort.

A genetic test for screening purposes

Our results demonstrate that mutant DNA molecules from most endometrial cancers and some ovarian cancers can be found in routinely collected Pap specimens. 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 specimens. In a screening setting, the presence and genotype of tumors would obviously not be known before evaluation. We therefore designed a prototype test based on Safe-SeqS that could assess several genes and could be used in a screening setting (Fig. 2).

This multiplexed approach included 50 primer pairs that amplified segments of 241 to 296 base pairs (bp) containing frequently mutated regions of DNA. The regions to be amplified were chosen from the most commonly mutated genes in endometrial and ovarian cancers (Table 2) and included exons from *APC*, *AKT1*, *BRAF*, *CTNNB1*, *EGFR*, *FBXW7*, *KRAS*, *NRAS*, *PIK3CA*, *PPP2R1A*, *PTEN*, and *TP53.* In control experiments, 46 of the 50 amplicons were shown to provide information on a minimum of 2500 templates, as the number of sequenced templates can be determined directly from Safe-SeqS data (Fig. 2). Given the accuracy of Safe-SeqS, this number was adequate to comfortably detect mutations existing in >0.1% of template molecules (37). 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.

This test was applied to Pap specimens from 14 patients with cancer—12 endometrial and 2 ovarian—as well as 14 Pap specimens collected from healthy controls. The two ovarian cancers used were stages IA and IV. The endometrial cancers were stage I ($n = 10$), stage II ($n = 1$), and stage IV ($n = 1$). 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 specimens 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, where only one Safe-SeqS primer pair per amplicon was used (tables S3 and S6). No mutations were detected in the 14 Pap specimens from women without cancer (Fig. 4).

DISCUSSION

Georgios Papanicolaou published his seminal work, entitled *Diagnosis of Uterine Cancer by the Vaginal Smear*, in 1943 (38). 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.

An important development from the last several years is the recognition that all human cancers result from mutations in a limited set of genes and an even more limited set of pathways through which these genes act (39). The whole-exome sequencing data we present, combined with previous genome-wide studies, provide a compelling example of the common genetic features of cancer (Table 2). Through the analysis of particular regions of only 12 genes (table S5), we could detect at least one driver mutation in most of nine different gynecologic cancers (Table 1). Although several of these 12 genes were tumor suppressors, and therefore difficult to therapeutically target, knowledge of their mutational patterns provides actionable opportunities for cancer diagnostics.

The most important finding in this paper is that diagnostically useful 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. In our study, 100% of endometrial cancers ($n = 24$) and 41% of ovarian cancers ($n = 22$) shed cells into the cervix that could be detected in materials collected as part of routine Pap specimens. In both types of cancer, tumors of low grade were detected (Fig. 3). These findings, in conjunction with technical advances allowing the reliable detection of mutations present in only a very small fraction of DNA templates, are the foundations of the PapGene test.

This study provides proof of principle for endocervical DNA testing for gynecologic cancers, but there are important limitations that must be addressed before this approach can be used in the clinic. The test, even in its current format, appears to be promising as a screening tool for endometrial cancer because the data in Fig. 3 show that even the lowest stage endometrial cancers could be detected through the analysis of DNA in Pap specimens. However, only 41% of ovarian cancers could be detected in Pap specimens, even when the mutations in their tumors were known. In eight of the nine Pap specimens 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 (table S3). Further improvements in the technology could increase the technical sensitivity of the PapGene test and allow it to detect more ovarian cancers. One improvement would involve an increase in the number of potential gene targets assessed by the PapGene test. Development of an improved collection method may also be important to improve sensitivity. The current method of liquid specimen collection is designed for the detection of cervical cancer and, as such, uses a brush that collects cells from the ectocervix and only minimally penetrates the endocervical canal. A small cannula introduced into the endometrial cavity, similar to the Pipelle endometrial biopsy instrument, could theoretically be used to obtain a more highly enriched sample of cells coming from the endometrium, fallopian tube, and ovary (40). Specificity must also be further addressed in the future, and a greater number of healthy controls need to be evaluated, although it is encouraging that none evaluated so far had detectable mutations. This result is consistent with the idea that mutation-based screening should be exquisitely specific because mutations should not be found in normal cells. As noted in the Introduction, specificity is a major limitation of current screening tests in general and for ovarian cancer in particular.

The quantitative nature of the PapGene test also opens the possibility of using it to monitor the response to hormonal agents (for example, progestins) when treating young women with low-risk endometrial cancers. Some of these women choose to preserve fertility, undergoing medical therapy rather than hysterectomy (41), and PapGene testing could be performed at regular intervals to monitor them for local cancer recurrence or progression.

Even if tumors were identified at an advanced stage, detection of presymptomatic ovarian cancers could be of benefit. 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 now, surgeons attempt to remove any visible tumor. With each improvement in surgical debulking, survival has lengthened (42). The earlier these advanced-stage ovarian cancers are diagnosed, the lower the overall tumor burden and the higher the chance of optimal debulking. Furthermore, it is possible that 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.

An essential aspect of any screening approach is that it should be relatively inexpensive and easily incorporated into standard medical practice. Evaluation of HPV DNA is already part of routine Pap specimen testing because HPV analysis increases the test's sensitivity (43, 44). The DNA purification component of the PapGene test is identical to that used for HPV, so this component is feasible. The preparation of DNA, multiplex amplification, and sequencing constituting the PapGene test can be performed at a cost comparable to a routine HPV test in the United States today. Note that the increased sensitivity provided by the Safe-SeqS component of the PapGene test can be implemented on any massively parallel sequencing instrument, not just those used in this study. With the reduction in the cost of massively parallel sequencing expected in the future, PapGene testing should become even less expensive.

There are millions of Pap smear tests performed annually in the United States. 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. Although 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 in 1943 (38). Even today, when many cervical cytology specimens are screened with automated technologies, at least 2 to 8% of samples require evaluation by a skilled cytopathologist (45). In contrast, the analysis of PapGene testing is done completely in silico, and the readout of the test is objective and quantitative.

In sum, these data highlight the high specificity of mutation-based diagnostics paired with the sensitivity of interrogating local-regional bodily secretions for tumor-derived DNA. PapGene testing has the capacity to increase the use of conventional cytology screening through the unambiguous detection of DNA from endometrial and ovarian carcinomas, and lays the foundation for a new generation of screening tests.

MATERIALS AND METHODS

Patient samples

All samples for this study were obtained according to protocols approved by the Institutional Review Boards of The Johns Hopkins Medical Institutions (Baltimore, MD), Memorial Sloan-Kettering Cancer Center (New York, NY), University of Sao Paulo (Sao Paulo, Brazil), and ILSBio, LLC (Chestertown, MD). Demographic, clinical, and pathologic staging data were collected for each case. All histopathology was centrally re-reviewed by board-certified pathologists. Staging was based on 2009 FIGO criteria (46). Purified DNA from tumor and normal tissues as well as liquid-based Pap smears were quantified in all cases with quantitative PCR using the primers and conditions previously described (47). Unless otherwise indicated, all patient-related values are reported as means ± 1 SD. Additional details are provided in the Supplementary Materials and Methods.

MSI testing

Tumor samples were designated as follows: MSI-high if two or more mononucleotide repeats varied in length compared to the germline DNA; MSI-low if only one locus varied; and microsatellite stable if there was no variation compared to the germ line. Pentanucleotide loci confirmed identity in all cases. Additional details are provided in the Supplementary Materials and Methods.

Preparation and sequencing of captured Illumina DNA libraries

Preparation of Illumina genomic DNA libraries and selection for exomic DNA were performed according to the manufacturer's recommendations. Exomic capture was performed with the SureSelect Human Exome Kit V 4.0 (Agilent), whereas the custom solid-phase capture assay was performed by modification of previously described methods (48, 49). Paired-end sequencing with an Illumina GA IIx Genome Analyzer provided 2 × 75 base reads from each fragment. Known polymorphisms recorded in dbSNP Build 130 (50) in the sequence tags that passed filtering were removed from the analysis. Identification of high-confidence mutations was performed as described previously (30). Additional details are provided in the Supplementary Materials and Methods.

Assessment of low-frequency mutations

Primers were designed as described previously (37) with Primer3 (51). Templates were prepared for sequencing as described previously (37), with modifications that facilitated the amplification of multiple gene regions in a single well of a 96-well PCR plate. With the primers described in table S4, 66 ng of templates was amplified in two rounds of PCR (Fig. 2) for the single amplicon assays. The multiplexed assays were performed in similar fashion with six independent amplifications—each containing 66 ng of DNA (that is, ~400 ng total)—per sample with the primers described in table S5. High-quality sequence reads were analyzed as previously described (37), employing quality scores that reflected the

probability that an individual base call was made in error (52). The template-specific portion of the reads was matched to a reference sequence set by a custom script (available from the authors upon request). Additional details are provided in the Supplementary Materials and Methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank C. Blair, K. Judge, J. Schaeffer, J. Ptak, N. Silliman, L. Dobbyn, and M. Thiess for technical assistance and M. Ekdahl for operational assistance.

Funding: This work was supported by Swim Across America, The Commonwealth Fund, the Hilton-Ludwig Cancer Prevention Initiative, the Virginia and D. K. Ludwig Fund for Cancer Research, the Experimental Therapeutics Center of Memorial Sloan-Kettering Cancer Center, Chia Family Foundation, The Honorable Tina Brozman Foundation, The United Negro College Fund–Merck Graduate Science Dissertation Fellowship, the Burroughs Wellcome Career Award for Medical Scientists and the National Colorectal Cancer Research Alliance, National Cancer Institute contract N01-CN-43309, and NIH grants CA129825, CA062924, and CA43460.

REFERENCES AND NOTES

1. Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N, Wiener H, Herbert A, von Karsa L. European Guidelines for Quality Assurance in Cervical Cancer Screening. Second edition—Summary document. Ann Oncol. 2010; 21:448–458. [PubMed: 20176693]
2. DeMay, RM. Practical Principles of Cytopathology. ASCP Press; Chicago: 2010.
3. Howlader, N.; Noone, AM.; Krapcho, M.; Neyman, N.; Aminou, R.; Altekruse, SF.; Kosary, CL.; Ruhl, J.; Tatalovich, Z.; Cho, H.; Mariotto, A.; Eisner, MP.; Lewis, DR.; Chen, HS.; Feuer, EJ.; Cronin, KA. SEER Cancer Statistics Review, 1975–2009. Vintage 2009 Populations. http://seer.cancer.gov/csr/1975_2009_pops09/
4. Bray F, Ren JS, Masuyer E, Ferlay J. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. Int J Cancer. 201210.1002/ijc.27711
5. International Agency for Research on Cancer. GLOBOCAN 2008 v1.2. Cancer Incidence and Mortality Worldwide: IARC CancerBase No 10. http://globocan.iarc.fr
6. Sams SB, Currens HS, Raab SS. Liquid-based Papanicolaou tests in endometrial carcinoma diagnosis. Performance, error root cause analysis, and quality improvement. Am J Clin Pathol. 2012; 137:248–254. [PubMed: 22261450]
7. Smith P, Bakos O, Heimer G, Ulmsten U. Transvaginal ultrasound for identifying endometrial abnormality. Acta Obstet Gynecol Scand. 1991; 70:591–594. [PubMed: 1785276]
8. Mitchell H, Giles G, Medley G. Accuracy and survival benefit of cytological prediction of endometrial carcinoma on routine cervical smears. Int J Gynecol Pathol. 1993; 12:34–40. [PubMed: 8418076]
9. Carlson KJ, Skates SJ, Singer DE. Screening for ovarian cancer. Ann Intern Med. 1994; 121:124–132. [PubMed: 8017726]
10. Cramer DW, Bast RC Jr, Berg CD, Diamandis EP, Godwin AK, Hartge P, Lokshin AE, Lu KH, McIntosh MW, Mor G, Patriotis C, Pinsky PF, Thornquist MD, Scholler N, Skates SJ, Sluss PM, Srivastava S, Ward DC, Zhang Z, Zhu CS, Urban N. Ovarian cancer biomarker performance in prostate, lung, colorectal, and ovarian cancer screening trial specimens. Cancer Prev Res. 2011; 4:365–374.
11. Meden H, Fattahi-Meibodi A. CA 125 in benign gynecological conditions. Int J Biol Markers. 1998; 13:231–237. [PubMed: 10228907]
12. Buys SS, Partridge E, Black A, Johnson CC, Lamerato L, Isaacs C, Reding DJ, Greenlee RT, Yokochi LA, Kessel B, Crawford ED, Church TR, Andriole GL, Weissfeld JL, Fouad MN, Chia D, O'Brien B, Ragard LR, Clapp JD, Rathmell JM, Riley TL, Hartge P, Pinsky PF, Zhu CS, Izmirlian G, Kramer BS, Miller AB, Xu JL, Prorok PC, Gohagan JK, Berg CD. PLCO Project

Team, Effect of screening on ovarian cancer mortality: The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. JAMA. 2011; 305:2295–2303. [PubMed: 21642681]

13. ACOG Committee on Practice Bulletins, ACOG Practice Bulletin. Clinical Management Guidelines for Obstetrician-Gynecologists. Number 60, March 2005. Pregestational diabetes mellitus. Obstet Gynecol. 2005; 105:675–685. [PubMed: 15738045]
14. Partridge E, Kreimer AR, Greenlee RT, Williams C, Xu JL, Church TR, Kessel B, Johnson CC, Weissfeld JL, Isaacs C, Andriole GL, Ogden S, Ragard LR, Buys SS. PLCO Project Team, Results from four rounds of ovarian cancer screening in a randomized trial. Obstet Gynecol. 2009; 113:775–782. [PubMed: 19305319]
15. American Cancer Society. Detailed guide: Ovarian cancer—Can ovarian cancer be found early?. http://www.cancer.org/cancer/ovariancancer/detailedguide/ovarian-cancer-detection
16. U.S. Preventive Services Task Force. Screening for ovarian cancer: Recommendation statement. Am Fam Physician. 2005; 71:759–762. [PubMed: 15756773]
17. American College of Obstetricians and Gynecologists. ACOG Committee Opinion: Number 280, December 2002. The role of the generalist obstetrician-gynecologist in the early detection of ovarian cancer. Obstet Gynecol. 2002; 100:1413–1416. [PubMed: 12468197]
18. National Comprehensive Cancer Network. Practice Guidelines in Oncology: Ovarian cancer and genetic screening. http://www.nccn.org/professionals/physician_gls/PDF/genetics_screening.pdf
19. Lindor NM, Petersen GM, Hadley DW, Kinney AY, Miesfeldt S, Lu KH, Lynch P, Burke W, Press N. Recommendations for the care of individuals with an inherited predisposition to Lynch syndrome: A systematic review. JAMA. 2006; 296:1507–1517. [PubMed: 17003399]
20. Marques JP, Costa LB, Pinto AP, Lima AF, Duarte ME, Barbosa AP, Medeiros PL. Atypical glandular cells and cervical cancer: Systematic review. Rev Assoc Med Bras. 2011; 57:234–238. [PubMed: 21537713]
21. Insinga RP, Glass AG, Rush BB. Diagnoses and outcomes in cervical cancer screening: A population-based study. Am J Obstet Gynecol. 2004; 191:105–113. [PubMed: 15295350]
22. Sharpless KE, Schnatz PF, Mandavilli S, Greene JF, Sorosky JI. Dysplasia associated with atypical glandular cells on cervical cytology. Obstet Gynecol. 2005; 105:494–500. [PubMed: 15738014]
23. DeSimone CP, Day ME, Tovar MM, Dietrich CS III, Eastham ML, Modesitt SC. Rate of pathology from atypical glandular cell Pap tests classified by the Bethesda 2001 nomenclature. Obstet Gynecol. 2006; 107:1285–1291. [PubMed: 16738153]
24. Geier CS, Wilson M, Creasman W. Clinical evaluation of atypical glandular cells of undetermined significance. Am J Obstet Gynecol. 2001; 184:64–69. [PubMed: 11174481]
25. Malpica A, Deavers MT, Lu K, Bodurka DC, Atkinson EN, Gershenson DM, Silva EG. Grading ovarian serous carcinoma using a two-tier system. Am J Surg Pathol. 2004; 28:496–504. [PubMed: 15087669]
26. Ries, LAG.; Young, JL.; Keel, GE.; Eisner, MP.; Lin, YD.; Horner, M-J. SEER Survival Monograph: Cancer Survival Among Adults: US SEER Program, 1988–2001. Patient and Tumor Characteristics. http://seer.cancer.gov/publications/survival/
27. Ichigo S, Takagi H, Matsunami K, Murase T, Ikeda T, Imai A. Transitional cell carcinoma of the ovary (Review). Oncol Lett. 2012; 3:3–6. [PubMed: 22740845]
28. Kommoss F, Kommoss S, Schmidt D, Trunk MJ, Pfisterer J, du Bois A. Arbeitsgemeinschaft Gynaekologische Onkologie Studiengruppe Ovarialkarzinom, Survival benefit for patients with advanced-stage transitional cell carcinomas vs. other subtypes of ovarian carcinoma after chemotherapy with platinum and paclitaxel. Gynecol Oncol. 2005; 97:195–199. [PubMed: 15790458]
29. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature. 2011; 474:609–615. [PubMed: 21720365]
30. Jones S, Wang TL, Shih IM, Mao TL, Nakayama K, Roden R, Glas R, Slamon D, Diaz LA Jr, Vogelstein B, Kinzler KW, Velculescu VE, Papadopoulos N. Frequent mutations of chromatin remodeling gene *ARID1A* in ovarian clear cell carcinoma. Science. 2010; 330:228–231. [PubMed: 20826764]

31. Jones S, Wang TL, Kurman RJ, Nakayama K, Velculescu VE, Vogelstein B, Kinzler KW, Papadopoulos N, Shih IM. Low-grade serous carcinomas of the ovary contain very few point mutations. J Pathol. 2012; 226:413–420. [PubMed: 22102435]
32. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, Jia M, Shepherd R, Leung K, Menzies A, Teague JW, Campbell PJ, Stratton MR, Futreal PA. COSMIC: Mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res. 2011; 39:D945–D950. [PubMed: 20952405]
33. Barrow E, Robinson L, Alduaij W, Shenton A, Clancy T, Lalloo F, Hill J, Evans DG. Cumulative lifetime incidence of extracolonic cancers in Lynch syndrome: A report of 121 families with proven mutations. Clin Genet. 2009; 75:141–149. [PubMed: 19215248]
34. Oda K, Stokoe D, Taketani Y, McCormick F. High frequency of coexistent mutations of *PIK3CA* and *PTEN* genes in endometrial carcinoma. Cancer Res. 2005; 65:10669–10673. [PubMed: 16322209]
35. Kuhn E, Wu RC, Guan B, Wu G, Zhang J, Wang Y, Song L, Yuan X, Wei L, Roden RB, Kuo KT, Nakayama K, Clarke B, Shaw P, Olvera N, Kurman RJ, Levine DA, Wang TL, Shih IM. Identification of molecular pathway aberrations in uterine serous carcinoma by genome-wide analyses. J Natl Cancer Inst. 2012; 104:1503–1513. [PubMed: 22923510]
36. Hamilton CA, Cheung MK, Osann K, Chen L, Teng NN, Longacre TA, Powell MA, Hendrickson MR, Kapp DS, Chan JK. Uterine papillary serous and clear cell carcinomas predict for poorer survival compared to grade 3 endometrioid corpus cancers. Br J Cancer. 2006; 94:642–646. [PubMed: 16495918]
37. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. Proc Natl Acad Sci USA. 2011; 108:9530–9535. [PubMed: 21586637]
38. Traut HF, Papanicolaou GN. Cancer of the uterus: The vaginal smear in its diagnosis. Cal West Med. 1943; 59:121–122. [PubMed: 18746585]
39. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med. 2004; 10:789–799. [PubMed: 15286780]
40. Cooper JM, Erickson ML. Endometrial sampling techniques in the diagnosis of abnormal uterine bleeding. Obstet Gynecol Clin North Am. 2000; 27:235–244. [PubMed: 10857117]
41. Gunderson CC, Fader AN, Carson KA, Bristow RE. Oncologic and reproductive outcomes with progestin therapy in women with endometrial hyperplasia and grade 1 adenocarcinoma: A systematic review. Gynecol Oncol. 2012; 125:477–482. [PubMed: 22245711]
42. Bristow RE, Tomacruz RS, Armstrong DK, Trimble EL, Montz FJ. Survival effect of maximal cytoreductive surgery for advanced ovarian carcinoma during the platinum era: A meta-analysis. J Clin Oncol. 2002; 20:1248–1259. [PubMed: 11870167]
43. Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A, Ratnam S, Coutlée F, Franco EL. Canadian Cervical Cancer Screening Trial Study Group, Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. N Engl J Med. 2007; 357:1579–1588. [PubMed: 17942871]
44. Naucler P, Ryd W, Törnberg S, Strand A, Wadell G, Elfgren K, Rådberg T, Strander B, Johansson B, Forslund O, Hansson BG, Rylander E, Dillner J. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. N Engl J Med. 2007; 357:1589–1597. [PubMed: 17942872]
45. Willis BH, Barton P, Pearmain P, Bryan S, Hyde C. Cervical screening programmes: Can automation help? Evidence from systematic reviews, an economic analysis and a simulation modelling exercise applied to the UK. Health Technol Assess. 2005; 9:1–207. [PubMed: 15774236]
46. Pecorelli S. Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. Int J Gynaecol Obstet. 2009; 105:103–104. [PubMed: 19367689]
47. Rago C, Huso DL, Diehl F, Karim B, Liu G, Papadopoulos N, Samuels Y, Velculescu VE, Vogelstein B, Kinzler KW, Diaz LA Jr. Serial assessment of human tumor burdens in mice by the analysis of circulating DNA. Cancer Res. 2007; 67:9364–9370. [PubMed: 17909045]
48. Wu J, Matthaei H, Maitra A, Dal Molin M, Wood LD, Eshleman JR, Goggins M, Canto MI, Schulick RD, Edil BH, Wolfgang CL, Klein AP, Diaz LA Jr, Allen PJ, Schmidt CM, Kinzler KW,

Papadopoulos N, Hruban RH, Vogelstein B. Recurrent *GNAS* mutations define an unexpected pathway for pancreatic cyst development. Sci Transl Med. 2011; 3:92ra66.

49. He J, Wu J, Jiao Y, Wagner-Johnston N, Ambinder RF, Diaz LA Jr, Kinzler KW, Vogelstein B, Papadopoulos N. IgH gene rearrangements as plasma biomarkers in non-Hodgkin's lymphoma patients. Oncotarget. 2011; 2:178–185. [PubMed: 21399237]
50. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: The NCBI database of genetic variation. Nucleic Acids Res. 2001; 29:308–311. [PubMed: 11125122]
51. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 2000; 132:365–386. [PubMed: 10547847]
52. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 1998; 8:186–194. [PubMed: 9521922]

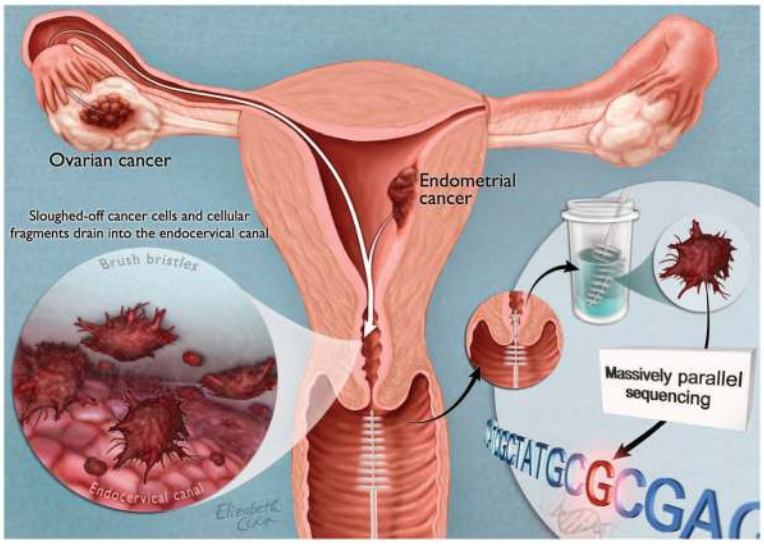

Fig. 1.

Schematic demonstrating the principal steps of the procedure described in this study. Tumor 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. By means of massively parallel sequencing, this DNA is queried for mutations that indicate the presence of a malignancy in the female reproductive tract.

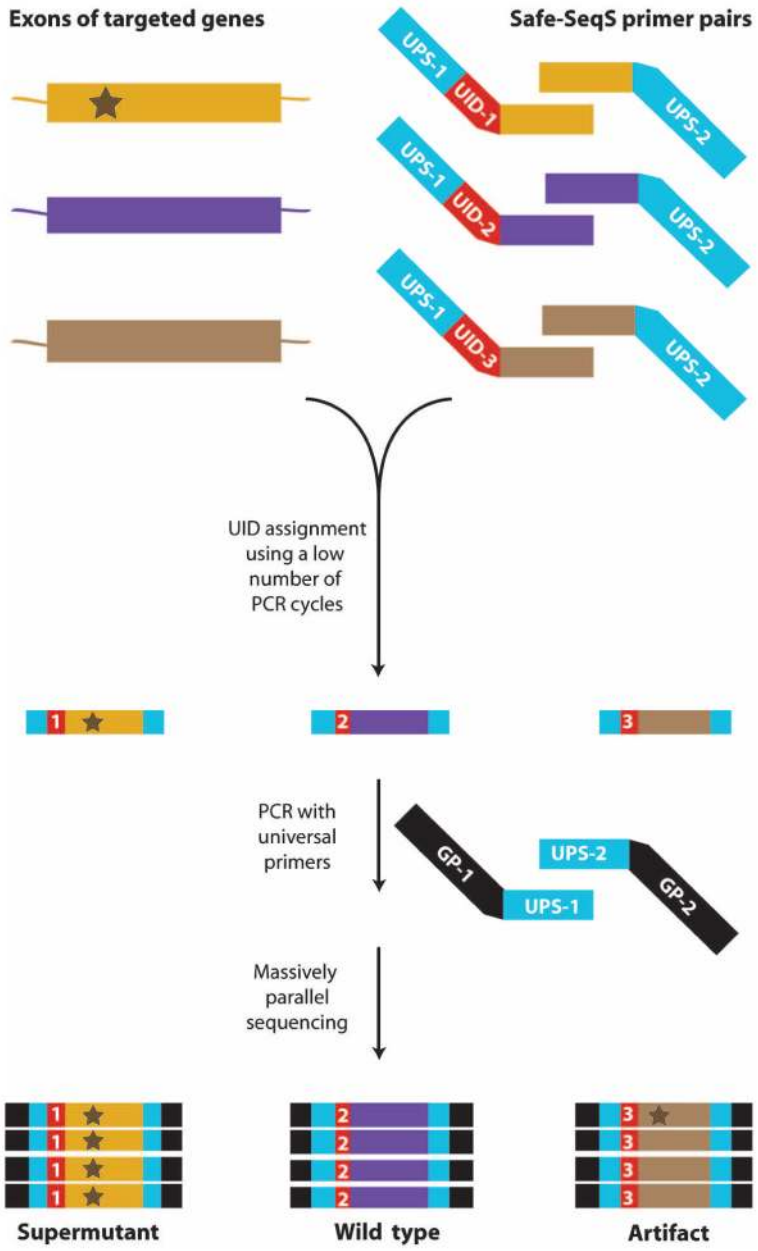

Fig. 2.

Diagram of the modified Safe-SeqS assay used, which allowed for the simultaneous detection of mutations in 12 different genes. Top left: DNA templates from three exons of different genes (yellow, purple, and brown rectangles) to be queried for mutations. Note that only one of the templates contains a mutation (star) that exists before any sample preparatory steps or sequencing. Top right: Safe-SeqS primer pairs contain binding sites for universal primers ("UPS," blue), a unique identifier ("UID," red), and gene-specific sequences (colors match the targeted exon). Next, the templates and primers are combined into a single PCR compartment and a UID is attached to each targeted template, along with UPS binding sites, after a low number of PCR cycles ("UID assignment"). The Safe-SeqS primers are removed, and subsequent PCR is performed with primers containing UPS sites, as well as the sequences required for attachment to the sequencing instrument ("GP," black)

to prepare the templates for massively parallel sequencing. When mutations preexist in template DNA before sample preparation, all of the sequenced daughter molecules sharing the same UID will contain the same mutation (a “supermutant”). In contrast, artifactual mutations caused by sample preparation or sequencing are unlikely to be observed in most other daughter molecules sharing the same UID (“Artifact”). Note that only one of two DNA strands is depicted for clarity.

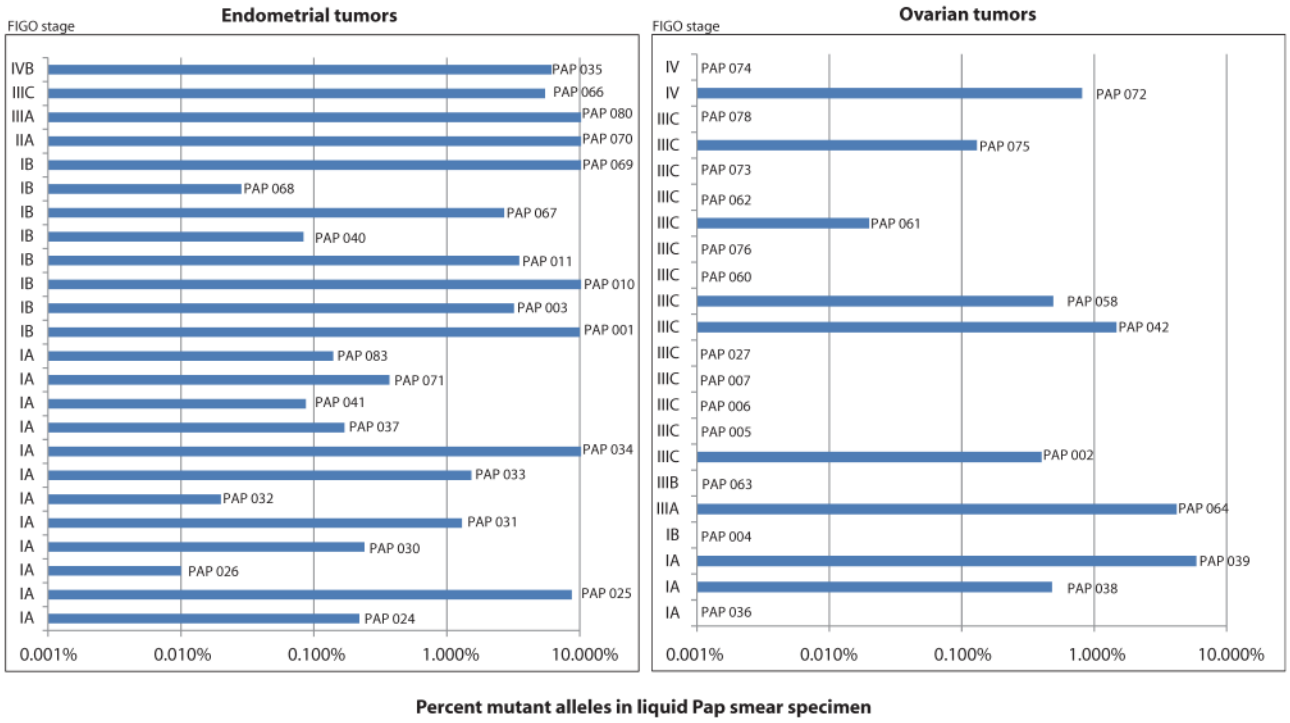

Fig. 3.

Percent mutant alleles in liquid Pap smear specimens. The fraction of mutant alleles from each of 46 Pap specimens is depicted. The stage of each tumor is listed on the *y* axis. The *x* axis demonstrates the percent mutant allele fraction as determined by traditional Safe-SeqS. Mutant allele frequencies are higher than 10% in some cases but are depicted at 10% in this figure for clarity. Precise mutation frequencies are reported in table S3 for all samples.

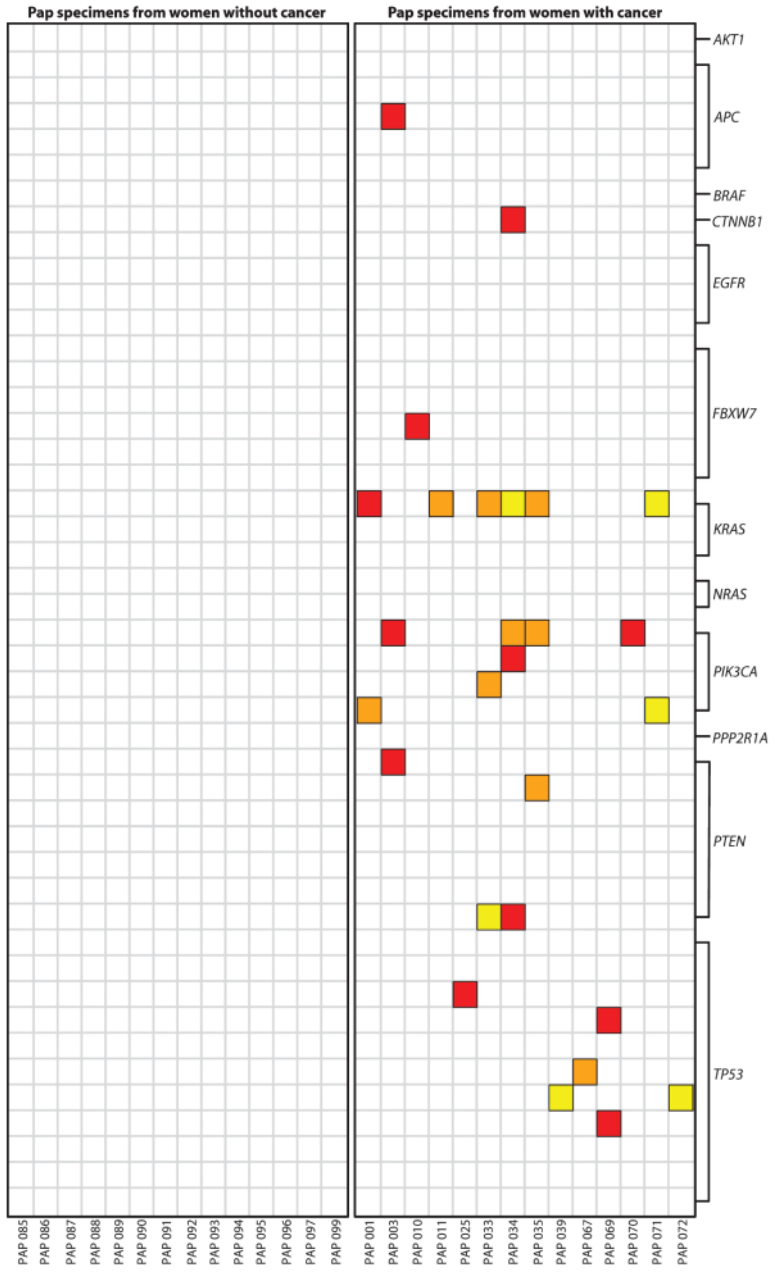

Fig. 4.

Heat map depicting the results of multiplex testing of 12 genes in Pap specimens. The PapGene test interrogates 46 gene regions, with each block on the *y* axis representing one region analyzed for the indicated gene. The 28 samples assessed (14 from control women without cancer and 14 from women with 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 indicating a mutant fraction of 1 to 10%, and red indicating a mutant fraction of >10%.

Table 1

Epidemiology of ovarian and endometrial tumors in the United States. The estimated numbers of new cases and deaths in the United States from the major subtypes of ovarian and endometrial cancers are listed. N/A, not available.

Tissue	Type	Subtype	Subtype distribution (%)	Estimated new cases, 2012	5-Year survival (%)	References
Ovarian	Epithelial	High-grade serous	60	13,368	9	(3, 25)
		Endometrioid	15	3,342	71	(3, 26)
		Clear cell	10	2,228	62	(3, 26)
		Low-grade serous	8	1,782	40	(3, 25)
		Mucinous	2	446	65	(3, 26)
		Transitional cell	2	446	57	(27, 28)
		Other	3	668	N/A	(3, 25–28)
Endometrial	Type I: endometrioid	Endometrioid	85	40,060	91	(3, 26)
	Type II: non-endometrioid	Papillary serous	10	4,713	45	(3, 26)
		Clear cell	5	2,357	68	(3, 36)

Sci Transl Med. Author manuscript; available in PMC 2013 August 30.

Table 2

Genetic characteristics of ovarian and endometrial cancers. The frequencies of the commonly mutated genes in ovarian and endometrial cancers are listed.

Tissue	Type	Subtype	Somatically mutated genes (frequency)	Reference
Ovarian	Epithelial	High-grade serous	*TP53* (96%)	(29)
			CSMD3 (6%)	(29)
			FAT3 (6%)	(29)
			BRCA1 (3%)	(29)
			BRCA2 (3%)	(29)
		Endometrioid	*TP53* (68%)	(32)
			ARID1A (30%)	(32)
			CTNNB1 (26%)	(32)
			PTEN (17%)	(32)
			PIK3CA (15%)	(32)
			KRAS (10%)	(32)
			PPP2R1A (11%)	(32)
			CDKN2A (12%)	(32)
			BRAF (8%)	(32)
		Clear cell	*ARID1A* (57%)	(30)
			PIK3CA (40%)	(30)
			PPP2R1A (7%)	(30)
			KRAS (4.7%)	(30)
		Low-grade serous	*BRAF* (38%)	(31)
			KRAS (19%)	(31)
		Mucinous	*TP53* (56%)	(32)
			KRAS (40%)	(32)
			PPP2R1A (33%)	(32)
			CDKN2A (16%)	(32)
			PTEN (11%)	(32)
Endometrial	Type I: endometrioid	Endometrioid	*PTEN* (64%)	Current study
			PIK3CA (59%)	Current study
			ARID1A (55%)	Current study
			CTNNB1 (32%)	Current study

Tissue	Type	Subtype	Somatically mutated genes (frequency)	Reference
			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* (82%)	(35)
			PIK3CA (24%)	(35)
			FBXW7 (20%)	(35)
			PPP2R1A (18%)	(35)
		Clear cell	*TP53* (45%)	(32)
			PPP2R1A (33%)	(32)
			PIK3CA (29%)	(32)
			PTEN (13%)	(32)
			PIK3R1 (9%)	(32)
			KRAS (5%)	(32)