

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
4 May 2006 (04.05.2006)

PCT

(10) International Publication Number
WO 2006/047484 A2(51) International Patent Classification: **Not classified**(21) International Application Number:
PCT/US2005/038315

(22) International Filing Date: 24 October 2005 (24.10.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/620,926	22 October 2004 (22.10.2004)	US
60/631,240	29 November 2004 (29.11.2004)	US
60/644,568	19 January 2005 (19.01.2005)	US
60/679,969	12 May 2005 (12.05.2005)	US
60/679,968	12 May 2005 (12.05.2005)	US

(71) Applicant (for all designated States except US): **RED-PATH INTEGRATED PATHOLOGY, INC.** [US/US];
816 Middle Street, 2nd Floor, Pittsburgh, Pennsylvania
15212 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **FINKELSTEIN,**

Sydney David [US/US]; 311 Marberry Drive, Pittsburgh,
Pennsylvania 15215 (US). **SWALSKY, Patricia** [US/US];
311 Marberry Drive, Pittsburgh, Pennsylvania 15215 (US).

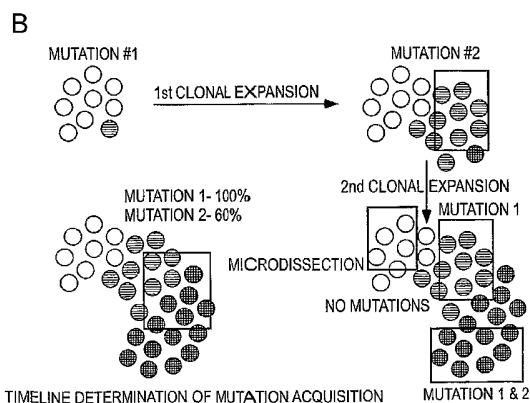
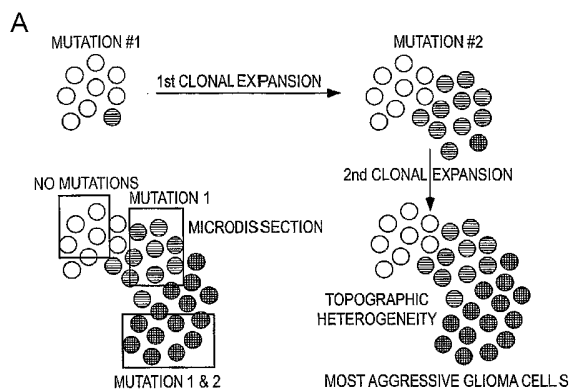
(74) Agents: **MEYER, Mercedes K.** et al.; Drinker Biddle &
Reath LLP, 1500 K Street, N.W., Suite 1100, Washington,
District Of Columbia 20005-1209 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY,
MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO,
NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK,
SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

[Continued on next page]

(54) Title: MOLECULAR ANALYSIS OF CELLULAR FLUID AND LIQUID CYTOLOGY SPECIMENS FOR CLINICAL DIAGNOSIS, CHARACTERIZATION, AND INTEGRATION WITH MICROSCOPIC PATHOLOGY EVALUATION



(57) Abstract: The application relates to methods, materials, kits, and devices for characterizing fluid specimens of any type or source that may or may not contain cells. The methods, materials, kits, and devices can be used for analyzing fluid obtained from various organs or near a purported tumor site, such as via breast ductal lavage or aspiration of a pancreatic cyst. The methods, materials, kits, and devices generate molecular information that can be used in conjunction with clinical, imaging, and microscopic pathology evaluation to provide a superior pathology diagnosis for clinical and research.



WO 2006/047484 A2



FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *without international search report and to be republished
upon receipt of that report*

**MOLECULAR ANALYSIS OF CELLULAR FLUID AND LIQUID CYTOLOGY
SPECIMENS FOR CLINICAL DIAGNOSIS, CHARACTERIZATION, AND
INTEGRATION WITH MICROSCOPIC PATHOLOGY EVALUATION**

5 FIELD OF THE INVENTION

The application relates to methods, materials, kits, and devices for characterizing biological fluid specimens that may or may not contain cells. The methods, materials, kits, and devices generate molecular information that can be used in conjunction with clinical, imaging, and microscopic pathology evaluation to provide a superior pathology diagnosis
10 for clinical and research purposes.

BACKGROUND

Prevention of cancer by its diagnosis and treatment during pre-cancerous stages of development offers the most desired option for reduction and elimination of cancer
15 morbidity and mortality. If detection cannot be achieved in the pre-cancer stage, then diagnosis at an earlier stage of cancer progression offers the best hope for complete cure. In order to achieve this end, sampling techniques such as breast ductal lavage, urine cytology, analysis of fixed cytology fluids such as but not limited to that obtained by liquid cytology preparation, pancreatic juice collection or removal of cerebrospinal fluid become
20 increasingly used to secure representative material of the lining cells or actively proliferating cells that may contain neoplastic cellular elements. However, the effectiveness of cellular fluid analysis is diminished by the inability of a cytology evaluation to provide discriminating, definitive information on the samples obtained thereby (Huber *et al.*, 2004 *Lung Cancer* 45 Suppl. 2: S209-13). Microscopic examination is increasingly recognized as
25 inadequate to address this need to for early and definitive diagnosis of cancer and so other platforms are required to achieve this objective.

At the same time, advances in modern medical imaging techniques, such as but not limited to highly sensitive computerized tomography, magnetic resonance imaging, and endoscopic ultrasound, have resulted in the earlier detection of cancer in smaller sized
30 specimens. Imaging has made possible to aspirate alterations representing precancer and early stage cancer with a high degree of precision when such lesions are small in overall size (Chen *et al.*, 2005 *Biomed. Instrum. Technol.* 39(1):75-85). Representative specimens of these early cancer alterations can now be more easily obtained; however, the gains are offset by the inability to affirm a definitive diagnosis of cancer on these specimens. Hence the
35 need for improved systems for definitive diagnosis of cancer and related states.

Fluid collections that bathe and drain sites of potential cancer formation are

increasingly diagnosed for earlier and better diagnosis. Examples of such fluid collection can include cyst formation within organs, breast ductal lavage fluid, urine, cerebrospinal fluid, and pleural fluid. This list can easily be extended to virtually any part of the body. As a result, pathologists are increasingly called upon to evaluate fluid specimens. However, 5 due to the paucicellular nature of these samples, an indefinite or inadequate diagnosis is often rendered. For example, over 50% of pancreaticobiliary cytobrush cytology specimens are not definitively diagnosed (Khalid et al., 2005 *Clin. Lab. Med.* 25(1): 101-106). Moreover in the 50% with definitive diagnosis, the diagnosis occurs when obvious cancer has already occurred and treatment is limited or ineffective.

10 The human genome project has led to an explosion of genetic information that will eventually lead to the full cataloging of all known human genes, including those involved in cancer development and progression. Specific DNA structural alterations, such as point mutations, genomic deletions, gene copy number losses and gains, gene rearrangements, and epigenetic DNA methylation alterations have been characterized but are not currently used 15 in specimen diagnosis. Thus the demonstration of mutational information can serve to enhance definitive diagnosis when demonstrated to be present in correlation with cellular microscopic alterations. Yet, much of this new genetic information relating to cancer growth is not used in the context to improve definitive disease diagnosis, where it could play a most valuable role. This can be attributed to many factors including the small size and 20 paucicellular nature of the specimen, effect of chemical fixation to optimize morphologic microscopic interpretation, and lack of definitive microscopic criteria for definitive diagnosis.

A rate-limiting step in patient treatment is pathology evaluation of representative specimens wherein the pathologist can give a definitive diagnosis of the presence, type, 25 grade, and character of the neoplasia to the treating physician. Based on this information, the treating physician can then ascertain the best course of treatment. Until recently, this fundamental aspect of disease management was based on microscopic evaluation alone, without taking advantage of the insights derived from molecular genetic analysis. The result of this is subjectivity in specimen diagnosis that often leads to an indeterminate pathology 30 diagnosis, often labeled as "atypical", "suspicious", "not definitive for", or "inadequate for diagnosis". This is especially true for pre-cancerous conditions when microscopic criteria have not yet been universally agreed upon, but where an early diagnosis could be most beneficial (Moinfar et al., 2000 *Cancer* 88(9): 2072-81).

For this specific reason, a patient's diagnosis is often placed on hold, resulting in 35 great cost to the healthcare system and potentially life threatening consequences to the patient, due to delayed or improper treatment. It is not uncommon for a patient to be told by one pathologist that the lesion they may have is benign, while a second pathologist

diagnoses the very same specimen as malignant. These disagreements can be objectively resolved by the incorporation of molecular analysis into pathology practice, as described herein.

Yet, impediments currently exist in implementing molecular and microscopic pathology information. Current clinical research in the area of early cancer detection emphasizes mutation detection. Technologies, such as gene chip microarrays for altered RNA expression, and comparative genomic hybridization for DNA genomic deletion and/or amplification, have been instrumental in delineating a great variety of putative mutational changes in early cancer. However, these methods alone still remain ill-suited for use in assessing specimens of the type procured for the purpose of early cancer detection. For example, cDNA production for expression microarrays and comparative genomic hybridization (CGH) require abundant amounts of fresh tissue or fresh DNA. These requirements cannot be satisfied by the minute amount of cellular material and nucleic acid material present in and derived from fluid sampling techniques.

Furthermore, all tissue or fluid removed for the purpose of disease diagnosis must first be evaluated by a pathologist for adequacy and representative characteristics. Pathologists require that the specimens be optimally fixative-treated and stained for microscopic review. To violate this fundamental rule for specimen processing would invite mishandling of specimens with serious therapeutic and medical malpractice implications. Effective application of molecular analysis requires that the fluid specimens be integrated into pathology practice so that no information is lost.

Reliance upon mutation detection alone has limitations. Precancerous states may have few detectable mutations, but the paucity of cellular material of these states can lead to false positive artifacts in mutation detection that in turn can be misinterpreted as more advanced forms of cancer (*see, e.g., Miller et al., 2002 Genetics 160(1): 357-66*). Reliance upon mutation detection alone will lead to a measure of unreliability in cancer diagnosis that will only impede effective clinical translation of molecular discovery. For this reason, other parameters of early cancer formation, such as the cellular proliferation rate and quality/quantity of DNA released from actively replicating cells, may be useful in the overall molecular analysis of neoplastic and related lesions.

There is also an increased need for personalized patient diagnosis and treatment to achieve better outcomes for the individual. Putting aside issues related to subjective microscopic interpretation, pre-cancer and cancer alterations do not behave in a uniform or predictable manner. Two patients may have lesions which appear to be identical under the microscope, but which may pursue radically different natural courses with one quickly progressing to cancer while the other undergoes slow or no progression over time. Discrimination cannot be reliably found in identifying unique microscopic features. Rather,

the distinction relates to intrinsic molecular alterations unique to each patient's neoplasm that serves as a more rational and effective basis for neoplasia diagnosis, classification, characterization, and treatment. This need to understand precancer and cancer biology at the individual patient level has become a cornerstone of modern cancer treatment and pathology evaluation of specimens must provide the information needed to make more personalize treatment decisions (Mocellin et al., 2005 *Trends Mol. Med.* 2005 11(7):327-35).

One example wherein fluid specimens are frequently used is breast ductal lavage for the purpose of diagnosing a breast cancer. Breast cancer remains a leading cause of mortality despite awareness of its occurrence in the general population, as well as extensive research into its molecular pathogenesis using *in vitro*, animal, and clinical human systems. Pharmaceutical research directed to developing anticancer agents has proceeded actively in recent years. However, the need for effective agents and treatments to treat all types of breast cancer and metastatic breast cancer remains.

Diagnosis of breast tissue is also impacted by observer variability in the subjective interpretation of the sample (Harvey et al., 2002 *Pathology*, 34(5): 410-6; Newman, "Ductal lavage: what we know and what we don't," 2004 *Oncology* 18(2): 179-85, and discussion 185-6, 189, 192).

There is a general consensus, for example, that prevention of breast cancer by its diagnosis and treatment during pre-cancerous stages of development offers the most desired option for managing this disease. If detection cannot be achieved in the pre-cancer stage, then detection at an earlier stage of progression offers the best hope for complete cure. In order to achieve this end, sampling techniques such as breast ductal lavage, have been introduced to secure representative material of the breast duct lining cells or actively proliferating cells in contact with the breast ductal drainage system (O'Shaughnessy, 2003 *Surg. Clin. North Am.* 83(4): 753-69; Locke et al., 2004 *Breast Cancer Res.* 6(2): 75-81; Kenney et al., 2004 *Curr. Oncol. Rep.* 6(1): 69-73; Khan, 2004 *Curr. Treat. Options Oncol.* 5(2): 145-51; Dooley et al., 2001 *J. Natl. Cancer Inst.* 93(21): 1624-32). Based on a washing-out concept similar to bronchoalveolar lavage of the lung (Liloglou, "Cancer-specific genomic instability in bronchial lavage: a molecular tool for lung cancer detection," 2001 *Cancer Res.* 61(4): 1624-8), breast ductal lavage gathers representative material from within the breast ductal system that is then examined to arrive at a definitive diagnosis for early disease. However, the effectiveness of the breast ductal lavage approach is diminished by the inability of pathology evaluation to provide discriminating, definitive information on the samples obtained (Khan, 2004; Dooley et al., 2001). Most current clinical research in the area of early breast cancer detection emphasizes mutation detection. New technologies, such as gene chip microarrays for altered RNA expression and comparative genomic

hybridization for DNA genomic deletion or amplification, have been instrumental in delineating a great variety of putative mutational changes in early breast cancer. However, these methods alone still remain ill-suited for use in assessing specimens of the type procured for the purpose of early cancer detection. For example, cDNA production and comparative genomic hybridization (CGH) generally require abundant amounts of fresh tissue to operate effectively. These needs cannot be satisfied by techniques such as breast ductal lavage and breast fine needle aspiration. Even more fundamentally, all tissue or fluid removed for the purpose of disease diagnosis must first be evaluated by a pathologist for adequacy and representative characteristics; the pathologist requires that the specimens be optimally fixative-treated and stained for microscopic review (*see e.g., Lindsey, 2004 Lancet Oncol. 5(12): 704; Fabian et al., 2004 J. Nat'l. Cancer Inst., 96(20) 1488-9*). To violate this fundamental rule would be to invite mishandling of specimens with serious therapeutic and medical malpractice implications. Effective application of molecular discovery requires that it be integrated into pathology practice so that no information is lost or jeopardized.

Another problem encountered with standard pathology practices is that tissue heterogeneity is often ignored when applying molecular methods to the evaluation of tissue specimens (Hollingsworth et al., 2004 *Am. J. Surg.* 2004 187(3): 349-62; Newman, "Current issues in the surgical management of breast cancer: a review of abstracts from the 2002 San Antonio Breast Cancer Symposium, the 2003 Society of Surgical Oncology Annual Meeting, and the 2003 American Society of Clinical Oncology meeting," 2004 *Breast J.* 10 Suppl. 1: S22-5). Due to the time, labor, and cost associated with cDNA or CGH analysis, usually only a single sample is evaluated for a given patient. This testing of a single sample ignores the fact that breast cancer progression is a stochastic process that proceeds heterogeneously (Kenney et al., 2004 *Curr. Oncol. Rep.* 6(1): 69-73; Moifar et al., 2000 *Cancer Res.* 60(9): 2562-6). Averaging across the heterogeneous tissue by homogenizing a large breast tissue sample sharply contrasts with time honored methods for pathology evaluation of breast tissue, which seeks to detect and characterize microscopic cellular heterogeneity (Kenney et al., 2004; Moifar et al., 2000). This has led to microdissection of more than one tissue target for molecular analysis (O'Connell et al., 1998 *J. Nat'l Cancer Inst.* 90(9): 697-703; O'Connell et al., 1994 *Breast Cancer Res. Treat.* 32(1): 5-12; Larson et al., 1991 *Cancer Res.* 51(12): 3151-5; Sasatomi et al., 2002 *Cancer Res.* 62(9): 2681-9; and Finkelstein et al., 2003 *Hepatology* 37(4): 871-9). Microdissection serves to isolate and optimally purify the predetermined tissue target so that stochastic cancer-associated changes can be detected. The use of multiple targets is congruent with established principles of pathology practice that seek to better understand and correlate topographic variation across the extent of a given neoplastic process.

Reliance upon mutation detection results in certain limitations. Precancerous lesions of the breast and early breast cancer may have few detectable mutations. The relative paucity of cellular material of these states can lead to false positive artifacts in mutation detection that in turn can be spuriously interpreted as more advanced forms of cancer. Reliance upon mutation detection alone will lead to a measure of unreliability in cancer diagnosis that will only serve as a further impediment to effective clinical translation of molecular discovery.

Additionally, accurate diagnosis may not always occur in the instance of reactive cellular proliferative states of the breast, which often is misdiagnosed as breast cancer (*see e.g., Siim et al., 1988 Br. J. Surg. 75(9): 920-1; Branton et al., 2003 Int. J. Surg. Pathol. 11(2): 83-7*). Such misdiagnosis often results in unnecessary treatment and unintended morbidity of the patient. The very same integration of molecular discovery in the context of early cancer detection is ideally suited to an alternative application for reducing the misdiagnosis of reactive states for cancer. This must, however, be accomplished in a manner that is accurate and reliable.

Putting aside issues related to subjective microscopic interpretation, breast pre-cancer and cancer genetic changes do not behave in a uniform or predictable manner (Tavassoli, 2001 *Virchows Arch.* 438(3): 221-7; Masood, 2002 *Microsc. Res. Tech.* 59(2): 102-8). Two patients with the same microscopic-appearing precancerous lesion may pursue radically different natural courses, with one quickly progressing to cancer, while the other patient shows slow or no progression over time. Unique microscopic features are not useful in discriminating between neoplasms. Rather, the distinction relates to intrinsic molecular alterations unique to each patient's neoplasm that serves as a more rational and effective basis for breast neoplasia diagnosis, classification, characterization, and treatment.

The same paradigm holds true for early breast cancer detection. Two tumors that appear to be equivalent may behave in a vastly different manner, including different responses to the same therapy. Microscopic appearance alone is now recognized as inadequate for individual lesion characterization.

Another example of where fluid specimens are taken is urinary bladder cancer via urine cytology. The urinary bladder is a common site of cancer formation which may vary in biological aggressiveness from indolent to aggressive forms with long and short survival respectively. Examination of the urine has served as a means to detect the presence of urinary bladder cancer based on shed cells from the neoplastic lesions within the organ collecting in the urine. While a positive urine cytology for cancer has great specificity (close to 100%) the sensitivity is poor (*see, e.g., Bassi et al., 2005 Urol. Int.* 75(3): 193-200). More importantly, the early stage of bladder cancer development tends to be

universally missed which is precisely the stage of the disease where early detection would have the greatest impact.

Because urinary bladder cancer exists as a spectrum of biological aggressiveness, for patient diagnosis and effective treatment, early detection is needed, especially for the early aggressive forms (Gildea et al., 2000 *Genes Chromosomes Cancer* 2000 27(3): 252-63). While all forms would benefit from early detection, enhanced detection of the indolent forms would have far less impact on healthcare than the aggressive forms. This cannot be accomplished using urine cytology, where the indeterminate rate is exceedingly high inhibiting application of effective treatment.

Many of the concepts described for breast cancer apply equally well to urinary bladder cancer. Microscopic criteria for definitive diagnosis of the early stages have not been agreed upon and are subjective to apply. Small specimen size, such as that found in urine cytology, limits microscopic interpretation. The neoplastic process is heterogeneous in topographic extent, and it is not uncommon to have both early and later forms of the disease present at the same time. It is regrettable that the urine, so easily obtained and containing vital information concerning the state of the urothelium with respect to neoplastic progression, cannot be effectively evaluated using microscopic analysis alone. As such it is ideally suited for an integrated molecular pathology approach such as that described here.

Another example of fluid samples used in cytology wherein historically the information was been limited in what could be determined from liquid cytology cervix scrapings for uterine cancer and uterine cervical dysplasia. Since the introduction of the cervical pap smear technique, the incidence of uterine cervical cancer has dropped greatly. In theory, the disease is fully preventable since evolution of the disease occurs over years affording the opportunity to detect the developing stages before cancer formation. This has not occurred, as a significant number of women die each year of the disease because of inadequate microscopic detection of neoplasia.

Because the pap smear test is performed on such a large population of patients, the contrary problem of over-diagnosis is also important. Once again this can be attributed to the subjective nature of cytology interpretation which relies upon observer experience and ability to render a definitive diagnosis. Over the years several classifications for cervical dysplasia have been proposed and implemented. Current classification known as the updated "Bethesda Classification" has been specifically designed to incorporate pathologist subjectivity. However, patients with cervical cancer continue to slip through the screening program. Here is a representative example where incorporation of molecular findings together with microscopic observations can only serve to improve the current detection capability.

An uterine cervical scraping is not a fluid specimen, but rather a scraping together of cellular material. In the past, the scrapings were smeared on glass slides and stained to form cytology preparations for diagnosis. In so doing, there were a small but significant number of smears which turned out to be inadequate for microscopic review due to smear preparation.

Thus, better methods and means of diagnosing cancer including breast cancer, earlier detection of human malignancy, and a greater ability to distinguish between the types of cancer, pre-cancer, or other abnormal conditions are needed. The materials, methods and kits disclosed address these need and enable the more definitive diagnosis of cancer in cellular fluid specimens.

SUMMARY

The present materials, methods, and kits provide the information needed to advance pathology practice through the molecular characterization of fluid specimens. The results obtained from the fluid specimens can be assessed alone or in combination with the pathology determination (*e.g.*, data from a fixed tissue or histopathology).

Herein are provided methods, compositions, kits, and devices for diagnosing and characterizing fluid specimens that may or may not contain cells thereby improving patient diagnosis and treatment as well as for other clinical and research purposes. The present materials, methods, and kits provide a new modality with which to characterize fluid samples. Although the patient is preferably human, such diagnosis can also be performed on other animals for veterinary diagnosis, for example. It may be expected to have profound consequences for future diagnosis of cancer and pre-cancer states and, by consequence, therapeutic management.

The embodiments described herein serve to provide new methods and means of a method for characterizing a biological fluid sample from a patient for evidence of cancer comprising (a) performing molecular analysis of DNA from the biological fluid sample comprising (i) performing an optical density analysis of the aspirate to determine DNA quantity; (ii) performing a quantitative PCR analysis of the biological fluid sample to determine DNA quality; (iii) performing competitive template PCR to determine DNA quality; and (b) performing mutation analysis of the DNA in the biological fluid sample comprising: (i) determining the presence of mutations in a tumor suppressor gene and/or the presence of a cancer related genetic marker; (ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic

copy; (iii) determining point mutations in the K-ras oncogene and/or point mutations in at least one other cancer-associated gene; (iv) determining other structural alterations in DNA; (v) determining the percentage of mutated DNA from steps (b) (ii) and (b) (iii); and (vi) determining the specific temporal sequence of mutation accumulation based on step (v).

- 5 Finally, embodiments enable the detection of focal areas of more advanced cancer formation and development in specimens that are heterogeneous and contain varying stages of cancer formation.

The method may further comprise confirming the results of steps (a) and (b) by comparing said results with a pathologic analysis of resected tissue obtained from the patient. The method may comprise, for example, analyzing the cyst carcinoembryonic antigen (CEA) level of the biological fluid sample. Also the method describe here enables the objective determination of focal areas of cancer progression within a complex, heterogeneous evolving neoplastic lesion which may contain preexisting, precursor areas as well as more recently advanced higher grade tumor.

15 The DNA in the biological fluid sample may be free-floating or free and adherent to the surface of cells or tissue constituents of the cyst. The cycles of quantitative PCR (qPCR) performed in step (a) (ii) of the method may be greater than a threshold unique for that specific type of DNA. The cycles of quantitative PCR performed in step (a) (ii) may be greater than 30 if no malignancy is present. Preferably, the cycles of quantitative PCR performed in step (a) (ii) may be 29 to 30. More preferably, the cycles of quantitative PCR performed in step (a) (ii) may be less than or equal to 29 if a malignancy is present.

20 The optical density (OD) may be about 2.0 to about 7.5. The optical density may be about 7.5 or higher. Preferably, the loss of heterozygosity has an allele ratio which is two standard deviations beyond the average for the ratio of the specific pairing of polymorphic alleles.

25 Preferably, the patient is a mammal. More preferably, the patient is a human, a domesticated animal, or an agricultural animal. The patient may be suffering from a cancer or dysplasia, a pre-cancerous state, or a non-neoplastic condition. Preferably, the cancer is selected from the group consisting of carcinoma, an epithelial malignancy, a sarcoma, a mesenchymal malignancy, a lymphohematopoietic cancer, an urinary tract cancer, a cervical cancer, a pancreatic cancer, a neuroepithelial cancer, a central nervous system cancer (e.g., a glioma) or other cancer listed herein.

35 Preferably, the pre-cancerous state is selected from the group consisting of epithelial precancerous states including, but not limited to, mucinous cystadenoma, leukoplakia, or serous cystadenoma, and colon polyp, mesenchymal precancerous lesions, neuroglial precancerous lesions, lymphohematopoietic precancerous lesions, and precancerous lesions of other cellular types. Preferably, the non-neoplastic condition is selected from the group

consisting of epithelial non-neoplastic lesions including but not limited to pancreatitis, pancreatic pseudocyst, mesothelial cyst of the pancreas, lymphoepithelial cyst of the pancreas, an ischemic necrosis of the pancreas, mastitis, and non-neoplastic lesions of mesenchymal, neuroglial, lymphohematopoietic, and other cell types.

5 The other DNA structural alterations may be gene amplification, gene translocation, gene rearrangement, and/or epigenetic modification of DNA by DNA methylation.

Also provided is a method for diagnosing and/or determining the prognosis of a cancer, a dysplasia, a pre-cancerous state, or a non-neoplastic condition in a patient comprising: (a) performing molecular analysis of DNA from a biological fluid sample of the
 10 patient comprising: (i) performing an optical density analysis of the biological fluid sample to determine DNA quantity; (ii) performing a quantitative PCR analysis of the biological fluid sample to determine DNA quality; and (iii) performing competitive template PCR to determine DNA quality; and (b) performing mutation analysis of the DNA in the biological fluid sample comprising: (i) determining the presence of one or more mutations in a tumor
 15 suppressor gene and/or the presence of a cancer related genetic marker; (ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or
 20 structural alterations of each allelic copy; (iii) determining point mutations in the K-ras oncogene and/or point mutations in at least one other cancer-associated gene; (iv) determining other structural alterations in DNA; (v) determining the percentage of mutated DNA from steps (b)(ii) and (b)(iii); and (vi) determining the specific temporal sequence of mutation accumulation based on step (v); and diagnosing and/or determining the prognosis
 25 of a cancer or dysplasia, a pre-cancerous state, or a non-neoplastic condition in the patient in need based on the results of steps (a) and (b). Finally, embodiments enable the detection of focal areas of more advanced cancer formation and development in specimens that are heterogeneous and contain varying stages of cancer formation.

Preferably, the cancer or dysplasia for the methods, compositions, and kits is
 30 selected from the group consisting of a carcinoma, an epithelial malignancy, a sarcoma, a mesenchymal malignancy, a lymphohematopoietic cancer, a neuroepithelial cancer, and a central nervous system cancer. Preferably, the pre-cancerous state is selected from the group consisting of wherein the pre-cancerous state is an epithelial precancerous state including but not limited to mucinous cystadenoma, leukoplakia, or serous cystadenoma,
 35 and colon polyp, mesenchymal precancerous lesions, neuroglial precancerous lesions, lymphohematopoietic precancerous lesions, and precancerous lesions of other cellular types. Preferably, the non-neoplastic condition is epithelial non-neoplastic lesions including but

not restricted to pancreatitis, pancreatic pseudocyst, mesothelial cyst of the pancreas, lymphoepithelial cyst of the pancreas, an ischemic necrosis of the pancreas, mastitis, and non-neoplastic lesions of mesenchymal, neuroglial, lymphohematopoietic, and other cell types.

5 Also provided is a method for determining a course of treatment for a cancer or dysplasia, a pre-cancerous state, or a non-neoplastic condition of a patient comprising: (a) performing molecular analysis of DNA from a biological fluid sample from the patient comprising: (i) performing an optical density analysis of the biological fluid sample to determine DNA quantity; (ii) performing a quantitative PCR analysis of the biological fluid
10 sample to determine DNA quality; and (iii) performing competitive template PCR to determine DNA quality; and (b) performing mutation analysis of the DNA in the biological fluid sample comprising: (i) determining the presence of mutations in a tumor suppressor gene and/or the presence of a cancer related genetic marker; (ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or
15 other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy; (iii) determining point mutations in the K-ras oncogene and/or point mutations in at least one other cancer-associated gene; (iv) determining other structural
20 alterations in DNA; (v) determining the percentage of mutated DNA from steps (b)(ii) and (b)(iii); and (vi) determining the specific temporal sequence of mutation accumulation based on step (v), and wherein the results of steps (a) and (b) are used to determine a course of treatment for a patient suffering from a cancer or dysplasia, a pre-cancerous state, or a non-neoplastic condition. Also, embodiments enable the detection of focal areas of more
25 advanced cancer formation and development in specimens that are heterogeneous and contain varying stages of cancer formation

 Also provided is a kit for determining the diagnosis and/or prognosis of a patient comprising: (a) a device for performing molecular analysis of DNA in the biological fluid sample from the patient; (b) a device for performing mutation analysis of the DNA in the
30 biological fluid sample from the patient; (c) reagents for performing the molecular analysis and the mutation analysis. The device may further optionally comprise a device for recording data obtained from the molecular analysis of DNA and the mutation analysis of the DNA. Preferably, the device for recording data records patient age, patient sex, patient medical history, prior cancer history of patient, patient weight, patient family history of
35 cancer and disease, diagnosis determined from the mutation analysis of the DNA and molecular analysis of the DNA, a temporal sequence of mutation accumulation, and proposed treatment based on determined diagnosis.

In another embodiment, a method is disclosed for determining and characterizing an anomaly such as but not limited to that arising in the breast in a patient comprising: (a) performing a molecular analysis of DNA from a biological sample from a patient's breast tissue comprising: (i) performing optical density analysis of the biological sample to
 5 determine DNA quantity; (ii) performing quantitative PCR analysis of the aspirate to determine DNA quality; (iii) performing competitive template PCR to determine DNA quality; and (b) performing mutation analysis of the DNA of the biological sample comprising the steps of: (i) determining the presence of mutations in a tumor suppressor gene and/or the presence of a cancer related genetic marker in the DNA of the biological
 10 sample; (ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy; (iii) determining point
 15 mutations in the K-ras oncogene and/or point mutations in at least one other cancer-associated gene; (iv) determining copy number alterations such as homozygous deletion and oncogene amplification; (v) determining other structural alterations in DNA; (vi) determining the percentage of mutated DNA from steps (b)(ii) and (c)(iii); and (vii) determining a temporal sequence of mutation accumulation based on step (v); and (c)
 20 assessing the data from steps (a) and (b) to determine the presence of a breast anomaly and characterize the type of breast anomaly. Also contemplated are aspects which the detection of focal areas of more advanced cancer formation and development in specimens that are heterogeneous and contain varying stages of cancer formation.

The biological sample may be an aspirate from a cyst from the breast or any other
 25 organ. A breast anomaly includes but is not limited to a breast cancer, a breast dysplasia, a pre-cancerous breast condition, or a non-neoplastic condition. A breast cancer includes but is not limited to invasive ductal carcinoma, invasive lobular carcinoma, breast sarcoma, ductal adenocarcinoma, breast acinar cell carcinoma, metastatic cancer involving the breast, recurrent breast cancer, cystosarcoma phyllodes, Paget's disease, and metastatic breast
 30 cancer. A metastatic breast cancer includes but is not limited to an axillary metastasis with no apparent primary tumor or an axillary adenopathy of a metastatic adenocarcinoma.

A breast dysplasia can be atypical ductal hyperplasia or atypical lobular hyperplasia. A pre-cancerous breast condition includes but it not limited to mucinous cystadenoma, serous cystadenoma, mucinous duct ectasia, intraductal papillary mucinous neoplasm, and
 35 breast intraepithelial neoplasia (PanIN grades 1-3), and a papilloma of the breast. A non-neoplastic condition of the breast includes but it not limited to a pseudocyst, fibrocystic disease of the breast, a fibroadenoma, inflammatory mastitis, soft tissue trauma, a

lymphedema, and a mesothelial cyst.

The method may further comprise confirming the results of steps (a) and (b) by comparing said results with a pathologic analysis of a cystic lesion surgically obtained from the patient. The method may also further comprise analyzing the carcinoembryonic antigen (CEA) level of the aspirate.

In a further embodiment, a method is disclosed for diagnosing and/or determining the prognosis of a breast anomaly in a patient suffering from breast cysts comprising the steps of the previous method, wherein a diagnosis and/or prognosis of the patient is determined from analyzing steps (a) and (b). The patient can be one who suffers from breast cysts. The breast anomaly may be a breast cancer, a breast dysplasia, a pre-cancerous breast condition, or a non-neoplastic condition. The breast dysplasia may be atypical ductal hyperplasia or atypical lobular hyperplasia. The pre-cancerous breast condition may be a mucinous cystadenoma, a serous cystadenoma, a mucinous duct ectasia, an intraductal papillary mucinous neoplasm, a breast intraepithelial neoplasia (PanIN grades 1-3), and a papilloma of the breast. The non-neoplastic condition can be a breast pseudocyst, a fibrocystic disease of the breast, a fibroadenoma, inflammatory mastitis, soft tissue trauma, a lymphedema, and a mesothelial cyst.

In a further embodiment, a method is disclosed for determining an effective course of treatment for a breast anomaly in a patient using the steps discussed herein for diagnosing the type of breast anomaly, and wherein the breast anomaly is malignant, ascertaining the progression of the malignancy from information obtained from steps (a) and (b) to determine effective treatment for the diagnosed cancer. Preferably, the treatment is selected from the group consisting of chemotherapy, radiation therapy, hormone therapy, bone marrow transplant therapy, surgery, and a combination of any one or more of these therapies.

Preferably, the breast anomaly is a breast cancer, a breast dysplasia, a pre-cancerous breast condition, or a non-neoplastic breast condition. Preferably, the breast dysplasia is atypical ductal hyperplasia or atypical lobular hyperplasia. Preferably, the pre-cancerous breast condition is selected from the group consisting of mucinous cystadenoma, serous cystadenoma, mucinous duct ectasia, intraductal papillary mucinous neoplasm, and breast intraepithelial neoplasia (PanIN grades 1-3), and a papilloma of the breast. Preferably, the non-neoplastic condition is selected from the group consisting of breast pseudocyst, fibrocystic disease of the breast, fibroadenoma, inflammatory mastitis, soft tissue trauma, lymphedema, and mesothelial cyst.

In yet another embodiment, a system is disclosed for determining and characterizing an anomaly of a patient comprising: (a) a means for performing a molecular analysis of the purified DNA from the biological sample, wherein said molecular analysis means obtains an

optical density of the biological sample to determine DNA quantity, comprises means for performing quantitative PCR to determine DNA quality, and comprises means for performing competitive template PCR to determine PCR quality, and wherein said functionally associated means for performing the molecular analysis is performed serially or
 5 in parallel; (b) a DNA mutation analysis means for detecting (i) a mutation in a tumor suppressor gene, a cancer related genetic marker, wherein said DNA mutation analysis means analyzes the purified DNA of the biological sample; (ii) a loss of heterozygosity in a tumor suppressor gene or in other polymorphic markers linked to tumor suppressor genes with respect to their allelic balance; (iii) a mutation in K-ras or a point mutation in a breast
 10 cancer-associated gene; (iv) and determining copy number alterations; (v) and determining other structural alterations in DNA; (vi) and determining the percentage of mutated DNA from means for (b)(ii) and (b)(iii); and (c) means for chronologically listing accumulation of the mutations identified by means of (b)(v); and wherein said system generates an output that determines and characterizes the anomaly of the patient. The anomaly can be a breast
 15 anomaly, pancreatic anomaly or any of the diseases listed herein.

In yet a further embodiment, an analytical platform is disclosed for diagnosis of an anomaly in a patient comprising: (a) an integrated means for preparing DNA from a biological sample; (b) an integrated means for analyzing DNA quality from said prepared DNA from said biological sample; (c) an integrated means for analyzing DNA quantity
 20 from said prepared DNA from said biological sample; (d) an integrated means for analyzing said prepared DNA from said biological sample for one or more mutations; (e) an integrated means for determining a time course of accumulation of said one or more mutations; and (f) an integrated means for presenting data produced by said means of (a) to (e) for rendering a diagnosis.

25 A further embodiment enables the detection of focal areas of more advanced cancer formation and development in specimens that are heterogeneous and contain varying stages of cancer formation. This is based on the differential quality of DNA which is better from higher grade neoplasia, where cell turnover is more rapid than from lower grade, more slowly replicating cellular areas.

30 In a further embodiment, methods are provided that enable the analysis of urine for the detection and characterization of precancer and cancerous changes arising in the urinary bladder or anywhere along the flow of urine from its inception to elimination from the body. The methods apply not only the urine but to any collection of flow of bodily fluid that may be cellular, paucicellular or even lacking intact cells in which case the analysis operates on
 35 the free DNA in the fluid.

In a further embodiment, methods are provided that enable the analysis of liquid cytology samples derived from any site throughout the body and formulated into a liquid

specimen with fixative for the purpose of providing molecular information to complement cytology microscopic evaluation.

- Yet another embodiment provides for a method of analyzing a liquid cytology sample for the presence of a mutational change in a patient comprising: (a) performing a molecular analysis of DNA from the liquid cytology sample from the patient comprising:
- 5 (i) performing optical density analysis of the liquid cytology sample to determine DNA quantity;
 - (ii) performing quantitative PCR analysis of the liquid cytology sample to determine DNA quality;
 - 10 (iii) performing competitive template PCR to determine DNA quality; and
 - (b) performing mutation analysis of the DNA of the liquid cytology sample comprising:
 - (i) determining the presence of mutations in a tumor suppressor gene and/or the presence of a cancer related genetic marker in the DNA of the liquid cytology sample;
 - 15 (ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy;
 - 20 (iii) determining point mutations in a K-ras oncogene and/or point mutations in at least one other cancer-associated gene;
 - (iv) determining copy number alterations such as homozygous deletion and oncogene amplification;
 - (v) determining other structural alterations in DNA;
 - 25 (vi) determining the percentage of mutated DNA from steps (b)(ii) and (b)(iii); and
 - (vii) determining a temporal sequence of mutation accumulation based on step (v); and
 - (c) assessing the data from steps (a) and (b) to determine the presence of a mutational change in a patient and (d) optionally further determining whether a focal site of neoplastic progression of the neoplastic lesion is more advanced than other portions of the lesion. In another aspect, a method is provided for analyzing a fluid biological sample, with or without a cell, for detecting focal site of neoplasia progression in a patient with or without a large component of neoplastic or normal tissue comprising: (a) performing a molecular analysis of DNA on the liquid biological sample comprising:
 - 35 (i) performing optical density analysis of the biological sample to determine DNA quantity;
 - (ii) performing quantitative PCR analysis of the aspirate to determine DNA quality;

- (iii) performing competitive template PCR to determine DNA quality; and
- (b) performing mutation analysis of the DNA of the liquid biological sample comprising:
- (i) using DNA primers to amplify an amplicon of at least 300 nucleotides to 1500 nucleotides, thereby amplifying DNA from more actively replicating tissue sites and excluding degraded DNA from amplification;
 - (ii) determining mutation presence in a tumor suppressor gene and/or mutation presence in a cancer related genetic marker in the DNA of the liquid biological sample;
 - (iii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy;
 - (iv) determining point mutations in the K-ras oncogene and/or point mutations in at least one other cancer-associated gene;
 - (v) determining copy number alterations such as homozygous deletion and oncogene amplification;
 - (vi) determining other structural alterations in the DNA of the liquid biological sample;
 - (vii) determining the percentage of mutated DNA from steps (b)(ii) and (b)(iii); and
 - (viii) determining a temporal sequence of mutation accumulation based on step (v); and
- (c) determining the focal site of neoplasia progression in the patient by the steps of (a) to (b). It is noted that the amplicon can be at least 300 nucleotides or more, more preferably greater than 400 nucleotides, more preferably 500 nucleotides, and yet more preferably 600 nucleotides or longer. Optimally the length is some integer value between 300 nucleotides and 1500 nucleotides (a "long length amplicon"), which thereby exclude degraded DNA from being amplified.

- Another aspect provided herein is a method of detecting an allelic imbalance in a patient without an internal normal, non-neoplastic specimen from the patient comprising:
- (a) performing a molecular analysis of DNA from the biological sample comprising:
- (i) performing optical density analysis of the biological sample to determine DNA quantity;
 - (ii) performing quantitative PCR analysis of the DNA quality;
 - (iii) performing competitive template PCR to determine DNA quality; and
- (b) performing mutation analysis of the DNA of the biological sample comprising:
- (i) using DNA primers to amplify a polymorphic region of DNA;

- (ii) defining a combination of specific polymorphic alleles in specimens from the general population
- (iii) calculating quantitatively content ratio of polymorphic alleles in specimens from the general population;
- 5 (iv) defining a statistical average and normal distribution of the content ratio of each combination of polymorphic alleles in specimens from the general population;
- (v) determining a threshold for designation as within normal limits;
- (vi) defining content ratios exceeding threshold values as outside the range of content ratio variation in specimens from the general population; and
- 10 (c) calculating a percentage of cellular DNA subject to allelic imbalance in step (vi) using the average content ratio as a normalizing factor. In one embodiment, the amplification is at least 500 to 1500 nucleotides long. In another embodiment, the threshold is at least about 95%. In a different embodiment, the polymorphic region of DNA amplified is a microsatellite, a minisatellite or a single nucleotide polymorphism.

15

BRIEF DESCRIPTION OF THE DRAWINGS

It is to be understood that both the foregoing description and the following detailed description are exemplary and explanatory only, and are not restrictive of the material methods, devices, and kits. The accompanying drawings, which are incorporated herein by reference, and which constitute a part of this specification, illustrate certain embodiments and together with the detailed description, serve to explain the principles of the materials, methods, devices, and kits. The drawings are exemplary only, and should not be construed as limiting the materials, methods, devices and kits described herein.

FIG. 1. Left hand, upper panel of three tissue samples depicts an example of manual stereoscopic microdissection of pancreatic adenocarcinoma from tissue sections. Right hand, upper 4 panels depicts microdissection of cytology cell clusters directly from stained cytology slides after removing cover slips. Lower panel presents tabulated data of a topographic genotype (TG) performed on a pancreatic duct brushing having discrete clusters of representative cells. The data from two cases are shown in which the microdissected cytology was performed in triplicate and compared to the resected specimen. Near total concordance was found validating the approach used to integrated molecular analysis in cytology specimens.

FIG. 2. The figure shows an illustration of the time course of mutation acquisition.

FIG. 3. Panels 3A and 3B show diagrams delineating the time course of mutation acquisition.

FIG. 4. The figure shows the traditional approach to classifying gliomas. This approach is based solely on histopathologic features emphasizing the presence or absence of

oligodendroglial differentiation. When clearly present, this is generally not a problem. However, when oligodendroglial differentiation is present in partial or minor degrees, there can be significant subjectivity in its appreciation and impacting classification. For example, diagnosis of mixed glioma, oligoastrocytoma, astrocytoma with oligodendroglial features, etc., can be used in such instances. In our proposed integrated molecular pathology classification based on topographic genotyping of multiple targets, the critical element is whether 1p allelic loss affects a long segment of the telomeric portion of 1p as the first detectable mutation. This denotes the treatment responsive subset that will not necessarily be adversely influenced by later acquisition of aggressive mutational damage involving 9p, 10q, and 17p. The presence of 1p deletions of any genomic extent or timing will confer an oligodendroglial character to the morphology, which in turn may be subjectively appreciated.

FIG. 5. The figure shows analytic validation and quality control measure, operating at two levels. The first is with respect to the sample which is replicated in a fashion that must lead to very close findings. The second operates at the individual marker level with replication. The variability between individual results that are replicated can be used as a means to compare the significance of results.

FIG. 6. The figure shows an electropherogram illustrating marker replication. Panel 1A and 1B are two separate microdissections from the same slide for one tetranucleotide microsatellite marker. The results are as expected, nearly identical. Panels 2A and 2B are the same microdissections as those of panels 1A and 1B, but were analyzed using a second marker. Again, the results as expected are nearly identical.

FIG. 7. The figure shows allelic dropout. Note that the ratio of the allele peak heights is highly variable. Yet, these are replicate analysis of the same marker from the same source of DNA. Quality control measures are essential to guard against this phenomenon, which will result in false positive results.

FIG. 8. PCR Amplification Data of k-ras-2 for long length amplicons. The lower panel shows the number of Ct cycles performed to obtain the amplicon. Poor quality DNA produced no long length amplicons (samples A01 and F01).

DETAILED DESCRIPTION

Novel methods, reagents, and scientific approaches that improve the analysis of fluid specimens including liquid cytology specimens are presented herein. Improvements are provided in reaching a definitive diagnosis and prediction of cancer, and more effectively distinguishing neoplasia from the diagnosis of similar appearing, but distinct, breast non-neoplastic abnormalities and pre-cancerous conditions as determined from analyzing biological fluid samples. The methods and devices herein can be used in

combination with pathology determination made on fresh specimens (*e.g.*, biopsied tissue and frozen sections) as well as on minute, fixative-treated, stained specimens. Finally important quality control validation methods are provided to monitor analytic precision of the analysis.

5 These biological fluid samples can be very small in volume, which is often all that is available in clinical medical practice. The effect will be to improve the detection, diagnosis, management and prevention of cancer in a reliable, timely and cost effective manner. Individual patients will benefit greatly by earlier and more objective diagnosis, and significant financial savings will be realized for the healthcare system. The applicability to
10 cell-free fluid specimens is noteworthy, as currently there are only limited means available to analyze these types of specimens. As a result, this type of specimen is usually discarded despite the fact that it may well contain valuable discriminating information.

 While mutation detection is regarded as a means to characterize cancer, assessment of the quality and quantity of DNA has never been employed to achieve this objective.
15 Assessment of DNA is not intuitive. The emphasis had always been on detection of mutations, which is suitable for detecting advanced forms of cancer that bear abundant mutational damage. Many studies have documented the striking difference with respect to the presence or absence of mutations when comparing advanced cancer to a non-cancerous process of the same tissue type. (C. F. Taylor, et al., 2004 *Methods Mol. Med.* 92:44; R. G.
20 Cotton, 2000 *Eur. J. Pediatr.* 159 Suppl. 3: S179-82). This technique is preferably utilized in discriminating advanced stages of neoplasias wherein mutation accumulation is greater. The methods, materials, and kits can also be used to objectively separate early cancer from indolent forms of neoplasia (benign tumors) and from similar appearing non-neoplastic states.

25 Aspects of the methods here utilize any fluid sample, such as but not limited to liquid cytology. Liquid cytology can be scrapings for all type of brushings, aspiration, fluid collection and other forms of cytology specimens. The liquid cytology format provides the opportunity to obtain a residual fluid specimen with or without significant cellular material (*e.g.*, Lorincz, 2003 *Salud Publica Mex.* 45 Suppl. 3: S376-87). As such, it is ideally suited for
30 molecular analysis using the approach outline in this application. This is an unexpected bonus of the liquid cytology format; to improve standard cytology practice by reducing inadequate specimens and to afford the opportunity for integrated molecular pathology evaluation. It should be noted that direct microdissection of cytology smears can also be performed using methods such as Topographic Genotyping™ (see U.S. Patent No.
35 6,340,563). The methods described here provide valuable discriminating information but also enhance standard tissue histopathology/cytology when microdissection based analysis can be performed.

As for other examples of fluid analysis, the same considerations of tissue heterogeneity, small sample size, and requirement for fixative treatment apply. It is also important to note that neoplasia is a dynamic process that evolves over time with progression from the normal state, often through dysplasia and into malignancy (Marcucci et al., 2005 *Curr. Opin. Hematol.* 1(1): 68-75). The process is stochastic as focal areas with newly acquired mutations producing growth advantage clonally expand and proliferate more rapidly to dominate a particular lesion. At any point in time however, the cellular content and the DNA derived from it represent a mixture of preexisting precursor lesion and focally expanding progressively more aggressive neoplasia. Fluid accumulations or fluid in contact with such a process will reflect the full range of cellular progression. This reality makes it essential that diagnostic techniques that evaluate associated fluid collections, whether cellular, paucicellular, or cell-free, be evaluated in a way that makes clear the presence of this dynamic evolution. Provided herein are methods and materials to accomplish consistent analysis of cellular, paucicellular, or cell-free fluid samples.

Liquid cytology has emerged as a general platform to evaluate all types of cytology specimens whether derived from solid or fluid sources. Fine needle aspirates (FNA) of solid masses are typically ejected into liquid cytology fixative after which the specimen is handled in the same fashion to that described for cervical brushing. Similarly brushings from hollow viscid such as the biliary tract, pancreatic duct, ureter or other sites can be immersed in liquid cytology fixative which causes cells to become released from the brush and move into the liquid cytology fluid. The methods described can be used to more effectively any type of fluid or cytology sample including those from entirely solid lesions.

Other fluid sources, no matter how paucicellular, can be effectively analyzed to integrate molecular and microscopic information from the same case and from the same specimen. This would include and not be limited to collected cerebrospinal fluid, aspirated cyst fluid, pleural fluid, pericardial fluid, ascites fluid, collected pancreatic juice, collected urinary tract fluid from any site from its formation, collected bile, aspirated joint fluid, and fluid within or associated with sites of cancer formation. The fluid may be analyzed directly by the methods outline below, or via liquid cytology after the fluid has been mixed with a suitable fixative for cytology preparation.

Analytical platforms and methods of use are also described which enhance current improvements in imaging and advances in specimen sampling methods. The analytical platform includes any one or more of the following:

- (1) techniques to separate any cells or tissue present in the fluid from the cell free fluid for separate molecular pathology analysis;
- (2) techniques to measure the quantity of DNA serving as means to infer the proliferative rate of neoplastic states in continuity with fluid collection;

(3) techniques to measure the quality of DNA including the extent of its breakdown into progressively shorter pieces of nucleic acid serving as means to infer the proliferative rate of neoplastic states in continuity with fluid collection;

5 (4) mutation analysis of minute specimens affording determination of genomic and specific gene copy number imbalance;

(5) mutation analysis of minute specimens affording determination of point mutational change of nucleic acids in the sample;

(6) mutation analysis of minute specimens affording determination of specific gene amplification and/or homozygous gene loss;

10 (7) mutation analysis of minute specimens affording determination of microsatellite instability;

(8) mutation analysis of minute specimens affording determination of gene rearrangement;

15 (9) mutation analysis of minute specimens affording determination of genomic deletional expansion;

(10) mutation analysis of minute specimens affording determination of altered DNA methylation status uniquely innovated to operate on fixative-treated tissue sections and cytology smears;

20 (11) techniques to determine the time course of mutation accumulation providing a dynamic understanding of molecular change;

(12) quality control measures to monitor the reliability and representativeness of mutation detection consisting of recommendations for replicate analysis at both the specimen level (sample replication) and/or replication of the individual marker results (*i.e.*, marker replication);

25 13) techniques to optimally handle liquid cytology specimens so that they may provide highly detailed and correlative molecular information in the context of microscopic evaluation; and

30 14) techniques to detect the presence of focal transformation of increasing grade of neoplasia in fluid from a heterogeneous cellular specimen with both higher and lower grade areas of neoplasia.

In particular, a preferred analytical platform has been designed to correlate with and improve histopathologic and cytologic observation, complementing and improving existing pathology practice. The materials, methods and kits disclosed permit the objective diagnosis of invasive cancer, low- and high-grade forms of dysplasia, and neoplastic and
35 non-neoplastic conditions. The methodology comprises detailed molecular analysis incorporating DNA quality and quantity, point mutational analysis of oncogenes and other genes and other forms of mutational damage correlated with cancer, a broad spectrum of

tumor suppressor genes and other breast cancer associated genes linked to microsatellite loss of heterozygosity (LOH), and new approaches for accurate copy number determination of genes and oncogenes (copy number analysis). Another aspect provides methods to clearly discriminate cancer from non-cancer states, such as those produced by inflammation, infection, and trauma to a particular organ or tissue. Methods for diagnosing, determining prognosis of, and defining a course of treatment for cancer, high-grade dysplasia, pre-cancerous states, and non-neoplastic conditions are also provided.

An important component of relates to the quality control assessment of small amounts of DNA that, due to paucity of diagnostic material, may produce both false-positive and false-negative results for mutation detection (C. R. Miller, et al., 2002 *Genetics* 160(1): 357-66; J. C. Dreesen, et al., 1996 *J. Assist. Reprod. Genet.* 13(2): 112-4). Standard pathology practice addresses this need by extracting purified DNA from specimens and measuring its content by optical density (OD) determination. Unfortunately, extracted DNA from fixative-treated, stained specimens is vanishingly low, e.g., below about 1 ng for fixative-treated, paraffin-embedded specimens measuring 1 cm in size or less. Although quantitative polymerase chain reaction (qPCR) has been employed on fixative-treated tissue, qPCR assumes high quality starting nucleic acid content, which cannot be assumed when dealing with fixative-treated, stained specimens of limited size. Methods and devices which address the need for analytic validation and quality control are provided herein.

Non-neoplastic fluid collections have very low amounts of DNA present in the fluid, because the lining cells contain very few cells that replicate slowly. Low-grade, slow-growing, neoplastic anomalies would generate an associated fluid that contained relatively greater amounts of DNA reflecting the greater turnover of replicating lining cells. Finally, a malignant tumor is believed to possess the greater amounts of DNA of high quality and intactness in fluid localized near the tumor.

To confirm this belief, a novel approach as set forth herein was created to quantitatively define the amount and integrity of fluid DNA, and then demonstrate statistically significant thresholds that would separate non-neoplastic, indolent, and malignant states from each other. The combined applications of these concepts to pathology analysis of fluid specimens and the specific methods used were uniquely created to address these needs.

The approach includes combining the direct quantitation of extracted DNA using optical density (OD) measurement techniques, qPCR determination of DNA concentration, and competitive template PCR reaction involving defined genomic segments of the glucocerebrosidase (GCS) gene and its pseudogene (other genes and their associated pseudogenes can be used alone or in combination with *GCS* and its pseudogene) (Martinez-Arias et al., 2001 *Hum Mutat.* 17(3): 191-8). The qPCR reaction is not used to measure

DNA concentration, because that is accomplished by the OD measurement. The OD concentration is used to standardize the qPCR reaction and other mutational analyses to a starting concentration of about 5 ng/ μ L, although other levels for normalization can be used with equal effectiveness. The variation in the qPCR reaction then serves to measure the degree of intactness of DNA, since the highest qPCR values for DNA starting concentration will be seen when DNA has not undergone degradation. This use of the qPCR provides data on DNA integrity.

The CT-PCR reaction, performed in triplicate (or more) at varying concentrations, provides quantitative information on the quality of DNA with respect to degradation, and quantitative information on the representative amplifiability of the DNA. The CT-PCR reaction provides a sensitive means to quantitatively characterize DNA degradation in a fashion that is essentially independent of the various phases of nucleic acid amplification (exponential phase, plateau phase) and is effective on minute specimen samples. By means of gene/pseudogene sequences that are nearly identical, but differ by varying lengths of base deletion, the effect of DNA degradation can be accurately characterized. This can be done, for example, using the GCS gene and its pseudogene or other gene/pseudogene combination. By performing CT-PCR in triplicate at different concentrations of starting template, the effect of allelic dropout can be effectively controlled, and the degree of degradation can be accurately defined. Adequacy of amplifiable DNA is reflected by replicate reliable amplification of both the shorter and longer sized amplicons, while low amounts of inadequate levels of amplifiable DNA for mutational analysis is represented by variability in the effectiveness of relative amplification of the longer sized amplicon. The use of CT-PCR as described herein is a novel approach providing the critical information needed to effectively analyze minute clinical tissue specimens subject to tissue fixation and staining.

Also described herein are methods for determining specific gene copy number also specially adapted for use with minute fixative-treated, stained cellular samples and fluid specimens, but not restricted to these types of specimens. The DNA sequence of a gene of interest and a comparator gene are searched for homology to direct the creation of oligonucleotide primers for duplex amplification. More importantly, the amplicons generated are also designed to be of approximate but not equal length, thereby enabling a stoichiometric relationship to be developed between them. If both genes are present at the same copy number, then the amplicons generated will be equal. If the gene-of-interest undergoes amplification, then its amplicon is in relatively greater excess than the comparator gene. Conversely, if the gene-of-interest undergoes homozygous loss, then the comparator amplicons will dominate. This unique approach is one which can provide quantitative information on specific gene copy number effective on virtually any type of

specimen.

These innovations are combined with several standard procedures to produce a quantitative system for molecular analysis of clinical specimens for diagnostic and other purposes. The system is entirely complementary to standard pathology practice. It does not
5 compete with current pathology practices, but rather advances them, eliminating the roadblock that currently exists to advancing patient medical care.

1. Acronyms and Definitions

Acronyms and definitions as provided herein are art accepted unless indicated
10 otherwise.

1.1 Acronyms

In accordance with this detailed description, the following abbreviations and definitions apply unless indicated otherwise in the specification:

	ABL1	proto-oncogene tyrosine-protein kinase ABL1
15	ALCL	anaplastic large cell lymphoma
	ALL	acute lymphoblastic leukemia
	AML	acute myelogenous leukemia
	APC	adenomatous polyposis coli protein
	APL	acute promyelocytic leukemia
20	Bax	Bcl-2 associated x protein
	Bcl-2	apoptosis regulator Bcl-2
	Bcl _{x_L}	Bcl apoptosis regulator Bcl-X (also known as Bcl _{x_S})
	BL	Burkitt's lymphoma
	BRCA-1	breast cancer and ovarian cancer-1
25	BRCA-2	breast cancer and ovarian cancer-2
	CCND1	G1/S specific cyclin D1 (also known as Bcl-1)
	CDKN2A	cyclin-dependent kinase 4 inhibitor A
	CDK2	cell division protein kinase 2 (also known as cyclin dependent kinase 2)
30	CDK4	cell division protein kinase 4 (also known as cyclin dependent kinase 4)
	CDK5	cell division protein kinase 5 (also known as cyclin dependent kinase 5)
	cDNA	complimentary deoxyribonucleic acid
35	CEA	carcinogenic antigen
	CEL	carboxy ester lipase
	CGH	comparative genomic hybridization

	CLL	chronic lymphocytic leukemia
	CML	chronic myelogenous leukemia
	CRK	proto-oncogene C-cr ~~k~~
	CT	computed tomography
5	Ct	threshold value for qPCR produced detection
	CTCL	cutaneous T-cell lymphoma
	CT-PCR	competitive template PCR
	DCIS	ductal carcinoma <i>in situ</i>
	DLC1	Rho-GTPase-activating protein 7
10	DLCL	diffuse large cell lymphoma
	DPC4	mothers against decapentaplegic homolog 4 (also known as SMAD4)
	EGF	epidermal growth factor
	EGFR	epidermal growth factor receptor
15	EMSI	src substrate cortactin
	ErbB2	receptor tyrosine-protein kinase ErbB-2 precursor (also known as c-erbB-2, HER2, HER-2/neu)
	ETS1	c-ets-1 protein
	ETS2	c-ets-2 protein
20	FES	proto-oncogene tyrosine-protein kinase Fes/FPS
	FGF	fibroblast growth factor
	FGF-2	fibroblast growth factor-2
	FGF-3	fibroblast growth factor-3
	FHIT	Bis (5'-adenosyl)-triphosphatase
25	FISH	fluorescence <i>in situ</i> hybridization
	FL	follicular lymphoma
	FNA	fine needle aspiration
	FNAB	fine needle aspiration biopsy
	GCS	glucocerebrosidase
30	HER	human epidermal growth factor receptor
	HGH	hepatocyte growth factor
	HIF-1 α	hypoxia-inducible factor -1 α
	HIF-1 α	hypoxia-inducible factor-1 α
	Int2	Int-2 proto-oncogene protein precursor (also known as FGF-3 or fibroblast-growth factor-3)
35	ITS	insulinoma tumor suppressor gene locus
	LCAM	epithelial-cadherin precursor (also known as CDH1)

	LOH	loss of heterozygosity
	LPL	lymphoplasmacytoid lymphoma
	MAF	transcription factor maf
5	MAFK	transcription factor mafK (also known as P18, erythroid transcription factor NF-EZ, p18 subunit)
	MCA	mucinous cystadenoma
	MCAC	mucinous cystadenocarcinoma
	MCC	colorectal mutant cancer protein
	MCL	mantle cell lymphoma
10	Mdm2	ubiquitin-protein ligase E3 mdm2
	MEN-1	menin
	MIB1	mindbomb homolog 2
	MRI	magnetic resonance imaging
	Mtus1	mitochondrial tumor suppressor 1
15	N	clinical analysis
	NF-1	neurofibromatosis-1
	NF-2	neurofibromatosis-2
	NME	nucleoside diphosphate kinase A
	OD	optical density
20	PARK2	Parkin (Parkinson's disease protein 2)
	PCNA	proliferating cell nuclear antigen (also known as cyclin)
	PCR	polymerase chain amplification reaction
	pN	pathologic analysis
	PTEN	phosphatidylinositol-3,4,5-triphosphate 3-phosphatase
25	qPCR	quantitative polymerase chain reaction
	RAB8A	ras-related protein rab-8A (also known as MEL)
	RARB	retinoic acid receptor beta
	RASSF1	ras association domain family 1
	RB-1	retinoblastoma associated protein 1
30	REL	c-rel proto-oncogene protein
	ROC	receiver operating characteristic
	ROS1	proto-oncogene tyrosine-protein kinase ROS precursor
	SKI	Ski oncogene
	TEP1	telomerase protein component 1
35	TG	topographic genotyping
	TGF	tissue growth factor
	TGF α	tissue growth factor alpha

	TGFβ	tissue growth factor beta
	TNF	tumor necrosis factor
	TP53	cellular tumor antigen p53
	Tsc2	tuberin
5	TSG11	tumor suppressor gene on chromosome 11
	VEGF	vascular endothelial growth factor
	VHL	Von Hippel-Lindau disease tumor suppressor
	WT-1	Wilm's tumor protein
	YES1	proto-oncogene tyrosine-protein kinase YES

10

1.2 Definitions

It must be noted that as used herein, the singular forms "a", "an", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a patient" includes a plurality of patients and so forth.

15 By "patient" and "subject" are meant to include any mammal including, but not limited to, bovines, primates, equines, porcines, caprines, ovines, felines, canines, and any rodent (*e.g.*, rats, mice, hamsters, and guinea pigs). Preferred animals include agricultural animals, domesticated animals, and primates, especially humans.

20 By "anomaly" is meant a broad, encompassing term to indicate and disease related change in a lymphatic or hematopoietic cell or tissue of an organ. Thus, "anomaly" includes cancer or dysplasia, a pre-cancerous neoplastic state, or a non-neoplastic condition. Pre-cancerous states include proliferative lesions that can span a spectrum from low-grade to high-grade neoplasia.

25 By "non-neoplastic condition" and "non-neoplastic abnormality" are meant to indicate specimens from sites known not to contain neoplasia. The non-neoplastic condition may be inflammatory or any adaptive state that may include features of cell proliferation but needs to be clearly discriminated from neoplasia.

30 By "breast anomaly" is meant a broad, encompassing term to indicate a disease related change in the breast. Thus, "breast anomaly" includes a breast cancer or dysplasia, a pre-cancerous breast neoplastic state, or a breast non-neoplastic condition. The term "breast cancer" includes, but is not limited to, invasive ductal adenocarcinoma, invasive lobular carcinoma, invasive neuroendocrine cell carcinoma, sarcoma of the breast, metastatic cancer of the breast, breast cancer recurrence, or other forms of breast cancer. The term "pre-cancerous breast state" includes, but is not limited to, typical or atypical ductal hyperplasia, 35 typical or atypical lobular hyperplasia, breast papilloma, or other precancerous lesions of the breast. The term "non-neoplastic condition" includes, but is not limited to, fibrocystic disease, reactive proliferations of the breast, mastitis, soft tissue trauma or injury, and other

non-neoplastic states. Breast-related conditions are also discussed herein and are encompassed by the term "breast anomaly".

By "biological sample" is meant to include a frozen section, a biopsy of fresh tissue, fixative-treated tissue, stained tissue, and fixative and stained tissue or cells. By "biological fluid sample" is meant to include, but is not limited to, breast lavage samples, ascites fluid samples, fine needle aspirates (FNAs) from a cyst or other region of the subject's body, urine, blood, cerebrospinal fluid, a liquid cytology sample obtained by any medically available method, and/or saliva. The sample can contain cells or may contain only free-floating DNA (non-nuclear DNA) in the fluid sample.

2. Cancer and Conditions and Genetic Defects Relating to the Cancers and Conditions

2.1 Cancers and Other Conditions

The cancers include those linked to familial cancers and cancers of indeterminate origin. Also contemplated herein are the diagnosis, prognosis, and treatment determination for purposes of cancer treatment (*i.e.*, advanced versus early stage cancer), precancerous conditions (*e.g.*, dysplasia), and non-cancerous conditions (*e.g.*, hyperplasia). These are listed in Table 1 below as well as discussed herein. This list is not totally comprehensive and can be expanded to include additional old and newly recognized disease states. This speaks to the broad applicability of this application to advance specimen analysis.

The malignancies and precancerous conditions that can be diagnosed using the materials and methods described herein include leukemias and lymphomas, and cancers and precancerous conditions that can be readily screened by analysis of a biological fluid sample from a patient.

The cancers include but are not limited to cancers of the head and neck (*e.g.*, nasal cavity, paranasal sinuses, nasopharynx, oral cavity, oropharynx, larynx, hypopharynx, salivary glands, and paragangliomas), lung tumors (*e.g.*, non-small cell and small cell lung tumors), neoplasms of the mediastinum, brain cancers, cancers of the gastrointestinal tract (*e.g.*, colon, esophageal carcinoma, pancreatic carcinoma, gastric carcinoma, hepatobiliary cancers, cancers of the small intestine, cancer of the rectum, and cancer of the anal region), genitourinary cancers (*e.g.*, kidney cancer, bladder cancer, prostate cancer, cancers of the urethra and penis, and cancer of the testis), gynecologic cancers (*e.g.*, cancers of the cervix, vagina, vulva, uterine body, ovaries, fallopian tube carcinoma, peritoneal carcinoma, and gestational trophoblastic diseases), breast cancer, cancer of the endocrine system (*e.g.*, thyroid tumors, parathyroid tumors, adrenal tumors, pancreatic endocrine tumors, carcinoid tumors, carcinoid syndrome, and multiple endocrine neoplasias), sarcomas of the soft tissues and bone, benign and malignant mesothelioma, skin cancers, liver cancers, malignant

melanoma (*e.g.*, cutaneous melanoma and intraocular melanoma), neoplasms of the central nervous system, pediatric tumors (*e.g.*, neurofibromatosis, neuroblastoma, rhabdomyosarcoma, Ewing's sarcoma and peripheral neuroectodermal tumors, germ cell tumors, primary hepatic tumors, and malignant gonadal and extragonadal germ cell tumors),
 5 paraneoplastic syndromes, and solid cancers with unknown primary sites. By cancer or neoplasm is also meant to include metastatic disease and recurrence or relapse of a cancer(s). Also contemplated are virally induced neoplasms such as adenovirus-, HIV-1-, or human papilloma virus-induced neoplasms (*e.g.*, cervical cancer, Kaposi's sarcoma, and primary CNS lymphoma) and any secondary cancer appearing in lymph.

10 Lymphomas include, but are not limited, to Hodgkin's lymphoma, non-Hodgkin's lymphoma (*e.g.*, B-cell non-Hodgkin's lymphoma and T-cell non-Hodgkin's lymphoma), cutaneous T-cell lymphomas (CTCL), lymphoplasmacytoid lymphoma (LPL), mantle cell lymphoma (MCL), follicular lymphoma (FL), diffuse large cell lymphoma (DLCL), Burkitt's lymphoma (BL), Lennert's lymphoma (lymphoepithelioid lymphoma), Sezary
 15 syndrome, anaplastic large cell lymphoma (ALCL), and primary central nervous system lymphomas.

Leukemias include but are not limited to chronic myelogenous leukemia (CML), acute myelogenous leukemia (AML), acute promyelocytic leukemia, acute lymphoblastic leukemia (ALL), prolymphocytic leukemia, hairy cell leukemia, T-cell chronic lymphocytic
 20 leukemia, plasma cell neoplasms, chronic lymphocytic leukemia (CLL), and myelodysplastic syndromes (*e.g.*, chronic myelomonocytic leukemia).

25

TABLE 1: TISSUE/ORGAN ANOMALIES

Cancer	Pre-Cancerous Condition	Benign Tumors	Non-Neoplastic Condition
Carcinoma and other epithelial malignancies	Epithelial dysplasia of varying degree. Possess significant mutational damage.	Uncontrolled growth but without metastatic potential. Possess small amounts of mutations.	Simulators of neoplasia that must be rigorously distinguished. Lack detectable mutational damage.
Sarcoma and other mesenchymal malignancies	Mesenchymal dysplasia of varying degree. Possess significant mutational damage.	Uncontrolled growth but without metastatic potential. Possess small amounts of mutations.	Simulators of neoplasia that must be rigorously distinguished. Lack detectable mutational

			damage.
Lymphohematopoietic cancers	Premalignant cellular proliferations and dysplastic syndromes. Possess significant mutational damage.	Uncontrolled growth but without metastatic potential. Possess small amounts of mutations.	Simulators of neoplasia that must be rigorously distinguished. Lack detectable mutational damage.
Neuroepithelial cancer, glioma and other central nervous system cancers	Premalignant cellular proliferations and dysplastic syndromes. Possess significant mutational damage.	Uncontrolled growth but without metastatic potential. Possess small amounts of mutations.	Simulators of neoplasia that must be rigorously distinguished. Lack detectable mutational damage.

The term cancer can also include aneuploid and diploid cancers, familial and hereditary cancers, virus-induced cancers, chemotherapeutic/radiation-induced cancers, cancers caused by environmental factors, sporadic cancers, and other types indicated herein.

- 5 Also contemplated is a metastatic cancer, which includes but is not limited to a cancer in any organ presenting as a metastasis but with no apparent primary tumor.

Benign lesions are generally characterized as proliferative or non-proliferative in nature. Non-proliferative lesions are generally not associated with an increased risk of cancer. Proliferative lesions without atypia generally result in a small increase in risk.

- 10 Atypical hyperplasia is associated with a greater risk of cancer development (*i.e.*, relative risk of about 4 to about 5). Premalignant conditions include, but are not limited to, premalignant organ cancer.

- Also contemplated herein is the analysis of cancer recurrence. Cancer recurrence includes, but is not limited to, local recurrence after surgery, recurrence after combined
 15 surgery and radiation therapy, recurrence after combination treatment of chemotherapy, radiation, surgery, bone marrow transplant, and/or other treatment modalities and combinations thereof.

2.2 Genetic Defects

- Familial or hereditary cancers are those inherited from one generation to the next.
 20 The first-identified, hereditary cancer linked gene is the tumor-suppressor gene, *p53*, in Li-Fraumeni syndrome. Mutations of *p53* had previously been described in the context of progression of sporadic (non-hereditary) cancers.

- Following the hypothesis that a point mutation might be inherited in one allele of a putative locus, and that loss of heterozygosity or other genetic lesion might occur in the
 25 other allele of that locus later in life, led researchers to the eventual localization and characterization of *BRCA1* in human chromosome 17q21. Mutations in *BRCA1* have been

verified in familial breast and ovarian cancer patients, but surprisingly mutations have not commonly been detected in sporadic breast cancers. Robert B. Dickson et al., "Cancer of the Breast," IN CANCER – PRINCIPLES & PRACTICE OF ONCOLOGY, 1541-1556 (5th ed., DeVita et al., ed., 1997).

5 Thereafter, the gene *BRCA2* was identified on chromosome 13q13, which is associated with familial female and male breast cancer, but not ovarian cancer. Familial breast cancer genes include but are not limited to *p53*, *BRCA1*, *BRCA2*, and in some instances E-cadherin. Other inherited cancer genes include *APC*, *MCC*, *WT-1*, *RB-1*, *VHL*, *MEN1*, *NF1*, *NF2* as well as others, all of which would function equivalently in the system
10 described here.

 In addition to familial/hereditary genes, other genes have been established to be involved in breast cancer development and progression, including "cancer progression related genes" and other "tumor suppressor genes". Breast cancer and other forms of human malignancy include progression genes, but are not limited to ErbB2, Erb2/Erb3, myc, cyclin
15 D1, RB-1, and p53. The human epidermal growth factor receptor (HER) family of receptor tyrosine kinases includes HER1 (EGFR/ErbB1), HER2 (ErbB2/neu), HER3 (ErbB3), and HER4 (ErbB4); over-expression of HER 1-4 have been observed in breast cancer. Tovey et al., 2004 *Breast Can. Res.* 6(3): R246-R251. Genetic mutations in tumor suppressor genes and tumor progression genes all relate to acquisition and progression of cancer. Tumor
20 suppressor genes such as RB-1 and p53 have been established. However, additional tumor suppressor genes have been indicated to be localized on chromosomes 7q, 9p, 11p, 11q, 17q, and 18q. This is only a very small subset of the total pool of known cancer related genes. Thus, the pool of tumor suppressor genes may be expected to expand in time with new gene discovery. All such genes are amenable to the analysis described herein.

25 Another cytogenetic alteration in cancer involves gene amplification. The initial step is thought to be the formation of extrachromosomal, self-replicating units, termed double-minute chromosomes. These genetic elements later become permanently incorporated into chromosomal regions and are termed homogeneous staining regions. An amplified genetic unit (amplicon) is typically much larger than the wild-type gene.

30 Other genes that appear to be involved in cancer development and progression include "survival factors." Survival factors include, but are not limited to, estrogen, epidermal growth factor (EGF), angiogenesis factors, tissue growth factor alpha (TGF α), fibroblast growth factor-2 (FGF-2), insulin, extracellular matrix, Bcl-2, and Bcl_{xL}. Genes and the proteins encoded thereby involved in cell death (*i.e.*, death factors) also are
35 commonly involved with cancer development and progression. Cancer death factors include but are not limited to tissue growth factor beta (TGF β), tumor necrosis factor (TNF), myc, p53, Cyclin A/CDK2, Bcl_{xs}, Bax, hypoxia-inducible factor-1 α (HIF-1 α), and ICE protease.

Cell cycle regulators may also be aberrantly expressed in cancers. Cell cycle regulators include but are not limited to estrogen, progesterone, growth factors (*e.g.*, FGF, VEGF, TGF β , pleiotropin, and HGF), myc, p53, extracellular matrix, and cell-cell junction proteins (*e.g.*, LCAM). These proteins in turn may interact with a cell cycle kinases (*e.g.*,
 5 CDK2, CDK4, CDK5), a cyclin (*e.g.*, D family, E family, A family, or B/A family).

Evidence is accumulating which shows that the development of cancer is a complex process that includes the dysregulation of proliferative factors, activation of oncogenes, dysregulation of inhibiting factors, and loss of suppressor gene function. The malignant progression of cancer is characterized by a progressive deterioration of the normal
 10 mechanisms of cell cycle progression and tissue compartmentalization until the highly abnormal and dysregulated state of metastatic disease is reached. The methods and materials herein for analyzing fluid samples provide a flexible approach that enables different types of mutational damage to be detected in minute specimens. The information from fluid samples can be used alone, or in combination with pathology data obtained by standard microscopy,
 15 FISH analysis, histopathology, and so forth.

Other genetic factors include sex steroid receptors (*e.g.*, estrogen and progesterone) and ploidy (DNA histogram). Less commonly assessed factors can include, but are not limited to, proliferative indices (*e.g.*, Ki-67, PCNA/cyclins, thymidylate synthetase, MIB1), topoisomerase II, histone H3, transforming growth factors (alpha and beta), epidermal
 20 growth factors, insulin-like growth factors and their associated binding proteins, oncogene products (*e.g.*, *Her2/neu* or *c-erbB2*, *int2*, *c-myc*, *ras*, *RB*, *Bcl2*, *CCND1*, *EMS1*), p53 protein, invasion related markers (*e.g.*, cathepsin D, laminin receptor, stomelysin-3, uPA/PA-1), angiogenesis factors, pS2, NM23, heat-shock proteins, and MDR1.

Genetic research on cancer, including those molecular changes that may be
 25 important during the earlier phases of development and progression, has documented a wide range of tumor suppressor genes, oncogenes, and other molecular changes, some of which are summarized in Tables 2 and 3. Other genes and molecular alterations are also discussed however the list can be extended with equivalent applicability.

TABLE 2: REPRESENTATIVE GENETIC MUTATIONS IN CANCER

Tumor Suppressor Genes	Oncogenes	Other
TP53	ErbB2, Erb2/Erb3	DNA methylation alterations
Retinoblastoma, RB-1	Int2/FGF3	Homozygous deletion of p16
BRCA1	c-myc	Microsatellite instability
BRCA2	Mdm2	Translocations
CDKN2A	Jun	Gene rearrangements
PTEN	Fos	-----

As exemplified in Table 3, detection of cumulative mutations represents an important aspect of the molecular pathology approach for early detection of cancer. It is recognized however that the total detectable mutation complement in an individual patient may in fact be relatively few in total number, which is in keeping with an early stage of cancer development. At the same time, related conditions, such as pre-cancer states and non-neoplastic states, will themselves possess few or no detectable mutations respectively. Thus, the need to incorporate DNA quantity and quality analysis is addressed by the embodiments disclosed herein.

TABLE 3: CANCER ASSOCIATED GENETIC ALTERATIONS ASSOCIATED WITH STAGE OF PROGRESSION

Genetic alterations associated with early primary cancer development	Genetic alterations associated with later progression of primary cancer	Genetic alterations associated with metastatic spread of cancer
BRCA1, BRCA2	TP53	NME1
HER2/neu		

The methods, materials, and kits described herein provide examination of biological fluid, such as aspirates. Fluid samples often do not result in accurate diagnosis when examined solely by microscopic evaluation of cytology slides. Another option is to look for mutations in the free-floating DNA by using surrogate markers for cellular neoplastic transformation. The free-floating DNA would have originated from the nuclear DNA of cells lining the area from which the fluid was collected or which was in direct contact with the fluid. It is these cells that produced the fluid that provide the basis for diagnosis and characterization of the fluid sample and its associated condition.

At this point, it is useful to distinguish clearly between the different types of DNA that will apply to this situation of laboratory evaluation. The DNA of interest is DNA of the cells that are responsible for the anomaly. It is this DNA that represents the anomaly of interest, including early cancer. The fluid will contain the DNA released from these critical cells of interest. However, the amount and intactness of the DNA will depend upon the active turnover of cells, as described above. The DNA in the fluid can be referred to as free-floating DNA. Due to the physicochemical properties of DNA, part or all of the DNA may not be free-floating, but instead may become attached to either the side of the container used to collect the fluid, or the DNA may attach non-specifically to the surface of any cells present in the fluid. Thus, when the analysis is performed on the fluid, the DNA being evaluated may be a combination of nuclear DNA in detached cells present in the fluid, as well as free-floating DNA from the lining cells and secondarily attached to detached cells.

The DNA will be in greater amounts and have better integrity (quality) if the underlying cause of the fluid accumulation arises from a malignancy.

Representative cells lining malignant cysts appear to manifest a high level of accumulated mutational damage reflective of the underlying malignancy. Thus, high levels of accumulated mutational damage would not be seen in cells in a benign cyst. Similarly, inflammatory processes would not be expected to manifest evidence of cumulative mutational change. A higher degree of cell turnover may also occur in the lining of a malignant fluid collection, thus releasing high amounts and quality of mutated DNA into the fluid.

3. Methods

The methods described herein are useful in the prognosis, treatment, and/or diagnosis of anomalies. The methods comprise obtaining a tissue sample or fluid sample from an organ or tissue, extracting DNA from the sample, quantifying the DNA, analyzing the quality of the DNA, and performing mutational and other molecular analyses on the DNA all for the purpose of providing a comprehensive characterization of neoplastic change if present. The assessment can be performed of fluid samples alone or in combination with pathologic analysis of tissue biopsies, resected tissue, frozen sections, blood work, and/or pathologic analysis of any other biological sample.

3.1 Methods of Obtaining Aspirate

Fluid specimen collection may be obtained by medically available methods. Two more common methods utilize imaging via ultrasound or by ultrasound, abdominal CT, or magnetic resonance imaging (MRI). However, any technique to collect fluid with or without associated imaging can be used.

Preferably, microscopic slides associated with the fluid specimen, if present, are first reviewed, as well as the clinical information pertinent to the individual patient. However, fluid genotyping can be performed without preliminary review of microscopic features. The fluid sample can be examined by the naked eye to note its color, clarity, and character. For example, if the fluid contains a abundant blood, this date should be taken into account, due to implications related to the blood. Blood has been shown to contain inhibitors of DNA amplification and thus may lead to an inability to effectively genotyping a fluid sample that contains abundant blood. The presence of the blood could affect the results in biological fluid samples for assessment of cancers that are not hematopoietic in origin, and thus should be accounted for when reviewing the final results of the disclosed methods. A good way to evaluate for excessive blood in the fluid sample is to first view the fluid under the microscope, where the excessive blood will be easily appreciated. Thus, the analysis of the results and the methods are preferably performed by one of skill in the art.

These preliminary observations are merely suggestions, and do not necessarily need to be performed as a prelude to topographic genotyping of fluid, as used herein.

Breast cyst/duct aspiration may be obtained by medically available methods. Two more common methods utilize imaging via ultrasound or by computed tomography (CT) or magnetic resonance imaging (MRI). However, any technique to collect breast fluid DNA with or without associated imaging can be used.

Preferably, the histopathology and cytopathology associated with the cyst/duct fluid is first reviewed with the clinical information pertinent to the individual patient. However, fluid genotyping can be performed without preliminary review of microscopic features. The fluid sample can be examined by the naked eye to note its color, clarity, and character. For example, if the fluid contains abundant blood, this data should be taken into account, due to implications related to the blood. Blood has been shown to contain inhibitors of DNA amplification and thus may lead to an inability to effectively genotype the fluid sample that contains abundant blood (Guo et al., 2005 *Bio-technol. Lett.* 27(1): 3-6). As the presence of the blood could affect the results, it should be accounted for when reviewing the final results of the methods described herein. These preliminary observations are optional, and do not necessarily need to be performed as a prelude to topographic genotyping of fluid samples, as used herein.

3.2 Methods of Obtaining a Brushing

Cell-containing brushing specimens may be obtained by medically available methods. Two more common methods utilize direct visualization under endoscopy conditions to perform, monitor and optimize the brushing; the second method is to perform the brushing with the assistance of imaging either by radiological techniques including ultrasound, abdominal CT, or magnetic resonance imaging (MRI).

Preferably, microscopic slides associated with the brushing specimen, if present, are first reviewed, as well as the clinical information pertinent to the individual patient. However, brushing genotyping can be performed without preliminary review of microscopic features. The fluid sample can be examined by the naked eye to note its color, clarity, and character. For example, if the brushing contains abundant blood, this data should be taken into account, due to implications related to the blood. Blood has been shown to contain inhibitors of DNA amplification and thus may lead to an inability to effectively genotyping a fluid sample that contains abundant blood. The presence of the blood could affect the results in biological brushing samples for assessment of cancers that are not hematopoietic in origin, and thus blood should be accounted for when reviewing the final results of the disclosed methods. A good way to evaluate for excessive blood in the brushing sample is to first view the fluid under the microscope, where the excessive blood will be easily appreciated. Thus, the analysis of the results and the methods are preferably performed by

one of skill in the art. These preliminary observations are merely suggestions, and do not necessarily need to be performed as a prelude to topographic genotyping™ of fluid, as used herein.

5 Breast cyst/duct aspiration may be obtained by medically available methods. Two more common methods utilize imaging via ultrasound or by computed tomography (CT) or magnetic resonance imaging (MRI). However, any technique to collect breast fluid DNA with or without associated imaging can be used.

10 Preferably, the histopathology and cytopathology associated with the cyst/duct fluid is first reviewed with the clinical information pertinent to the individual patient. However, fluid genotyping can be performed without preliminary review of microscopic features. The fluid sample can be examined by the naked eye to note its color, clarity, and character. For example, if the fluid contains abundant blood, this data should be taken into account, due to implications related to the blood. Blood has been shown to contain inhibitors of DNA amplification and thus may lead to an inability to effectively genotype the fluid sample that
15 contains abundant blood (Guo et al., 2005 *Biotechnol. Lett.* 27(1): 3-6). As the presence of the blood could affect the results, it should be accounted for when reviewing the final results of the methods described herein. These preliminary observations are optional, and do not necessarily need to be performed as a prelude to topographic genotyping of fluid samples, as used herein.

20 Liquid cytology specimens may be obtained by a variety of ways that culminate in a sample containing representative cellular material mixed with liquid cytology fixative (Dawson, 2004 *Cancer* 102(6): 340-4). The original specimen source may be of any type and from any part of the body. It need not be limited to fluid specimens but can include solid tissue aspirations that are then placed in suitable supporting fixative fluid. The fluid
25 character of the specimens derived from the admixture of solid cells fragments is mixed with suitable cytology fixative fluid. Having reached the point where the specimen resides in fluid, the procedure outline below would be an example of how liquid cytology specimens could be handled.

3.3 Separation of Cells From Cell Free Fluid in Liquid Specimens

30 An initial step in the analysis of a fluid specimen that contains cellular material is to separate the specimen into a cell-free fluid and a concentrated collection of the cellular material. This is not required if one is willing to analyze a fluid specimen within which intact cells are present. If the fluid specimen is paucicellular, the inclusion of sparse numbers of cells may not greatly influence the analysis of free DNA in the fluid. However,
35 if one desires a more rigorous individual analysis of the cell-free fluid and the cellular constituents, then a suitable separation technique needs to be performed on the biological

fluid sample. This commonly involves either a centrifugation or filter separation system although other separations techniques are not excluded.

The cellular collection may be then analyzed in the form in which it exists or it may be smeared or otherwise deposited on glass slides for microscopic analysis. Then, based on the insights derived from microscopic analysis, specific cells of interest may be microdissected for a more focused molecular analysis. This will allow the independent molecular analysis of cytology cells and cell-free fluid, each of which will provide valuable discriminating information. The analysis of microdissected cellular cytology smears is discussed below.

10 3.3.1 Methods of Extracting DNA from the Fluid

The molecular analysis begins by extracting DNA from the biological fluid. An intermediate step of centrifugation can be performed for the purpose of concentrating the cellular material from a larger specimen volume, when a liquid specimen is felt to contain only limited amounts of cells. The degree of dilution is noted so that specimens of different cellularity can be compared to each other.

The next step is accomplished by first treating the sample with proteinase (e.g., proteinase K or similar proteases or combination of proteases), and then capturing the DNA on, for example, a spin column while eluting the protein, salts, and other unwanted constituents from the sample. Preferably, DNA is extracted from a fluid by column separation, though other equivalent method(s) can be substituted, such as ethanol precipitation and alternative techniques. The latter may include but is not limited to filter separation, adherence of DNA to glass beads, hybridization based separation, immunologic based separation. The extraction of DNA can be readily accomplished using, for example, a Qiagen spin column. Techniques that extract DNA from a biological fluid or tissue samples would also suffice. If sufficient free DNA is present in the fluid sample, it may be used directly, thereby bypassing the step of DNA extraction. The extracted DNA need not be highly purified. For example, the DNA may be used directly following extraction methods, as have been described in the literature. Highly purified DNA or DNA prepared by another means can also be used.

Molecular analysis begins by extracting DNA from the fluid or liquid cytology sample. The sample is treated with a proteinase (e.g., proteinase K, pronase, subtilisin, or other proteinases, or combinations of proteinases), and then capturing the DNA on, for example, a spin column while eluting the protein, salts, and other unwanted constituents. Preferably, DNA is extracted from the fluid or liquid cytology sample by column separation, though other equivalent methods can be substituted. This can be readily accomplished using, for example, a Qiagen spin column. Techniques that extract DNA from a biological fluid, liquid cytology, or tissue samples would also suffice. If sufficient DNA is present in

the fluid sample, it may be used directly, thereby bypassing the step of DNA extraction. The extracted DNA need not be highly purified, but may be used directly following extraction methods (*see e.g.*, Rolston et al., 2001 *J. Mol. Diagn.* 3(4): 129-32; Marsh et al., 2003 *Liver Transpl.* 9(7): 664-71). Highly purified DNA or DNA prepared by another means can also be used.

Patients can have a combination of these tissue samples taken and examined, the data from which can then be collated to form a diagnosis, prognosis, and treatment plan for each individual patient. Ideally, the patient would have the benefit of data from all three types of tissue sampling and analysis. The three methods of analyzing these types of tissue samples, in fact compliment each other for validation purposes.

In particular, a preferred analytical platform has been designed to correlate with and improve histopathologic and cytologic observation, complementing and improving existing pathology practice. The materials, methods, and kits disclosed permit the objective diagnosis of invasive cancer, low- and high-grade forms of dysplasia, and neoplastic and non-neoplastic conditions. The methodology comprises detailed molecular analysis incorporating DNA quality and quantity, point mutational analysis of oncogenes and other genes and other forms of mutational damage correlated with cancer, broad spectrum of tumor suppressor genes, and other breast cancer associated genes linked to microsatellite loss of heterozygosity (LOH), and a new approaches for accurate copy number determination of genes and oncogenes (copy number analysis). Another aspect provides methods to clearly discriminate cancer from non-cancer states, such as those produced by inflammation, infection, and trauma to a particular organ or tissue. Methods for diagnosing, determining prognosis of, and defining a course of treatment for cancer, high-grade dysplasia, pre-cancerous states, and non-neoplastic conditions are also provided.

Another embodiment relates to the quality control assessment of small amounts of DNA that, due to paucity of diagnostic material, may produce both false positive and false negative results for mutation detection (C.R. Miller, et al., 2002 *Genetics* 160(1): 357-66; J. C. Dreesen, et al., 1996 *J. Assist. Reprod. Genet.* 13(2): 112-4). Standard pathology practice addresses this need by extracting purified DNA from specimens and measuring its content by optical density (OD) determination. Unfortunately, extracted DNA from fixative-treated, stained specimens is vanishingly low, *e.g.*, below about 1 ng for fixative-treated, paraffin-embedded specimens measuring about 1 cm in size or less. Although quantitative polymerase chain reaction (qPCR) has been employed on fixative-treated tissue, qPCR assumes high quality starting nucleic acid content, which cannot be assumed when dealing with fixative-treated, stained specimens of limited size. Methods and devices which address the need for analytic validation and quality control are provided herein.

Non-neoplastic fluid collections have very low amounts of DNA present in the fluid, because the lining cells contain very few cells that replicate slowly. Low-grade, slow-growing, neoplastic anomalies would generate an associated fluid that contained relatively greater amounts of DNA reflecting the greater turnover of replicating lining cells. Finally, a malignant tumor is believed to possess greater amounts of DNA of high quality and intactness in fluid localized near the tumor than benign tumors and precancerous lesions.

The cellular specimens are placed into a convenient volume of lysis buffer in order to create a lysate. Preferably, the lysis buffer consists of NP-40 diluted to a final concentration of about 1.0%. NP-40 is the preferred non-ionic detergent, however other non-ionic detergents can also be utilized. For example, Tween (*e.g.*, Tween-20 or Tween-40), Triton X-100, and Nikkol may be substituted for NP-40. Detergents such as sodium dodecylsulfate (SDS) must be avoided, because its presence in the lysate will interfere with nucleic acid amplification.

One aspect of the methods described herein is not to overload the crude lysate with cellular material. Excessive cellular material will prevent effective nucleic acid amplification. Since microdissected samples can vary greatly in the amount of cellular material present, and it is not convenient to measure the amount of cellular material that is removed, the best means to gauge the desired quantity of cellular material is to identify a lysate that is mildly turbid. Samples that have heavy turbidity are not optimal.

Another aspect of the method is to examine the histopathologic appearance of the sample being placed into the lysis buffer. If microscopic examination reveals a relatively high content of acellular material, such as collagen, then more cellular material can be used to prepare the crude lysate. If excessive cellular material has been added to the lysis buffer such that it is too turbid, additional lysis buffer can be added to achieve the desired degree of turbidity.

The lysate is then treated with a proteinase (*e.g.*, proteinase K, pronase, subtilisin, thermolysin, or papain, or combinations thereof). Preferably, the proteinase is proteinase K. Proteinase K (triturachium alkaline proteinase) is then added to the lysate to achieve a final concentration of about 2 mg/mL. One or more proteinases are added to break down histone and other proteins that may be in contact with the DNA, thereby preventing DNA amplification. Although different concentrations of proteinase K can be used, both greater and less than about 2 mg/mL, a preferred amount is about 2.0 mg/mL. The lysate is incubated at 37°C for about two hours to overnight. Other temperatures and durations of incubation may be substituted with comparable effects. At the completion of the proteinase digestion, the specimen is heated to 100°C for 5 minutes to inactivate any residual proteinase enzyme. The lysate may then be stored indefinitely at about -22°C or lower until

needed for the step of nucleic acid amplification. When nucleic acid amplification is to be performed on the sample, the sample is then defrosted.

The lysate (defrosted or fresh) is centrifuged 10,000 revolutions per minute in a table-top centrifuge to pellet the undigested material at the bottom of the tube. One
5 microliter aliquots of the crude lysate are removed from the lysate immediately above the pellet. The lysate in this region is preferred for performing direct PCR amplification, although other portions of the lysate from the sample may also be utilized. The 1- μ L samples should not however include the pellet itself, which contains tissue that will interfere with PCR amplification.

10 One microliter of the crude lysate is added to a 12.5 μ L of nucleic acid amplification reaction buffer. Oligonucleotide primers are added as desired. A stock mixture buffer is then used to add *Taq* polymerase, deoxyribonucleotides, and salts to the final amplification mix. In order to enhance amplification of the minute quantities of DNA, the magnesium concentration is significantly increased to a level of about 8 mM magnesium
15 chloride. Magnesium chloride for the addition of the magnesium ions is added in order to improve hybridization of the probe to the template DNA. Magnesium chloride or other equivalent reagent is added to obtain a final concentration of greater than 5.0 mM to 15 mM and more preferably from about 8 mM to about 10 mM. Manganese and alternative salt ions of equivalent valency may also be utilized (e.g., MnCl_2 , MgCl_2).

20 Another aspect of the amplification mix that enhances the amplifiability of minute quantities of DNA is an amplification mix having a final concentration of sucrose of about 12 g per 100 mL, with a range of sucrose being from about 10 g per 100 mL to about 25 g/100 mL. The nature of how the sucrose concentration enhances amplification of the nucleic acid is not fully understood, but may occur by assisting the relaxation of double
25 strands of DNA and in turn encouraging primer annealing.

Thus, for greatest amplification of DNA from a minute biological sample, the sample is placed directly into a NP-40 containing lysis buffer, so as to form a lysate for direct nucleic acid amplification. Then the lysate is treated with one or more proteinases for a certain period of time. The proteinase(s) is deactivated by boiling, after which the sample
30 centrifuged. The fluid immediately above the pellet is withdrawn in about 1 μ L aliquots for individual amplification reactions. The final amplification mixture contains significantly higher than normal MgCl_2 and sucrose to both facilitate double strand relaxation and primer hybridization.

The materials and methods above can be used with an analytical platform. The
35 analytical platform can include:

(1) techniques to separate any cells present in the fluid from the cell free fluid for separate molecular pathology analysis;

(2) techniques to measure the quantity of DNA serving as means to infer the proliferative rate of neoplastic states;

(3) techniques to measure the quality of DNA including the extent of its breakdown into progressively shorter pieces of nucleic acid serving as means to infer the proliferative rate of neoplastic states;

(4) mutation analysis of minute specimens affording determination of genomic and specific gene copy number imbalance;

(5) mutation analysis of minute specimens affording determination of point mutational change of nucleic acids in the sample;

(6) mutation analysis of minute specimens affording determination of specific gene amplification and/or homozygous gene loss;

(7) mutation analysis of minute specimens affording determination of microsatellite instability;

(8) mutation analysis of minute specimens affording determination of gene rearrangement;

(9) mutation analysis of minute specimens affording determination of genomic deletional expansion;

(10) mutation analysis of minute specimens affording determination of altered DNA methylation status uniquely innovated to operate on fixative-treated tissue sections and cytology smears;

(11) techniques to determine the time course of mutation accumulation providing a dynamic understanding of molecular change; and

(12) quality control measures to monitor the reliability and representativeness of mutation detection consisting of recommendations for replicate analysis at both the specimen level (sample replication) and/or replication of the individual marker results (*i.e.*, marker replication).

3.3.2 Method of Extracting DNA from Microdissected Cytology Cells

In order to gain the most advantage from correlative microscopic analysis, and in order to complement cytology practice, the cytology cells are stained using the Pap smear or other suitable stains and examined microscopically. Then the most representative cell and cell clusters are microdissected off the glass slide and collected together in a small volume (*i.e.*, about 25 μ L) for molecular analysis. Rather than attempt highly inefficient DNA extraction, the microdissected cytology cells are gathered together in a crude lysate in about 1% NP-40 for detergent breakdown of cell walls. Other nonionic detergents can also be used (*e.g.*, Tween, Triton-X, Nikkol, and the like) in amounts ranging from about 0.5% to about 5%, and every 0.1% value in between that range. To further the availability of DNA for nucleic acid amplification, the crude lysate is treated with a proteinase at a concentration

of about 2 mg/mL. A preferred proteinase is proteinase K in an amount of about 0.5 mg/mL to about 5 mg/mL and every 0.1 mg value in between that range. Similar amounts of other proteinases can be used alone and in cocktails (e.g., subtilisin, thermolysin, pronase, papain, other proteinases, and combinations thereof). The quality of DNA can then be assessed as described below. The extracted DNA from micro dissected cytology cells can also be enhanced using the methods discussed in Section 3.3.1.

3.4 Method of Quantifying DNA and DNA Quality Analysis

After purifying the DNA, optical density (OD) analysis is performed to quantify the DNA. One approach can be the Nanodrop technique, because it requires only one microliter to be sacrificed for the purpose of obtaining the DNA concentration. (Ding et al., 2004 *J. Biochem. & Mol. Biol.* 37(1): 1-10). Other techniques for quantifying DNA can also be utilized.

The higher the OD value indicates that a larger amount of DNA is present. The quantity of DNA extracted can vary. However, the higher the amount of DNA, the more likely a high-grade dysplasia or a malignancy is present. The exact level of the DNA measured in this step, and the significance to the biological process, is not universal but must be interpreted in the overall context of findings. It has been shown that samples with an OD value of about 2.0 ng/ μ L or less are considered to have insufficient DNA. Values in the range of about 2.0-10.0 ng/ μ L (and any 0.1 ng/ μ L value in between) indicate pre-cancerous lesions and/or low-grade indolent forms of neoplasia. Values over about 10.0 are generally indicative of the presence of a malignancy. For example, a value of about 359 was observed with a sample from a pancreatic cyst, which indicated that the tumor was growing very actively within the cyst or ductal space. Thus, ranges of between 10.0 to 500 ng/ μ L or to 1,000 ng/ μ L are contemplated as representative of a malignancy. However, the upper limit could be even higher than 500 or 1,000. These conclusions are not definitive, but only suggestive and are to be used with the other data obtained by the other steps of the methods described herein to come to a definitive assessment.

The DNA can be quantified by OD measurement at wavelengths of 230, 260, and 280 nm. The 260/280 and 260/230 ratios should be 1.7-2.0, in keeping with extraction of purified DNA and for the purpose to exclude protein and other contaminants. Other suitable settings can also be substituted. Any technique that quantifies the DNA in the sample can be a suitable substitute to OD measurement or as discussed herein.

3.5 Determining DNA Quality for DNA Quantity and Quality Analysis

A variety of techniques are used to quantify DNA quality. These fall into two categories, quantitative polymerase chain reaction (qPCR) and competitive template polymerase chain reaction (CT-PCR). Each technique provides valuable information on DNA quality. Both techniques used together on specimens provide highly valuable

information in this regard. The qPCR reaction can be performed using, for example, sybr green as the indicator in a suitable thermocycler capable of measuring fluorescence during the amplification process, as this is the simplest and least costly technique. Other techniques for qPCR determination using fluorescent labeled primers can also be substituted. Known quantitative controls and replicate analysis of samples can be used to standardize amplification reactions and is preferred. The use and configuration of controls and replicate analysis may be varied as determined by the user. For example, standardization of qPCR amplification of the first exon of the K-ras-2 gene may be used for this purpose. However, any PCR product from any gene or genomic segment may be used.

10 The DNA in the aspirate or biological fluid sample may be free-floating or free and adherent to the surface of cells or tissue constituents of the cyst. DNA possesses a physicochemical tendency to adhere to biological surfaces such as cell membranes and physical structures, such as glass or plastic. From these locations the DNA can be extracted and analyzed. This DNA is derived from cells that line or are in contact with the region from which fluid is collected (*e.g.*, aspirate or lavaged samples). The DNA present would then be representative of those cellular elements that constitute the fluid source. The free or surface-attached DNA is not visible by microscopic examination. However, it can be extracted and analyzed as a means to assess the etiology and character of a fluid collection. The same concepts hold true for fluid moving through a channel, such as the breast ductal or pancreaticobiliary ductal system. The sample will generally contain both intact cells with internal nuclear DNA and free genomic DNA (non-nuclear DNA). Both internal nuclear DNA and a free-genomic DNA can be used to determine the characteristics of the lining cells. In addition to searching for mutations, the quantity and amplifiable quality of the DNA can serve as end points for the analysis of DNA.

25 Two such gene/pseudogene pairings are exemplified, although the approach can be applied to any other pairings of genomic segments differing only by a deletional region or regions. These genes include the glucocerebrosidase gene (GenBank D13286) and its pseudogene (GenBank D13287), and the human carboxyl ester lipase (*CEL*) gene (GenBank M94579) and its pseudogene (GenBank M94580). Other similar examples can also be employed. In a preferred embodiment, the glucocerebrosidase gene and its pseudogene is sufficient, as it affords specific deletional regions of 19 bases and 55 bases in the exon 1 and 9 respectively.

35 The competitive template PCR (CT-PCR) reaction for the glucocerebrosidase gene/pseudogene pairing (exon 1 and 9) provide a novel and sensitive means to quantitative the degree of DNA degradation. The results from this assay, together with the data from the OD and qPCR steps, provide discriminating information on DNA degradation. With the current techniques used in the art, the conditions of Table 1, for example, are only poorly

discriminated in fine needle biopsy cytology specimens using microscopic evaluation. This information in turn can be used to distinguish the conditions listed in Table 1, or any breast conditions discussed herein.

- 5 Genetic research on breast cancer, including those molecular changes that may be important during the earlier phases of development and progression, has documented a wide range of tumor suppressor genes, oncogenes, and other molecular changes, some of which are summarized in Table 4A and 4B.

TABLE 4A: GENETIC MUTATIONS IN BREAST CANCER

Tumor Suppressor Genes	Oncogenes	Other
TP53	ErbB2, Erb2/Erb3	DNA methylation alterations
Retinoblastoma, RB-1	Int2/FGF3	Homozygous deletion of p16
BRCA1	EGFR	Microsatellite instability
BRCA2	Mdm2	Translocations and rearrangements
VHL	FES	
RASSF1	MAFK	
DPC4	SKI	
DLC1	YES1	
FHIT	ETS1, ETS2	
Tsc2	RAB8A	
PARK2	REL	
ITS	ABL1	
TSF11	LOC428352	
Wt1-B	CRK	
TEP1	ROS1	
MTUS1	MAF	
LOC464014	Jun	
LOC464012	Fos	
LOC464013	Other tumor oncogenes	
JST		
LOC472922		
Mtus1		
RARB		
CDKN2A		
PTEN		

10 **TABLE 4B: BREAST CANCER ASSOCIATED GENETIC ALTERATIONS ASSOCIATED WITH STAGE OF PROGRESSION**

Genetic alternations associated with primary breast cancer formation	Genetic alternations associated with primary breast cancer progression	Genetic alternations associated with metastatic spread of breast cancer
BRCA1, BRCA2	TP53	NME1
HER2/neu		

As exemplified in Tables 4A and B, detection of cumulative mutations represents an important aspect of the molecular pathology approach for early detection of breast cancer.

However, it is recognized that the total detectable mutation complement in an individual patient and cancer may in fact be relatively few in total number, in keeping with the early stage of cancer development. At the same time, breast cancer related conditions, such as breast pre-cancer states, and non-neoplastic states themselves possess few or no detectable mutations. Thus, the need to incorporate DNA quantity and quality analysis is addressed by the embodiments disclosed herein.

3.6 Quantitative PCR Amplification

The first step in the qPCR process is to adjust the extracted sample DNA concentration to a value of about 5 ng/ μ L so that the absolute amount of DNA present in each reaction is the same. This will still allow the integrity to vary, which is the purpose of the analysis. About 5 ng/ μ L is preferred, because it has been found to be a minimal value for robust amplification. However, other amounts may be used (*e.g.*, about 10 ng/ μ L to 100 ng/ μ L and any integer value in between; more preferably about 1.0 ng/ μ L to about 10 ng/ μ L and any 0.1 ng/ μ L value in between). All other values may be substituted quite freely and is up to the discretion of the investigator.

The number of qPCR cycles may be used as a marker of DNA quality. The lower the number of cycles required to reach a desired threshold is indicative of higher quality DNA. In general, if over 30 qPCR cycles are required, then the DNA quality is considered suboptimal due to, for example, allelic imbalance resulting from inadequate amounts of template DNA. Specifically, Ct values (*i.e.*, threshold values for quantitative PCR product detection) over 30 cycles is considered evidence of poor quality of DNA, especially if the DNA quantity present is above about 2.0 ng/ μ L. Ct values of about 29-30 are considered borderline and representative of poor quality DNA. Cycle values of 29 or less are indicative of good quality DNA and are generally characteristic of the breast anomaly being malignant. Additionally, values of 29 or less may indicate the presence of neoplasia. The lower the Ct value, the more likely neoplastic cell proliferation is malignant. These values may vary based on the conditions of the assay and amounts of DNA employed. Cancers that proliferate slowly, are well differentiated in growth pattern, and are relatively less cellular may be expected to show borderline Ct values that are significantly higher than 29-30 cycles.

3.7 Competitive Template Nucleic Acid Amplification to Assess DNA Quality

DNA quality may be further assessed by performing competitive template PCR amplification for a unique pair of genes (*e.g.*, glucocerebrosidase gene and its pseudogene), at a particular point where the two genes have virtually identical sequences. In the case of GCS and its pseudogene, there is a 55 base pair deletion in the pseudogene. A second region in exon 1 of the glucocerebrosidase gene provides an equivalent opportunity to

utilize a 19 base pair deletion in the same manner. This is not the only gene that can be used in this fashion. In fact, any pairing of gene or genomic segments of similar sequence, but different in length, can be substituted.

5 The PCR reaction creates two amplicons that are identical in sequence except for the deletional region. During the reaction, a competition exists between the two similar templates (but having different lengths). The degree of DNA degradation in the sample will be reflected by less effective amplification of the longer template as compared to the shorter template. This serves as a measure of DNA integrity. The amount of each product, short and long, may be quantitatively measured by capillary electrophoresis. Methods of
10 performing the PCR reaction and electrophoresis are well known in the art. A non-neoplastic process shows prominent DNA degradation, while a malignancy is associated with the presence of abundant, good quality DNA. Reagents are added to the final sample to enhance DNA availability and to enhance the ability to amplify the DNA.

While the use of two highly similar amplification targets is recommended, any
15 system that utilizes similar primers to amplify products of different lengths can be substituted. The genetic targets used should be selected to preferentially produce different sized products under similar amplification conditions in the biological fluid samples near the target tissue specimens, and thus are essentially independent of specimen fixation or preparation effects. The closer to the value one (1.00) the ratio of amplified product to each
20 product is, the better the quality of DNA (*i.e.*, the more likely the tissue from which it was derived is malignant or hyperplastic).

The procedure for PCR amplification has been well described and different variations can be utilized. As described in the Examples below, the recommended procedures of the manufacturers for the PCR reagents are followed (*e.g.*, GeneAmp kit,
25 Applied Biosystems). However, other commercial and non-commercial systems for PCR amplification readily can be substituted. It is preferred that the PCR reaction is performed in a manner that is highly robust. Robust, in this context, indicates the reliable generation of abundant amplified DNA that accurately reflects the starting composition mixture of normal and mutated DNA derived from a particular specimen, especially when using minute
30 samples such as dilute fluid specimens. Reagents such as dimethylsulfoxide or dextran sulfate may be added to the amplification reaction to enhance amplification. Other similar reagents can be substituted. Also, manipulations, such as nested PCR, may be performed to further enhance amplification. Other similar steps may be used, though they are not mandatory.

35 Based on the concept that a longer-sized PCR product is present in lesser amounts than a shorter-sized product due to greater chance for strand breakage as a result of DNA degradation, a competitive duplex PCR reaction of highly similar DNA sequence, but

differing in length was needed. This is accomplished by simply carrying out a short- and long-product PCR reaction in one container (*e.g.*, test tube) on one source of DNA. Unfortunately, such duplex reactions are not equivalent, because they use different primers and generate radically different products. Moreover, the status of the PCR reaction must be carefully controlled, because amplification is not the same during the different phases of the reaction (*i.e.*, exponential phase versus plateau phase). The methods described herein overcome these obstacles.

Another aspect contemplates the use of competitive template PCR (CT-PCR). This technique enhances quality determination of DNA and analytic validation. Replicate aliquots are used as substrate to amplify a segment of DNA present in the glucocerebrosidase gene (GCS) that is identical in base sequence to that in its pseudogene (or other gene/pseudogene pairings), except for a 55 base pair deletion present in the latter. A second deletion of 19 bases present in the GCS gene, but absent in the pseudogene, is also available for similar application. The result is two amplicons that theoretically should be generated in equal amounts. Preferential amplification of a shorter allele from among a mixture of templates is a common observation when amplifiable DNA is of low quality or otherwise rate limiting. The relative fluorescent intensities of the true GCS/pseudo GCS is thus a measure of effective ability to amplify the starting DNA, in turn a measure of DNA integrity. This determination is coupled with absolute determination of fluorescent content serving as a measure of total amplicon production. The closer to a value of one (1.0) for ratio of true GCS gene/pseudo GCS gene, the higher the absolute fluorescence, the more reproducible the replicate analysis, and the greater the confidence that minute starting quantities of template DNA are being accurately represented in the final analysis. Analysis in triplicate is performed to determine consistency in PCR amplification for the templates present in a particular sample. This is vital for the effective molecular pathology analysis of paucicellular fluid samples. These techniques can be combined with analysis of microdissected tissue samples and cytology samples.

If the microdissected cellular sample is found to have excellent quality starting DNA, as described by the criteria above, then mutational analysis may proceed using threshold for significant allelic imbalance that are at the 95% confidence level for the range of variation in normal tissue specimens for each unique pairing of polymorphic markers. However, if there is evidence of limiting DNA, as reflected by CT-PCR, greater allowance must be given to imbalances that favor the preferential amplification of the shorter microsatellite allele. When CT-PCR indicates a relatively mild impairment, then mild imbalances favoring the shorter allele may be ignored. When greater impairment is indicated by CT-PCR, the specimen can be regarded as possessing low quantities of relatively poor quality DNA. As a further check, specific imbalances for genomic loci can

be repeated in duplicate, and the same criteria applied as that for true GCS/pseudo GCS. These insights are essential for effective analysis of clinical specimens of the type used to detect early breast cancer.

- 5 The CT-PCR reaction is performed in replicate using increasing amounts of starting DNA. When performed in this manner the results will inform the user concerning the total amount of amplifiable DNA in the test sample and the extent of degradation of DNA into smaller sized fragments.

**TABLE 5: CT-PCR QUANTITATIVE CHARACTERIZATION
OF DNA QUANTITY AND QUALITY**

10

	LOWEST CONCENTRATION (PEAK HEIGHT RATIO)	MIDDLE CONCENTRATION (PEAK HEIGHT RATIO)	HIGHEST CONCENTRATION (PEAK HEIGHT RATIO)
Adequate DNA Concentration Good DNA Quality	0.8-1.0	0.8-1.0	0.8-1.0
Low DNA Concentration Good DNA Quality	0.4-0.8	0.8-1.0	0.8-1.0
Very Low DNA Concentration Good DNA Quality	PCR FAILURE	ALLELIC DROPOUT	0.4-0.8
Adequate DNA Concentration Poor DNA Quality	0.4-0.8	0.4-0.8	0.4-0.8
Adequate DNA Concentration Very Poor DNA Quality	PCR FAILURE	ALLELIC DROPOUT	0.4-0.8
Low DNA Concentration Poor DNA Quality	ALLELIC DROPOUT	0.2-0.6	0.4-0.8

- 15 The presence of adequate amounts of good quality DNA is reflected by ratios of allelic height peaks in the about 0.8 to about 1.0 range. This indicates that the longer length allele is essentially amplified at an equivalent rate to that of the shorter allele. When the DNA content is suboptimal for balanced PCR amplification, there will be a tendency to delete the longer allele, and thus the ratio will tend to fall. As a higher concentration of starting DNA is used, one will see an improvement in the ratio into the range of about 0.8-1.0 that is associated with adequate amounts of good quality DNA. With very low quantities of DNA, "allelic dropout" and PCR failures will be seen with wide fluctuations in

the amplifiability of individual alleles, as exemplified in Table 5.

When DNA is degraded, there will be a range of DNA fragments from short to long in the specimen. Very short fragments may not be adequate for PCR amplification, while slightly longer fragments will tend to favor amplification of the short over the longer sized allele. Thus a stable lower peak height ratio value will be seen for all concentrations (Table 5). The combination of low amounts of poor quality DNA will show the expected effects as exemplified in Table 5. As shown, the presence of very low amounts and/or very poor quality DNA will result in failures and allelic imbalance that will not be possible to specifically trace to a particular cause.

10 3.8 **Mutation analysis and other molecular analysis of DNA in aspirate for Allelic Imbalance (loss of heterozygosity)**

Because chromosomal allelic loss, commonly referred to as loss of heterozygosity (LOH), is a major cause of tumor suppressor gene inactivation, detection of LOH from microsatellite markers closely linked to key tumor suppressor genes serves as an excellent surrogate marker for gene inactivation. The materials and methods herein use a panel of LOH markers on specimens, together with analysis of DNA quality and quantity to predict the presence of cancer or neoplasia, and to diagnose and treat these conditions. Other LOH markers can be substituted or used in combination; the materials and methods are not limited to the described LOH markers.

20 PCR amplification may be used to generate amplicons of less than 200 nucleotides using synthetic oligonucleotide primers flanking each microsatellite (amplicons can range in size from about 40 to less than 200 and every integer value in between). Allele peak heights and lengths may be used to define the presence or absence of an allelic imbalance (*i.e.*, LOH) for a given sample. Allelic imbalance is reported when the ratio of polymorphic allelic bands for a particular marker is beyond about 95% confidence limits for the variation in peak heights for individual allele pairings derived from an analysis using non-neoplastic specimen samples. In general, the value of the ratio is below 0.5 or above 2.0. Preferably, the allele ratio is two standard deviations beyond the average for the ratio of the specific pairing of polymorphic alleles. This will provide the lowest threshold for detection of significant allelic imbalance (LOH). However, other algorithms for defining LOH can be used, so long as they are applied uniformly across different specimens such as using the allele ratio of the non-neoplastic sample as a denominator that is divided into the allele ratio for the lesional samples. It is understood that minor amounts of LOH will not be detected. However, these minor LOH mutations may not be causally related to clonal expansion or provide significant malignant growth properties.

Allelic imbalance mutations are treated as genomic deletions associated with tumor suppressor genes. The ratio of allele peak heights is a measure of an admixture of mutated

and non-mutated cells (or mutated and non-mutated DNA), and varies according to the individual pairing of specific, microsatellite marker alleles. Allele ratios of 2.0 or 0.5 is said to be present when 50% of the total DNA is derived from cells possessing the loss. The deviation from the ideal normal ratio of about 1.0 indicates which specific allele is affected.

5 Allele ratios below about 0.5, or above about 2.0, are mathematically correlated with the proportion of cells affected by genomic loss. The order of mutation acquisition may be arranged in a temporal sequence of mutation accumulation reflecting the proportion of cells affected by specific microsatellite marker loss. Markers displaying more extreme ratios are considered to have been acquired earlier in the disease process. This conclusion is based on

10 the premise of clonal expansion. A "clonal expansion" occurs when tumor cell populations progressively replace each other by accruing mutations, which are causally associated with increasing malignant phenotype.

Parallel progression of cellular neoplasia and DNA mutational damage in, for example, breast cancer has not been described at the molecular level. The materials,

15 methods, and kits described herein provide examination of fluid samples from breast tissue and other tissues. Breast cyst aspirates are usually not available for accurate diagnosis. Another option is to look for mutations in the free-floating DNA that exists in the breast cyst aspirates by using surrogate markers of cellular neoplastic transformation. The free-floating DNA would have originated from the nuclear DNA of cells lining the cyst or in direct

20 contact with the cyst fluid. It is these cells that produced the cyst that provide the basis for diagnosis and characterization of the breast or other tissue cyst and associated condition.

At this point, it is useful to distinguish clearly between the different types of DNA that will apply to this type of laboratory evaluation. The DNA of interest is the DNA of the cells that are responsible for the disease or breast anomaly. It is this DNA that represents

25 the anomaly of interest. The cyst fluid will contain the DNA released from these cells of interest. However, the amount and intactness of the DNA will depend upon the active turnover of cells as described above. The DNA in the cyst fluid can be referred to as free-floating DNA. Due to the physicochemical properties of DNA, part or all of the DNA may not be free-floating, but instead may become attached to either the side of the container used

30 to collect the cyst fluid, or the DNA may attach non-specifically to the surface of any cells present in the cyst fluid. Thus, when the analysis is performed on the cyst fluid, the DNA being evaluated may be a combination of nuclear DNA in detached cells present in the cyst fluid, as well as free-floating DNA from the lining cells and secondarily attached to detached cells. The DNA will be in greater amounts and have better integrity (quality) if the

35 underlying cause of the cyst is malignant.

Representative cells lining malignant cysts appear to manifest a high level of accumulated mutational damage reflective of the underlying malignancy. Thus, high levels

of accumulated mutational damage would not be seen in cells of a benign cyst. Similarly, an inflammatory process in the breast would not be expected to manifest evidence of cumulative mutational change. A higher degree of cell turnover may also occur in the lining of a malignant cyst, thus releasing high amounts and quality of mutated DNA into the cyst fluid.

Because chromosomal allelic loss, commonly referred to as loss of heterozygosity (LOH), is a major cause of tumor suppressor gene inactivation, detection of LOH from microsatellite markers closely linked to key tumor suppressor genes serves as an excellent surrogate marker for gene inactivation. The present materials, methods, and kits use a panel of LOH markers on breast specimens and other specimens, together with analysis of DNA quality and quantity to predict the presence of breast cancer or neoplasia, and to diagnose and treat these conditions. Breast related LOH markers are provided in Table 6.

TABLE 6

Chromosome	Cancer Type	Markers	Associated Function	Associated LOH	Cite
18p11.3	DCIS	D18S59, D18S481, D18S452	Tumor progression	Prior allelic deletion 3p, 9p, 11p, 17p, and 17q	Kittiniyom et al., 2001 <i>Breast Cancer Res.</i> 3(3): 192-198

Copy number analysis is used to search for and quantitatively characterize homozygous deletion and/or specific gene amplification. The simplest approach to accomplishing this objective is to perform a duplex PCR reaction for a gene of interest coupled with a comparator gene that is assumed not to vary with respect to cancer progression. This approach is not effective, because the two individual nucleic acid amplification reactions operate independently. More importantly, the independent nature of the reactions reside mainly in the operating characteristics of the amplification primers with respect to hybridization and denaturation. The Human Genome Database (Internet: <URL:http://www.gdb.org>) was searched for primer sequences that were highly similar, especially with respect to the 3' end of the primers for the gene of interest and the comparator gene. Primer sequences for duplex PCR amplification were identified, in which the 3' end of the primers were highly similar. For example, the amplification products of both reactions would be very close in size so that the amplification reactions would operate in a nearly equivalent fashion. This is exemplified with the *CDKN2A* gene, which can undergo homozygous deletion in many types of human cancer (Liggett et al., 1998 *J. Clin. Oncol.* 16(3): 1197-206). The comparator gene containing a sequence of near identity to *CDKN2A* was the *APP* gene situated on chromosome 22 (Table 7). The sequences in bold represent the primers used. Obviously the reverse primer would be an antisense sequence to

that displayed in Table 7 (e.g., for CDKN2A the reverse primer would be 5'-CATTGCCAGGTATAAGAGCAGACTC-3').

TABLE 7

CDKN2A	
5' -TTTGAGAAATGGAGTCCGTCCTTCCAATGACTCCCTCCCCATTTTCCTATCTGCCTACA	
GGCAGAATTC TCCCCGTCCTATTAAATAAACCTCATCTTTTCAGAGTCTGCTCTTATAACC	
AGGCAATGTACACGTCTGAG-3'	
APP	
5' -AACTATAGTCCATTATCAAAACCAGGAAACTGACCTTGTGTGCCCACTGTC TTTTATGTCC	
CTAGTTTTCCTAATGACTAATGATTTTCATTTAAGATCTCTCATTTCGTGTTCTC TTGGAAAGAG	
AATCAAATACCAGAGTCTCTGTGCTGAGCTGCCAAAATGCCATAACTGTTTAACTTACTTCACAAA	
GGTACATAGATGACGATT-3'	

5 Note that the 10 bases at the 3' end of the upstream primers for each gene are identical in sequence. Similarly there is identity of the seven bases at the 3' end of the downstream primer. This homology together with the design of each primer to have similar melting temperatures ensures that the duplex reaction will proceed in parallel relationship to each other. The ratio of peaks for each product using fluorescent labeled primers on

10 capillary electrophoresis provides a quantitative relationship with respect to the relative content. Using normal tissue, cytology cells, and fluid as sources of non-neoplastic DNA and sample variation, it is possible to define thresholds for significant gene loss or gain in a pathologic sample. This approach is ideally suited for use with minute specimen samples. It opens the way for copy number analysis of individual genes in clinical samples of the type

15 obtained in routine management of patients.

While mutation detection is regarded as a means to characterize cancer, assessment of the quality and quantity of DNA has never been employed to achieve this objective. This assessment of DNA is not intuitive. In pathology, the emphasis had always been on detection of mutations, which is suitable for detecting advanced forms of cancer that bear

20 abundant mutational damage. Many studies have documented the striking difference with respect to the presence or absence of mutations when comparing advanced cancer to a non-cancerous process of the same tissue type. This however would not be of value in the early detection of cancer where accumulated mutations would be relatively few. An alternative end point is required to objectively separate early cancer from indolent forms of neoplasia

25 (benign tumors) and from similar appearing non-neoplastic states.

One embodiment provides for a method is provided for predicting the presence of a breast anomaly in a patient with breast cysts comprising:

(a) performing molecular analysis of DNA from the aspirate from a patient breast cyst comprising:

30 (i) performing an optical density analysis of the aspirate to determine DNA quantity;

- (ii) performing a quantitative PCR analysis of the aspirate to determine DNA quality;
- (iii) performing competitive template PCR to further quantify DNA quality; and
- 5 (b) performing mutation analysis of the DNA in the aspirate comprising:
- (i) determining the presence of mutations in a tumor suppressor gene and/or the presence of a cancer related genetic marker;
- (ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to
- 10 tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy;
- (iii) determining point mutations in a cancer-associated gene;
- (iv) determining copy number imbalances of specific genes, such as
- 15 homozygous deletion and oncogene amplification;
- (v) determining other structural alterations in DNA;
- (vi) determining the percentage of mutated DNA from steps (b)(ii) and (b)(iii);
- (vii) determining a temporal sequence of mutation accumulation based on
- 20 step (v) for the patient; and/or
- (viii) detecting the presence of focal neoplastic progression in a heterogeneous cellular lesion with lower and higher grade forms of neoplasia in association with fluid accumulation, thereby predicting the presence of a breast anomaly.
- 25 The DNA in the aspirate may be free-floating, or free and adherent to the surface of cells or tissue constituents of the cyst. DNA possesses a physicochemical tendency to adhere to biological surfaces such as cell membranes and physical structures, such as glass or plastic. From these locations the DNA can be extracted and analyzed. This DNA is derived from cells that line or are in contact with the region from which fluid is collected
- 30 (e.g., aspirate or lavaged samples). The DNA present would then be representative of those cellular elements that constitute the cyst. The free or surface attached DNA is not visible by microscopic examination. However, it can be extracted and analyzed as a means to assess the etiology and character of, for example, a breast cystic alteration. The same concepts hold true for fluid moving through a channel, such as the breast ductal system. The sample
- 35 will generally contain both intact cells with internal nuclear DNA and free genomic DNA (non-nuclear DNA). Both internal nuclear DNA and a free genomic DNA can be used to determine the characteristics of the lining cells. In addition to searching for mutations, the

quantity and amplifiable quality of the DNA can serve as end points for the analysis of DNA. Preferably, the cycles of quantitative PCR performed in step (a) (ii) is greater than a threshold unique for that specific type of DNA. Preferably, the cycles of quantitative PCR performed in step (a)(ii) is less than or equal to 29, although it may also be 29-30 or thirty or more in pre-cancerous or non-malignant samples.

The OD is about 3.0 to about 10.0, with an OD of greater than 10.0 in malignant tissues. Preferably, the allele ratio is two standard deviations beyond the average for the ratio of the specific pairing of polymorphic alleles.

Use of two markers within each locus was used to increase the likelihood that at least one of the markers would be polymorphic within a subject, and thus informative for LOH analysis or copy number alteration.

PCR amplification may be used to generate amplicons of less than 200 nucleotides using synthetic oligonucleotide primers flanking each microsatellite. Allele peak heights and lengths may be used to define the presence or absence of an allelic imbalance (*i.e.*, LOH) for a given sample. Allelic imbalance is reported when the ratio of polymorphic allelic bands for a particular marker is beyond about 95% confidence limits for the variation in peak heights for individual allele pairings derived from an analysis using non-neoplastic specimen samples. In general, the value of the ratio is below 0.5 or above 2.0. Preferably, the allele ratio is two standard deviations beyond the average for the ratio of the specific pairing of polymorphic alleles. This will provide the lowest threshold for detection of significant allelic imbalance. However, other algorithms for defining LOH can be used, so long as they are applied uniformly across different specimens, such as using the allele ratio of the non-neoplastic sample as a denominator that is divided into the allele ratio for the lesional samples. It is understood that minor degrees of LOH will not be detected. However, the inability to detect is not a drawback, because these minor LOH mutations may not be causally related to clonal expansion or provide significant malignant growth properties.

Allelic imbalance mutations are treated as genomic deletions associated with tumor suppressor genes. The ratio of allele peak heights is a measure of an admixture of mutated and non-mutated cells (or mutated and non-mutated DNA), and varies according to the individual pairing of specific microsatellite marker alleles. Allele ratios of 2.0 or 0.5 is said to be present when 50% of the total DNA is derived from cells possessing the loss. The deviation from the ideal normal ratio of about 1.0 indicates which specific allele is affected. Allele ratios below about 0.5 or above about 2.0 are mathematically correlated with the proportion of cells affected by genomic loss. The order of mutation acquisition may be arranged in a temporal sequence of mutation accumulation reflecting the proportion of cells affected by specific microsatellite marker loss (*see* FIGS. 3A, 3B, and 4). Markers

displaying more extreme ratios are considered to have been acquired earlier in the disease process. This conclusion is based on the premise of clonal expansion; a "clonal expansion" occurs when tumor cell populations progressively replace each other by accruing mutations which are causally associated with increasing malignant phenotype.

5 Determination of the temporal sequence of mutation accumulation had never before been effectively performed on fixative-treated clinical specimens or in a clinical context. Delineation of the time course of mutation acquisition is provided herein and is important to the diagnosis, prognosis, and treatment of the patient's disease. The identical profile of mutations between different individual patients with the same microscopic form of cancer
10 may not be expected to behave in exactly the same manner. The order by which the mutations are accumulated may greatly influence the final biological behavior of the breast abnormality. This is especially true for response to treatment where earlier acquisition of treatment responsive mutations will be associated with greater therapeutic responsiveness in the treated patient due to the presence of those mutations. Table 8 below shows the
15 detection and characterization of microsatellite instability in microdissected tissue samples. MSI stands for "microsatellite instability", which is a manifestation of DNA repair enzyme deficiency. DNA repair enzyme deficiency is usually caused by point mutational damage to one of several specific DNA repairs. The most common of these genes is MLH1. A mutation of MLH1 accounts for 95% of all DNA repair gene mutations. When the DNA
20 repair function fails, microsatellites throughout the genome show instability when the cells undergo replication, becoming either short or longer in length. Demonstration of microsatellites that shift the size is an indirect, but sensitive, measure of the presence of DNA repair enzyme deficiency. By "+++" is meant that there is a great deal of microsatellite size shifting, whereas only a single "+" indicates mild shifting. Mild shifting
25 is indicative of mild DNA repair enzyme deficiency. As cancers evolve, the degree of microsatellite instability worsens, whereas early in colon polyp development, the microsatellite instability will be mild. Using the advances described herein, the extent of microsatellite instability is capable of being characterized. "LGD" stands for "low grade dysplasia". With the materials and methods described herein, LGD is incapable of being
30 accurately diagnosed, if at all.

TABLE 8

PHS01-26895	BAT 25	BAT 26	D2S123	D5S346	D17S250
Adenomatous polyp (LGD); Area 1	MSI+	NA	Normal	Normal	NA
Adenomatous polyp (LGD) Area 2	MSI++	NA	Normal	Normal	NA
Colon adenocarcinoma Area 1	MSI+++	MSI+++	MSI+++	MSI+++	NA
Colon Adenocarcinoma Area 2	MSI+++	MSI+++	MSI+++	MSI+++	NA

In this 29-year old patient suspected of cancer susceptibility, two separate areas from a tubulovillous adenoma and a co-existing colonic adenocarcinoma were microdissected and tested using the recommended panel of markers. Note the presence of microsatellite
5 instability that is more extensive in the colorectal cancer. By applying topographic genotyping (TG) in this manner to fluid and cytology samples, diagnosing physicians will have new abilities at determining patients having cancer susceptibility, thereby facilitating early cancer detection and impacting cancer prevention. Knowing the existence of cancer susceptibility can greatly affect the manner in which a patient is treated and subsequently
10 followed for further cancer formation. For example, it is shown that gliomas, which acquire the 1p/19q deletion earlier in tumorigenesis, are more responsive to treatment than gliomas that do not have deletions early in tumorigenesis.

Assuming a model of allelic loss with minimal non-neoplastic cell DNA inclusion, the percentage of mutated DNA may be determined for each marker. When two or more
15 mutations, (e.g., K-ras-2 and/or allelic imbalance mutations), are detected, their time course of accumulation is inferred by the proportion of total DNA manifesting the alteration. In the case of K-ras-2 oncogene point mutation, involvement of 100% of cells is considered to be present when the intensity of the mutated base on sequencing is equal to or greater than the normal sequence base pair. This may be determined qualitatively by visual comparison or
20 more precisely by quantitation. Alternatively, any method that provides relative amounts of mutated cell population can be used.

During the final evaluation of all the data from the methods, the presence of mutational change must first be analyzed with respect to the quantity and quality of DNA. It is important to do this analysis of mutational change, because the presence of low
25 amounts of poor quality DNA can produce false positive detection of deletion type mutations due to a phenomenon in the PCR reaction called allelic drop-out.

These analytical methods may be used to diagnose breast cancer, pre-cancerous breast states or non-neoplastic conditions of the breast as well as any precancerous to cancer condition from another organ from which a fluid sample can be taken. The analytical
30 methods may also be used to determine the prognosis for the patient. The analytical methods may be further used to determine a course of therapy or combination of therapeutic modalities for the patient. Because the methods allow for sensitive prediction of cancer and pre-cancerous conditions, a physician or other of skill in the art can use the methods to more accurately determine the prognosis of the breast condition. Furthermore, if a physician has
35 advanced warning that cancer is present, treatment may be begun earlier and more carefully tailored to the condition (including surgery, radiation therapy, chemotherapy, and biologic therapy). To this end, the physician skilled in the art will be familiar with diagnosis,

prognosis and treatment of cancers, precancerous conditions, and non-neoplastic conditions. See, for example, DeVita, Jr. et al., CANCER: PRINCIPLES & PRACTICE ON ONCOLOGY, volume 2 (5th Ed., New York, 1997). Additionally, the methods described herein can be used to evaluate animal models and their responsiveness to treatment modalities.

5 **3.8.1 Nucleic Acid Amplification**

Nucleic acid amplification can be carried out, for example, as follows. One microliter aliquots can be removed for PCR amplification of individual polymorphic microsatellite markers. Other sized aliquots of DNA can also serve equally well. Nucleic acid amplification can be carried out according to manufacturer's instructions (*e.g.*,
10 GeneAmp kit, Applied Biosystems, Foster City, CA). Other variations on the PCR reaction can apply equally. Fluorescent labeled oligonucleotide primers are employed for quantitative determination of allelic imbalance based on the peak height ratio of polymorphic microsatellite alleles. Other quantitative system for PCR and/or qualitative approach can be substituted.

15 **3.8.2 Allelic Imbalance Determination**

Allelic imbalance determination can be carried out as follows. Post-amplification products are electrophoresed and relative fluorescence determined for individual allele peak height (GeneScan ABI3100, Applied Biosystems, Foster City, CA). The ratio of peaks are calculated by dividing the value for the shorter-sized allele by that of the longer-sized allele.
20 Thresholds for significant allelic imbalance have been determined beforehand in extensive studies using normal (*i.e.*, non-neoplastic) specimens representing each unique pairing of individual alleles for every marker used in the panel. Peak height ratios falling outside of two standard deviations beyond the mean for each polymorphic allele pairing were assessed as showing significant allelic imbalance. In each case, the non-neoplastic tissue targets are
25 used to establish informativeness status and then to determine the individual pattern of polymorphic marker alleles. Having established significant allelic imbalance, it is then possible to calculate the proportion of cellular DNA that was subject to hemizygous loss. For example, a polymorphic marker pairing whose peak height ratio was ideally 1.00 in normal tissue with a standard deviation in non-neoplastic tissue of 0.23, could be inferred to
30 have 50% of its cellular content affected by hemizygous loss if the peak height ratio was 0.5 or 2.0. This requires that a minimum of 50% of the DNA in a given sample be derived from cells possessing deletion of the specific microsatellite marker. The deviation from ideal normal ratio of 1.0 indicated which specific allele was affected. In a similar fashion allele ratios below 0.5 or above 2.0 could be mathematically correlated with the proportion of cells
35 affected by genomic loss. Other algorithms for quantitative determination of allelic imbalance can be used with equal effectiveness.

Examples of the normalizing data that was obtained through the analysis of normal microdissected samples is shown in the following four tables (Tables 9-12). For each unique pair of alleles, the average value for the ratio of peak heights is obtained. The variation within each set of samples is used to calculate the standard deviation. 95% confidence limits are then determined. Having established that a particular sample is significantly imbalanced, the use of the average value for specific allele pairings can then be used to determine the percentage of mutated cells accounting for the degree of imbalance.

The data in Tables 9-12 was obtained in as follows. In order to define minimum threshold for the detection of significant allelic imbalance indicative of mutational change, the range of variation in the ratio of peak heights for all combinations of specific alleles in the general population using normal microdissected tissue specimens was determined. A minimum of four different normal tissue specimen from four different human subjects were required to compute average and standard deviation values. The more common allele combinations were represented by much larger number of subjects. The threshold for minimal absolute values for allelic imbalance due to mutational change was then defined as the average plus or minus 2 standard deviations. These are recorded to be used as absolute values to detect minimal mutation induced allelic imbalance. Finally, the values at minimum threshold levels were also converted into percentages of mutated, microdissected cells using the average value for each allele pairing as a normalizer.

TABLE 9

MARKER	A1	A2	R	#	%[HET]	AVE	SD	LOW	%	HIGH	%	ADJ
D10S.520-HEX	14	37		409	68.7%	1.4132	0.3321	0.7490	47.0%	2.0773	32.0%	2.6664
D10S.520-HEX	122			1	0.2%	NA	NA					
D10S.520-HEX	138	146	0.9982	1	0.2%	0.9982	NA					
D10S.520-HEX	142	166	1.8760	1	0.2%	1.8760	NA					
D10S.520-HEX	146			1	0.2%	NA	NA					
D10S.520-HEX	146	154	1.2612	1	0.2%	1.2612	NA					
D10S.520-HEX	146	158	2.0377	4	1.0%	1.5380	0.3891	0.7598	50.6%	2.3163	33.6%	3.1136
D10S.520-HEX	146	162	1.8129	4	1.0%	1.3996	0.2855	0.8285	40.8%	1.9707	29.0%	2.3643
D10S.520-HEX	146	170	1.8487	1	0.2%	1.8487	NA					
D10S.520-HEX	150			5	1.2%	NA	NA					
D10S.520-HEX	150	158	1.3339	2	0.5%	1.2242	NA					
D10S.520-HEX	150	162	1.7292	4	1.0%	1.5817						
D10S.520-HEX	150	166	2.0864	1	0.2%	2.0864	NA					
D10S.520-HEX	154			13	3.2%	NA	NA					
D10S.520-HEX	154	158	1.9518	21	5.1%	1.4277	0.3233	0.7811	45.3%	2.0742	31.2%	2.6094
D10S.520-	154	162	1.9600	17	4.2%	1.5430	0.2792	0.9846	36.2%	2.1014	26.6%	2.4181

MARKER	A1	A2	R	#	%[HET]	AVE	SD	LOW	%	HIGH	%	ADJ
HEX												
D10S.520-HEX	154	166	2.0904	25	6.1%	1.5287	0.3473	0.8341	45.4%	2.2233	31.2%	2.8016
D10S.520-HEX	154	170	2.0263	5	1.2%	1.5645	0.4027	0.7591	51.5%	2.3700	34.0%	3.2246
D10S.520-HEX	154	174	2.0864	3	0.7%	1.4277	NA					
D10S.520-HEX	154	182	1.7136	1	0.2%	1.7136	NA					
D10S.520-HEX	158			22	5.4%	NA	NA					
D10S.520-HEX	158	162	2.0949	40	9.8%	1.4706	0.3990	0.6725	54.3%	2.2686	35.2%	3.2156
D10S.520-HEX	158	166	2.0873	27	6.6%	1.3138	0.3527	0.6084	53.7%	2.0193	34.9%	2.8373
D10S.520-HEX	158	170	1.5401	7	1.7%	1.3221	0.1874	0.9473	28.3%	1.6968	22.1%	1.8451
D10S.520-HEX	158	174	1.8733	5	1.2%	1.4109	0.3386	0.7336	48.0%	2.0882	32.4%	2.7134
D10S.520-HEX	162			40	9.8%	NA	NA					
D10S.520-HEX	162	166	1.9126	45	11.0%	1.3375	0.3055	0.7266	45.7%	1.9485	31.4%	2.4622
D10S.520-HEX	162	170	2.0019	18	4.4%	1.4223	0.2975	0.8273	41.8%	2.0173	29.5%	2.4452
D10S.520-HEX	162	174	1.8269	8	2.0%	1.4841	0.3441	0.7959	46.4%	2.1722	31.7%	2.7672
D10S.520-HEX	166			34	8.3%	NA	NA					
D10S.520-HEX	166	170	1.9447	23	5.6%	1.3002	0.2455	0.8093	37.8%	1.7911	27.4%	2.0889
D10S.520-HEX	166	174	1.6172	9	2.2%	1.2466	0.1875	0.8717	30.1%	1.6216	23.1%	1.7828
D10S.520-HEX	166	178	1.5868	1	0.2%	1.5868	NA					
D10S.520-HEX	166	186	1.8876	1	0.2%	1.8876	NA					
D10S.520-HEX	170			11	2.7%	NA	NA					
D10S.520-HEX	170	174	1.4976	5	1.2%	1.1945	0.1952	0.8041	32.7%	1.5849	24.6%	1.7744
D10S.520-HEX	174			1	0.2%	NA	NA					
D10S.520-HEX	178			1	0.2%	NA	NA					
D10S.520-HEX			$\Delta 0$	129	31.5%	NA						
D10S.520-HEX			$\Delta 04$	133	32.5%	1.3796	0.3320	0.7156	48.1%	2.0436	32.5%	2.6597
D10S.520-HEX			$\Delta 08$	76	18.6%	1.3764	0.3133	0.7499	45.5%	2.0030	31.3%	2.5264
D10S.520-HEX			$\Delta 12$	49	12.0%	1.4982	0.3162	0.8657	42.2%	2.1306	29.7%	2.5926
D10S.520-HEX			$\Delta 16$	15	3.7%	1.5041	0.3582	0.7877	47.6%	2.2206	32.3%	2.8722
D10S.520-HEX			$\Delta 20$	4	1.0%	1.5427	0.5662	0.4102	73.4%	2.6752	42.3%	5.8021
D10S.520-HEX			$\Delta 24$	2	0.5%	1.8623	NA					
D10S.520-HEX			$\Delta 28$	1	0.2%	1.7136	NA					

TABLE 10

MARKER	A1	A2	R	#	%[HET]	AVE	SD	LOW	%	HIGH	%	ADJ
D10S.1173-FAM	12	46		480	76.5%	1.4438	0.3241	0.7956	44.9%	2.0921	31.0%	2.6204
D10S.1173-FAM	124	140	1.5157	1	0.2%	1.5157	NA					
D10S.1173-FAM	128			1	0.2%	NA	NA					

MARKER	A1	A2	R	#	%[HET]	AVE	SD	LOW	%	HIGH	%	ADJ
D10S.1173-FAM	128	140	2.0231	2	0.4%	1.9854	NA					
D10S.1173-FAM	128	156	2.2872	2	0.4%	2.0880	NA					
D10S.1173-FAM	128	160	1.7646	1	0.2%	1.7646	NA					
D10S.1173-FAM	132	136	2.1770	1	0.2%	2.1770	NA					
D10S.1173-FAM	132	148	2.2964	2	0.4%	1.7887	NA					
D10S.1173-FAM	132	152	0.9612	1	0.2%	0.9612	NA					
D10S.1173-FAM	132	160	0.7901	1	0.2%	0.7901	NA					
D10S.1173-FAM	136			1	0.2%	NA	NA					
D10S.1173-FAM	136	144	1.8075	4	0.8%	1.4713	0.3116	0.8481	42.4%	2.0946	29.8%	2.5527
D10S.1173-FAM	136	148	1.9989	2	0.4%	1.6167	NA					
D10S.1173-FAM	136	152	1.9306	1	0.2%	1.9306	NA					
D10S.1173-FAM	136	156	1.9145	2	0.4%	1.5370	NA					
D10S.1173-FAM	140			5	1.0%	NA	NA					
D10S.1173-FAM	140	144	2.2913	9	1.9%	1.5296	0.3392	0.8513	44.3%	2.2079	30.7%	2.7484
D10S.1173-FAM	140	148	1.7114	9	1.9%	1.3886	0.2281	0.9324	32.8%	1.8447	24.7%	2.0678
D10S.1173-FAM	140	152	1.9542	10	2.1%	1.4612	0.3719	0.7174	50.9%	2.2050	33.7%	2.9762
D10S.1173-FAM	140	156	2.0704	6	1.3%	1.6186	0.2983	1.0220	36.9%	2.2153	26.9%	2.5637
D10S.1173-FAM	140	160	2.1555	2	0.4%	1.5297	NA					
D10S.1173-FAM	140	168	1.9572	1	0.2%	1.9572	NA					
D10S.1173-FAM	144			11	2.3%	NA	NA					
D10S.1173-FAM	144	148	1.8647	33	6.9%	1.4065	0.2840	0.8385	40.4%	1.9746	28.8%	2.3595
D10S.1173-FAM	144	152	2.1707	27	5.6%	1.3365	0.3546	0.6272	53.1%	2.0457	34.7%	2.8477
D10S.1173-FAM	144	156	2.1171	33	6.9%	1.4404	0.2834	0.8736	39.4%	2.0072	28.2%	2.3750
D10S.1173-FAM	144	160	2.2106	6	1.3%	1.5801	0.4081	0.7638	51.7%	2.3964	34.1%	3.2687
D10S.1173-FAM	144	164	1.7182	6	1.3%	1.4846	0.2169	1.0509	29.2%	1.9184	22.6%	2.0974
D10S.1173-FAM	144	168	1.7990	1	0.2%	1.7990	NA					
D10S.1173-FAM	148			42	8.8%	NA	NA					
D10S.1173-FAM	148	152	2.1788	50	10.4%	1.4629	0.3264	0.8100	44.6%	2.1158	30.9%	2.6420
D10S.1173-FAM	148	156	2.1614	45	9.4%	1.4112	0.2896	0.8321	41.0%	1.9904	29.1%	2.3934
D10S.1173-FAM	148	160	2.0270	17	3.5%	1.3903	0.2942	0.8019	42.3%	1.9787	29.7%	2.4105
D10S.1173-FAM	148	164	1.9255	5	1.0%	1.5573	0.3584	0.8405	46.0%	2.2740	31.5%	2.8852
D10S.1173-FAM	152			29	6.0%	NA	NA					
D10S.1173-FAM	152	156	2.0995	32	6.7%	1.4252	0.2923	0.8405	41.0%	2.0099	29.1%	2.4165
D10S.1173-FAM	152	160	1.8872	18	3.8%	1.3318	0.3202	0.6913	48.1%	1.9722	32.5%	2.5654
D10S.1173-FAM	152	164	1.9670	8	1.7%	1.5601	0.2742	1.0116	35.2%	2.1086	26.0%	2.4060
D10S.1173-FAM	152	168	2.2621	1	0.2%	2.2621	NA					

MARKER	A1	A2	R	#	%[HET]	AVE	SD	LOW	%	HIGH	%	ADJ
D10S.1173-FAM	156			12	2.5%	NA	NA					
D10S.1173-FAM	156	160	1.9573	17	3.5%	1.4907	0.2309	1.0289	~~3~~ 1.0%	1.9525	23.7%	2.1599
D10S.1173-FAM	156	164	1.3191	3	0.6%	1.0850	NA					
D10S.1173-FAM	156	168	0.8118	1	0.2%	0.8118	NA					
D10S.1173-FAM	160			10	2.1%	NA	NA					
D10S.1173-FAM	160	164	1.6621	6	1.3%	1.4123	0.3378	0.7367	~~4~~ 7.8%	2.0879	32.4%	2.7075
D10S.1173-FAM	164			2	0.4%	NA	NA					
D10S.1173-FAM	164	168	1.3328	1	0.2%	1.3328	NA					
D10S.1173-FAM			$\Delta 0$	133	27.7%	NA						
D10S.1173-FAM			$\Delta 04$	149	31.0%	1.4474	0.3020	0.8435	~~4~~ 1.7%	2.0513	29.4%	2.4837
D10S.1173-FAM			$\Delta 08$	106	22.1%	1.3698	0.3075	0.7548	~~4~~ 4.9%	1.9848	31.0%	2.4859
D10S.1173-FAM			$\Delta 12$	73	15.2%	1.4559	0.3166	0.8227	~~4~~ 3.5%	2.0890	30.3%	2.5764
D10S.1173-FAM			$\Delta 16$	22	4.6%	1.6484	0.3713	0.9058	~~4~~ 5.1%	2.3910	31.1%	2.9998
D10S.1173-FAM			$\Delta 20$	11	2.3%	1.4548	0.3972	0.6604	~~5~~ 4.6%	2.2491	35.3%	3.2044
D10S.1173-FAM			$\Delta 24$	1	0.2%	1.7990	NA					
D10S.1173-FAM			$\Delta 28$	4	0.8%	1.7308	0.3986	0.9337	~~4~~ 6.1%	2.5280	31.5%	3.2085
D10S.1173-FAM			$\Delta 32$	1	0.2%	1.7646	NA					

TABLE 11

MARKER	A1	A2	R	#	%[HET]	AVE	SD	LOW	%	HIGH	%	ADJ
D17S.974-NED	6	19		269	61.0%	1.2184	0.2319	0.7545	~~38~~ 1.1%	1.6822	27.6%	1.9674
D17S.974-NED	132			3	1.1%	NA	NA					
D17S.974-NED	132	140	1.0001	3	1.1%	1.2245	NA					
D17S.974-NED	132	144	1.4286	7	2.6%	1.3217	0.2613	0.7991	~~39~~ 5.5%	1.8442	28.3%	2.1859
D17S.974-NED	132	148	1.5529	3	1.1%	1.4896	NA					
D17S.974-NED	132	152	1.2918	2	0.7%	1.6792	NA					
D17S.974-NED	136			2	0.7%	NA	NA					
D17S.974-NED	136	140	1.3700	8	3.0%	1.1065	0.1631	0.7804	~~29~~ 5.5%	1.4327	22.8%	1.5690
D17S.974-NED	136	144	1.2691	6	2.2%	1.0780	0.1359	0.8062	~~25~~ 2.2%	1.3497	20.1%	1.4413
D17S.974-NED	136	148	1.4583	3	1.1%	1.3666	NA					
D17S.974-NED	136	152	1.3911	1	0.4%	1.3911	NA					
D17S.974-NED	140			64	23.8%	NA	NA					
D17S.974-NED	140	144	1.7523	66	24.5%	1.2194	0.2137	0.7919	~~35~~ 5.1%	1.6469	26.0%	1.8777
D17S.974-NED	140	148	1.8253	31	11.5%	1.1778	0.2552	0.6674	~~43~~ 3.3%	1.6882	30.2%	2.0785
D17S.974-NED	140	152	1.3913	3	1.1%	1.2956	NA					
D17S.974-NED	144			30	11.2%	NA	NA					

MARKER	A1	A2	R	#	%[HET]	AVE	SD	LOW	%	HIGH	%	ADJ
D17S.974-NED	144	148	1.4581	23	8.6%	1.1960	0.2445	0.7070	40.9%	1.6851	29.0%	2.0233
D17S.974-NED	144	152	1.5589	3	1.1%	1.3201	NA					
D17S.974-NED	148			6	2.2%	NA	NA					
D17S.974-NED	148	152	1.2369	5	1.9%	1.2004	0.1829	0.8345	30.5%	1.5663	23.4%	1.7267
D17S.974-NED			$\Delta 0$		39.4%	NA						
D17S.974-NED			$\Delta 04$	102	37.9%	1.2044	0.2157	0.7730	35.8%	1.6357	26.4%	1.8764
D17S.974-NED			$\Delta 08$	43	16.0%	1.1771	0.2372	0.7027	40.3%	1.6514	28.7%	1.9716
D17S.974-NED			$\Delta 12$	12	4.5%	1.3264	0.1851	0.9561	27.9%	1.6967	21.8%	1.8401
D17S.974-NED			$\Delta 16$	4	1.5%	1.4650	0.1576	1.1497	21.5%	1.7803	17.7%	1.8667
D17S.974-NED			$\Delta 20$	2	0.7%	1.6792	NA					

TABLE 12

MARKER	A1	A2	R	#	%[HET]	AVE	SD	LOW	%	HIGH	%	ADJ
D17S.1289-FAM	11	26		342	55.0%	1.1506	0.2128	0.7251	37.0%	1.5761	27.0%	1.8259
D17S.1289-FAM	126	142	1.3125	2	0.6%	1.0874	NA					
D17S.1289-FAM	130			16	4.7%	NA	NA					
D17S.1289-FAM	130	134	1.2073	3	0.9%	1.0784	NA					
D17S.1289-FAM	130	138	1.4978	13	3.8%	1.0899	0.1883	0.7132	34.6%	1.4665	25.7%	1.6654
D17S.1289-FAM	130	142	1.4732	43	12.6%	1.1345	0.1990	0.7366	35.1%	1.5325	26.0%	1.7475
D17S.1289-FAM	130	146	1.3904	23	6.7%	1.1186	0.1944	0.7298	34.8%	1.5074	25.8%	1.7146
D17S.1289-FAM	130	150	1.4635	2	0.6%	1.4225	NA					
D17S.1289-FAM	134			3	0.9%	NA	NA					
D17S.1289-FAM	134	138	1.1062	3	0.9%	0.9925	NA					
D17S.1289-FAM	134	142	1.4972	9	2.6%	1.1278	0.2533	0.6212	44.9%	1.6344	31.0%	2.0475
D17S.1289-FAM	134	146	1.2960	4	1.2%	1.2137	0.0831	1.0474	13.7%	1.3800	12.0%	1.4064
D17S.1289-FAM	134	150	2.0777	1	0.3%	2.0780	NA					
D17S.1289-FAM	138			8	2.3%	NA	NA					
D17S.1289-FAM	138	142	1.4432	19	5.6%	1.1436	0.1906	0.7624	33.3%	1.5248	25.0%	1.7155
D17S.1289-FAM	138	146	1.4968	13	3.8%	1.1479	0.2430	0.6618	42.3%	1.6339	29.7%	1.9909
D17S.1289-FAM	138	150	1.4819	4	1.2%	1.2330	0.1695	0.8940	27.5%	1.5719	21.6%	1.7004
D17S.1289-FAM	142			102	29.8%	NA	NA					
D17S.1289-FAM	142	146	1.4786	39	11.4%	1.1579	0.2092	0.7394	36.1%	1.5764	26.5%	1.8132
D17S.1289-FAM	142	150	1.4742	8	2.3%	1.2045	0.1992	0.8060	33.1%	1.6029	24.9%	1.7999
D17S.1289-FAM	142	154	1.3720	1	0.3%	1.3720	NA					
D17S.1289-FAM	146			15	4.4%	NA	NA					
D17S.1289-FAM	146	150	1.2587	1	0.3%	1.2587	NA					
D17S.1289-FAM	150			6	1.8%	NA	NA					
D17S.1289-FAM	158			2	0.6%	NA	NA					
D17S.1289-FAM	174			1	0.3%	NA	NA					
D17S.1289-FAM	178			1	0.3%	NA	NA					
D17S.1289-FAM			$\Delta 0$	229	54.9%	NA						
D17S.1289-FAM			$\Delta 04$	65	15.6%	1.1440	0.2001	0.7438	35.0%	2.6315	56.5%	1.7594
D17S.1289-FAM			$\Delta 08$	44	10.6%	1.1367	0.2178	0.7010	38.3%	2.5386	55.2%	1.8431
D17S.1289-FAM			$\Delta 12$	51	12.2%	1.1527	0.1916	0.7695	33.2%	2.6917	57.2%	1.7269
D17S.1289-FAM			$\Delta 16$	26	6.2%	1.1531	0.2701	0.6130	46.8%	2.3791	51.5%	2.1693
D17S.1289-FAM			$\Delta 20$	2	0.5%	1.4225	NA					

Tabular information such as that set forth in Tables 9-12 can be generated for every polymorphic marker and is not limited to microsatellites. This type of information can also be developed for single nucleotide polymorphisms (SNPs). The great advantage of such data is that it can be used in all situations without the need for a normal, non-neoplastic internal negative control specimen. This is especially helpful when dealing with cytology and biopsy specimens where the presence of non-neoplastic cells cannot be assumed to be present with certainty.

At this point, the proportion of cells or DNA accounting for the imbalance will be determined for each marker. This information is used with the information derived from microscopic analysis of the tissue section or cytology cells from where the samples originated. Particular attention here is paid to the extent of non-neoplastic cell inclusion, because these cells will provide normal DNA to the analysis. The inclusion of non-neoplastic cellular elements is not limited and is only mentioned in order to provide greater understanding for the quantitation of allelic imbalance. The inclusion of some degree of non-neoplastic cellular elements will not interfere with the final characterization of genomic deletion expansion. The proportion of cells or DNA affected by imbalance is determined for markers pertaining to a specific genomic deletion (*see* Table 9). A gradient of reduced mutation involvement indicates that a proportion of tumor cells with an expanded deletion are present in the area from which the sample was derived. This analysis is sensitive for detected genomic deletion up to the threshold representing two standard deviations from the mean for normal allele peak ratios.

3.8.3 Cytology Microdissection

Cytology microdissection is carried out in the same manner as topographic genotyping, with attention paid to the microscopic cytology appearance of the cells for selection to perform genotyping. For topographic genotyping™, *see* U.S. Patent No. 6,340,563 which is incorporated by reference in its entirety for all purposes. Targets clusters of cytology cells are markers with xylene resistant ink on the reverse face of the glass slide (FIG. 1). The coverslips are removed and designated cellular targets are microdissected off the glass slides using a manual approach or other technique for specimen microdissection, such as laser capture microdissection, *etc*). The microdissected cells are then processed for DNA quality determination and mutation detection as described herein.

The analysis of microdissected cellular material is carried out independently to that of the free DNA in the collected fluid sample. The results of each analysis may be compared to each other. While the two types of specimens would appear to be equivalent, they are in fact not equivalent. The microdissected cell genotyping informs the user of the alterations present in the cells themselves. The free-fluid genotyping informs the users of

the changes affecting a wider distribution of cells encompassing all the cells with the potential to contribute DNA into the fluid collection. For example, changes upstream of the fluid collection can be detected by analyzing the DNA that is collected in a downstream location. Other possible scenarios are shown in Table 13.

5 **TABLE 13: SIGNIFICANCE OF MOLECULAR ANALYSIS ACCORDING TO SPECIMEN TYPE**

Microdissected Cells	• Molecular data pertains to microdissected cells • Detailed molecular information on discrete collection of cells • Better opportunity to detect later acquired, lower amplitude mutational change • May give false understanding for the more actively proliferating and progressing areas due to sampling variation
Fluid Free DNA	• Molecular data pertains to all cells in contact with and contributing to fluid DNA • Potential to detect mutational damage upstream or from remote regions of the organ • Averaging across the DNA can result in later acquired, lower amplitude mutations not being detected • Less affecting by sampling variation

3.8.4 Genomic Deletion Expansion

A series of polymorphic markers are selected which may be in the form of
 10 microsatellites or single nucleotide polymorphisms. These are specifically designed to cover the location of the deletion as well as adjacent DNA both on the centromeric and telomeric side. A representative example is shown in Table 14 for the *APC* gene. The same approach can be applied to any gene of interest across all chromosomes. Also, the specific markers may vary according to the genomic roadmap, and the results will be the same. The
 15 greater the number of markers used which cover the DNA on either side of the genomic deletion, the more precise the existence and extent of genomic deletion expansion will be able to be characterized. Thus, any number of markers may be used. In Table 14, four markers are shown for simplicity, two on either side of the deleted gene. However, any number of markers can be used. Allelic imbalance analysis is then performed for each of
 20 the polymorphic markers. This involves PCR followed by electrophoresis to determine the balance status for polymorphic alleles. Any system of allelic imbalance analysis can be substituted and will yield the same results. At the conclusion of this step, the extent of the genomic deletion will be defined according to the status of the location of the various polymorphic markers. This will be described in more detail below with several clinical case
 25 examples provided.

TABLE 14: POLYMORPHIC MICROSATELLITE MARKERS SPANNING A DELETION GENOMIC REGION CONTAINING THE *APC* TUMOR SUPPRESSOR GENE

DISTAL MICROSATELLITE DISTANCE pter	PROXIMAL MICRO- SATELLITE DISTANCE pter	TUMOR SUPPRESSOR GENE DISTANCE pter	PROXIMAL MICRO- SATELLITE DISTANCE pter	DISTAL MICRO- SATELLITE DISTANCE pter
111.10 D5S2027	111.85 D5S1965	112.10 5q22.1 APC	112.20 D5S346	112.40 D5S1170
21.60 D9S736	21.80 D9S916	21.96 9p21 CDKN2A	22.10 D9S1814	22.20 D9S966
89.40 D10S579	89.50 D10S2394	89.60 10q23.31 PTEN	89.90 D10S541	90.00 D10S2339
7.10 D17S2159	7.25 D17S1763	7.50 17p13.1 TP53	7.60 D17S655	7.70 D17S1796

This analysis is then carried out in a quantitative manner to determine the proportion of cells and/or DNA affected by imbalance. The concepts described here can be applied to the source of DNA as well as to biological fluid specimens.

5 **TABLE 15: REPRESENTATIVE RESULTS FOR AN ANALYSIS OF GENOMIC DELETION FOR THE APC GENE REGION**

DISTAL MICROSATELLITE DISTANCE pter	PROXIMAL MICRO- SATELLITE DISTANCE pter	TUMOR SUPPRESSOR GENE DISTANCE pter	PROXIMAL MICRO- SATELLITE DISTANCE pter	DISTAL MICRO- SATELLITE DISTANCE pter
111.10 D5S2027	111.85 D5S1965	112.10 5q22.1 APC	112.20 D5S346	112.40 D5S1170
57%	83%		62%	NO IMBALANCE

As shown in Table 15, the APC gene region is subject to deletion damage. Polymorphic microsatellites are varying distances away from the deleted region, which demonstrates progressively small proportions of affected cells. Thus, the deletion is shown to be expanding, and the rate of expansion is quantitatively expressed by the diminishing proportion of cells manifesting this change in the topographic sites from where the source of DNA has been obtained. The use of additional markers for allelic balance representing genomic sites at different distance from the gene-of-interest will provide the user with a clearer understanding on how rapidly the genomic deletion is expanding.

3.8.5 Point Mutational Analysis

The fluid DNA and/or the microdissected cells can be evaluated for the presence of point mutational change as one of several different forms of cancer related damage.

TABLE 16: POINT MUTATIONAL CHANGE IN BETA-CATENIN

TTTCAGATTTGACTTTATTCTAAAAATATTTCATGCGT CATATCACAGATTCCTTTTTTTTAAATTAAAGTAACATT	BETA-CATENIN, EXON 3: POINT
---	--------------------------------

CCAATCTACTAATGCTAATACTGTTTCGTATTTATAGCT GATTTGATGGAGTTGGACA <u>TGGCCATG</u> GAACCAGAC <u>AGAAAAGCGGCTGTTAGTCACTGGCAG</u> CAACAGTCT <u>TACCTGGACTCTGGAATCCATTCTGGT</u> GCCACTACCA <u>CAGCTCCTTCTCTGAGTGGTAAAGG</u> CAATCCTGAGG <u>AAGAGGATGTGGATACCTCCCAAGTCC</u> TGTATGAGT <u>GGGAACAGGGATTTTCTCAGTCCTTCA</u> CTCAAGAAC <u>AAGTAGCTGGTAAGAGTATTATTTTCA</u> TTGCCTTACTG <u>AAAGTCAGAATGCAGTTTTGAGAACTAAA</u> AAGTTAGTG TATAATA	MUTATIONS ARE FOUND BETWEEN CODONS 32 TO 41
--	---

The domains used for primers in amplifying the region of the beta-catenin gene that is mutated are underlined with a single underline in Table 16. The section in bold is the region of the beta-catenin exon that is affected by mutational change.

- 5 The beta-catenin gene undergoes localized point mutational change in certain forms of human cancer. Oligonucleotide primers, designed according to the genomic sequence available from many different sources, are used to amplify the region potential subject to point mutational change. Another example is the defined acquisition of point mutational change in the *b-raf* gene, which occurs frequently in melanocytic tumors, papillary carcinoma of thyroid, and other neoplasms. Primers for assessing changes in *b-raf* are provided in Table 17.

TABLE 17: POINT MUTATIONAL CHANGE IN *B-RAF*

<u>B-RAF Protein</u>
N-MAALSGGGGGGAEPGQALFNGDMEPEA GAGAGAAASSAADPAIPEEVWNIQ MIKLTQEHIEALLDKFGGEHNPPSIYLEAYEYTSKLDALQREQQLLESLNGTD FSVSSASMDTVTSSSSSSLSVLPSSLSVFQNP TDVARSNPKSPQKPIVRVFLPNKQR TVVPARCGVTVRDSLKKALMMRGLIPECCA VYRIQDGEKKPIGWDTDISWLTGEE LHVEVLENVPLTTHNFVRKTFFTLAFCDFCRKL LFQGFRCQTCGYKFHQRCTEVP LMCVNYDQLDLLFVSKFFEHHPIPQEEASLAETALTS GSSPSAPASDSIGPQILTSPS PSKSIPIQPFRPADEHNRNQFGQRDRSSAPNVHINTIEPVNIDDLIRDQGFGRDGG STTGLSATPPASLPGSLTNVKALQKSPGPQREKSSSSSEDNRNMTLGRDSSDD WEIPDGQITVGQRIGSGSFGTVYKKGWHDVAVKMLNVTAPTQQLQAFKNEVG VLRKTRHVNILLFMGYSTKPQLAIVTQWCEGSSLYHHLHIETKFEMIKLIDIARQT AQGMDYLHAKSIHRDLKSNNIFLHEDLTVKIGDFGLATVYKSRWSGSHQFEQLSGS ILWMAPEVIRMQDKNPYSFQSDVYAFGIVLYELMTGQLPYSNNNRDQIIFMVGR GYLSPDL SKVRSNCPKAMKRLMAECLKKKRDERPLFPQILASIELLARSLPKIHRS ASEPSLNRAGFQTEDFS LYACASPKTPIQAGGYGAFPVH-COOH
<u>Exon 15 B-RAF</u>
TAAACTCTTCATAATGCTTGCTCTGATAGGAAAATGAGATCTACTGTTTTCCTT TACTTACTACACCTCAGATATA TTTCTTCATGAAGACCTCACAGTAAAAATA GGTGATTTTGGTCTAGCTACAGTGAAA TCTCGATGGAGTGGGTCCCATCA GTTTGAACAGTTGTCTGGATCCATTTTGTGGATGGTAAGAATTGAGGCTAT TTTTCCACTGATTAAATTTTGGCCCTGAGATGCTGCTGAGTTACTAGAAAGTC ATTGAAGGTCTCAACTATAGTATTTTCATAGTTCCCACT

A portion of the coding exon is highlighted in bold, representing sites of mutational

change that alters protein sequence.

Another example of point mutational change is detection of an EGFR point mutation in lung, non-small cell carcinoma and many other forms of neoplasia (Table 18).

TABLE 18: POINT MUTATIONAL CHANGE IN THE EPIDERMAL GROWTH FACTOR RECEPTOR GENE

EGFR: EXON 19	
CTGCGGGGGCGTCACAGCCCCCAGCAATATCAGCCTTAGGTGCGGCTCCACAG	
CCCCAGTGTCCCTCACCTTCGGGGTGCATCGCTGGTAACATCCACCCAGATCA	
CTGGGCAGCATGTGGCACCATCTCACAAT TG CCAGTTAACGTCTTCTCCTTCTC	
TCTCTGTCATAGGGACTCTGGATCCCAG AAG GTGAGAAAGTTAAAA TTCC	
CGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAACAAG	
AAATCCTCGATGTGAGTTTCTGCTTTGCTGTGTGGGGTCCATGGCTCTGA	
ACCTCAGGCCACCTTTTCTCATGTCTGGCAGCTGCTCTGCTCTAGACCCTGCT	
CATCTCCACATCCTAAATGTTCACTTTCTATGTCTTTCCCTTTCTAGCTCTAGTG	
GG	
EGFR: EXON 21	
CGTTCGCCAGCCATAAGTCCTCGACGTGGA GAG GCTCAGAGCCTGGCATGAAC	
ATGACCCTGAATTCGGATGCAGAGCTTCTTCCCATGATGATCTGTCCCTCACA	
GCAGGGTCTTCTCTGTTT CAG GGCATGA ACTA CTTGGAGGACCGTCGCTT	
GGTGCACCGCGACCTGGCAGCCAGGAAC CGT ACTGGTGAAAACACCGCAG	
CATGTCAAGATCACAGATTTTGGGCTGGCCAAACTGCTGGGTGCGGAAGA	
GAAAGAATACCATGCAGAAGGAGGCAAAGTAAGGAGGTGGCTTTAGGTC	
AGCCAGCATTTTCTGACACCAGGGACCA GG CTGCCTTCCCACTAGCTGTATT	
GTTTAACACATGCAGGGGAGGATGCTCTCCAGACATTCTGGGTGAGCTCGCAG	
CAGCTGCTGCTGGCAGCTGGGTCCAGCCAGGGTCTCCTGGTAGTGTGAGCCAG	
AGCTGCTTTGGGAACAG	

The coding exon is highlighted in bold, representing site of mutational change that alters protein sequence.

3.8.6 Copy Number Analysis

- Copy number analysis is used to search for and quantitatively characterize homozygous deletion and/or specific gene amplification. The simplest approach to accomplishing this objective is to perform a duplex PCR reaction for a gene-of-interest coupled with a comparator gene that is assumed not to vary with respect to cancer progression. This approach is not effective, because the two individual nucleic acid amplification reactions operate independently. More importantly, the independent nature of the reactions reside mainly in the operating characteristics of the amplification primers with respect to hybridization and denaturation. The Human Genome Database (Internet: <URL:http://www.gbd.org>) and Ensembl Genome Browser (Internet: <URL:http://www.ensembl.org>) was searched for primer sequences that were highly similar, especially with respect to the 3' end of the primers for the gene-of-interest and the comparator gene. This led to defining primer sequences for duplex PCR amplification in which the 3' end of the primers were highly similar. For example, the amplification

products of both reactions would be close in size, such that the amplification reactions would operate in a nearly equivalent fashion. This is exemplified with the *neu* gene, which can undergo homozygous deletion in many types of human cancer. The comparator gene containing a sequence of near identity to *neu* is the *APP* gene situated on chromosome 22 (Table 19).

TABLE 19: TARGET AMPLICONS OF NEU AND APP DESIGNED TO AMPLIFY USING PRIMER SETS WITH CLOSE 3' END HOMOMOLOGY

Neu	
5' -	GGATCCCTGATGGGGAGAATGTGAA AATTCCAGTGGCCATCAAAGT GTTGAGGGAAAACACAT CCCCAAAGCCAACAAAGAAATCTTAGA CGTAAGCCCCCTCCACCCTCTCCTGCTAGGAGGACAGGA AGGACCCCATGGCTGC AGGTCTGGGCTCTGGTCTCTCTTCATTGGGGTTTGGGGAGATATGACTCC CGCAAACCTAGACTA-3'
APP	
5' -	ATGCTTTATAAGGCCTCAACTCTTTTCTGGTTCATCAAAAAGTGCAGTTTAATTAGATGAAAA TAAGCAGCACTTAAATTCTAGATACAGAACTCCTACTCAAAAAAATATAGAAGTCATCTGTCCCCA GGAAGTGTGTATTTTATAC TCACCTAGT'TATTACACCAAGAAATAACACACAGAGATTGAACAATT CAGTACAAAGTTAGGATGAAACCTTCTGGC CATTCACGTAGCCATCGTATTCATCCTGAGAAAGGGAA ACCTAAAGACTTGGGAGAGTGTCTGGGAGGAAAGTGCTATTGAGAATGGGCTGTCCAAGCAAAGT GACAGGCAGGCCTCACTTAGCTAGAAGA GGGAGCTTTTCAGCCCTGCTCCACCATCAGGAATCTCC GTGACCTTGGACTACT GCCTTTAATCTCTTTTATCCTGGCTTCTAAAAGATACCGTCGAATTCATTC TTTTTCAAACCTATCAAAAATTTGAAATGTAGTGGTCAGGTTGAATGAAAGCCACGATAAGGAAA AGTGAGGGGGGTGC AGAACCTAGGAGTGGCTTACAGTGAGCAGCAGGAAGAGAAATTGGAAGATC AAAGGAAAGCATCAAAAATGATC ACAGGCCAGGCGTGGT-3'

The bolded domains denote the forward and reverse primers, which are designed to amplify portions of the noted gene. Each sequence in Table 19 is oriented in a 5' to 3' orientation. Note that the bases at the 3' end of the upstream primers for each gene are identical in sequence. Similarly, there is identity of the bases at the 3' end of the downstream primer. This homology together with the design of each primer having similar melting temperatures ensures that the duplex reaction will proceed in parallel relationship to each other. The ratio of peaks for each product using labeled primers on capillary electrophoresis provides a quantitative relationship with respect to the relative content. Using normal tissue, cytology cells, and biological fluid samples as a source of non-neoplastic DNA and sample variation, it is possible to define thresholds for significant gene loss or gain in pathologic samples. This approach is ideally suited for use with minute specimen samples, but can be used on larger samples. It opens the way for copy number analysis of individual genes in clinical samples of the type obtained in routine patient management.

3.8.7 Microsatellite Instability

Free DNA collected from fluid samples and microdissected cell clusters can be evaluated for shifts in microsatellite repeats to support the finding of DNA repair gene mutational damage. A simple yet thorough approach, ideally suited to microdissection

based analysis, is to carefully remove representative areas of cancer or precancerous lesions together with non-neoplastic tissue. The microdissected samples are then evaluated for shifts in microsatellite band positions.

The marker panel that can be used to develop this microdissection genotyping assays consists of two groups of 5 microsatellite markers. The criteria for determination of instability will follow closely that currently suggested in the literature. Patients will be classified as having either microsatellite stable, low level instability (1 out of 5 unstable markers), or high level instability (2-5 out of 5 unstable markers). The testing will be configured to use the second marker group in subjects initially found to have low level instability. Alternative microsatellites can be used with similar effectiveness, as may be required with different cancers.

3.8.8 Altered DNA Methylation Status

Certain common approaches for characterization of changes in DNA methylation status, such as methylation specific PCR, will not function properly when using minute amounts of fixative-treated DNA from clinical specimens. Disclosed herein is a more robust system in which bisulfite modification is performed on tissue sections and cytology specimens while still adherent to glass slides. Non-methylated cytosines are converted to thymidines following which the specimens are microdissected.

The PCR primers are carefully formulated to avoid overlapping potential methylated cytosines at C-G dinucleotide pairs. Methylation status is then assessed using capillary electrophoresis within the amplicons. The amplicons is formulated to include as many potential C-G dinucleotide pair overlaps thereby improving allele discrimination. This approach can be used for simple and effective methylation status detection with microdissected tissue samples. An example for the *CDKN2A* gene is shown in Table 20.

TABLE 20

1	Cggagagggg	gagaatagat	aaCgggCggt	ggggagtagt	atggagtCgg	CggCggggag
61	tagtatggag	ttttCggttg	attggttggt	taCggtCgCg	gttCgggggtt	gggtagagga
121	ggtgCgggCg	ttgttggagg	CgggggCggt	gtttaaCgta	tCgaatagtt	aCggtCggag
181	gtCgatttag	gttatgatga	tgggtagCgt	tCgagtggCg	gagttgttgt	tgttttaCgg
241	CgCggagttt	aattgCgtCg	atttCgttat	ttttattCga	ttCgtgtaCg	aCgttggtCg
301	ggagggtttt	ttggataCgt	tgggtggtgtt	gtatCgggtt	ggggCgCggt	tgggCgtgCg
361	Cgatgttttg	ggtCgtttgt	tCgtggattt	ggttgaggag	ttgggttatt	gCgatgtCgt
421	aCggtatttg	CgCgCggttg	Cgggggggtat	tagaggtagt	aattatgttt	gtatagatgt
481	CgCggaaggt	tttttagata	ttttCgattg	aaagaattag	agaggttttg	agaaatttCg
541	ggaaatttag	attattagtt	atCgaagggt	ttatagggtt	ataattgttt	tCgttataat
601	ttatttCgtt	ttCgtagttt	ttatttagaa	aatagagttt	ttaaaaatgt	tttggttttt
661	aaCgtagata	taagtttttt	tttatatatCg	taaatgttta	tttatattat	ttttatata
721	tttttataaa	aatgtaaaaa	agaaaatat	Cgtttttgtt	tttttattgt	gttgagggtt

781	tttggagtga	gtatttaCgt	tttaagCgta	tatttatgtg	ggtatttttt	gCgagtttCg
841	tagttttCgg	aagttgtCga	ttttatgata	agtattttgt	gaattaggga	agtttagggg
901	ggttattggt	tttttttgag	ttatatgtt	agtaaaggt	agaattaaag	tttaaataaa

In Table 20, the promoter sequence of the gene has been shown after bisulfite modification. Another example is shown for the MGMT gene, in Table 21.

TABLE 21

MGMT DNA METHYLATION DETECTION ASSAY	
OVERALL OBJECTIVE:	
To be able to quantitatively detect the ratio of methylated and unmethylated DNA for a particular gene. All four primers are used and the PCR product is run on GeneScan where peaks appear that are about 15 bases apart. The critical step is good bisulfite modification of tissue. The two sequences in the lower portion of the table display regions of the gene encoding epidermal growth factor receptor (EGFR) where point mutational damage can occur.	
IMMEDIATE GOAL:	
Mix together equal amounts of 1840 & 1841 to be used as the upstream primers. Do the same for the two downstream primers. Then apply to bisulfite modified microdissected tissue. Determine if a product is made. Product should always be made no matter what tissue used. Then if the reaction appears to work, the downstream primers can be labeled.	
1840U MGMT (METHYLATED BISULFITE MODIFIED TEMPLATE)	
5'-GTT GGG ATA GTT CGC GTT-3'	
1841U MGMT (UNMETHYLATED BISULFITE MODIFIED TEMPLATE)	
5'-GGA TAT GTT GGG ATA GTT TGT GTT-3'	
1842D MGMT (METHYLATED BISULFITE MODIFIED TEMPLATE)	
5'-CCA AAT ACC AAA AAA CGA CGA A-3'	
1843D MGMT (UNMETHYLATED BISULFITE MODIFIED TEMPLATE)	
5'-TTT TTC CAA ATA CCA AAA AAC AAC AAA-3'	
1	cccgcgcccc ggatatgctg ggacagcccg cgcccctaga acgctttgcg tcccgaagcc
61	cgcaggctct cgcggtgcgc accgtttgcg acttgcccc cccgcccccc ccgcccgc
121	ttggtacttg gaaaaatgga caaggattgt gaaatgaaac gcaccacact ggacagccct
1	ttCgCgtttC ggatatgttg ggatagttCg Cgttttttaga aCgttttgCg tttCgaCgtt
61	Cgtaggtttt CgCggtgCgt atCgtttgCg atttgttttt ttCgtttttt tCgtCgtttt
121	ttggtatttg gaaaaatgga taaggattgt gaaatgaaaC gtattatatt ggatagtttt

5

In Table 21, the promoter sequence of the gene has been shown after bisulfite modification.

A protocol for *in situ* bisulfite modification can be as follows:

1. Deparaffinize four micron thick tissue sections.
2. Treat tissue sections with Proteinase K (10 mg/ml, for 30 minutes).
3. Using Koplin jars (or equivalent container), incubate tissue section in 0.2 N NaOH for 20 minutes.
4. Incubate overnight in 3 M sodium bisulfite containing hydroquinone.
5. Rinse tissue sections in 0.3 N NaOH, three rinses for 10 minutes each.
6. Pass through graded alcohol and air dry. Tissue is ready for microdissection.

It is to be expected that *in situ* bisulfite conversion will not proceed to completion, and thus a proportion of DNA collected in the microdissection will not have undergone conversion. Primers are designed to take advantage of the presence of non-methylated cytosines at the 3' end of the oligonucleotide. The presence of unmodified DNA is not expected to interfere, given that PCR primers recognize only modified DNA. Moreover, the assay can be expected to operate in a stoichiometric manner in relationship to the quantity of modified DNA. This assumption is reasonable and supported by preliminary data.

Traditional DNA methylation assays based on bisulfite modification require the use of highly purified extracted DNA to allow full conversion. The use of consensus primers to amplify both methylated and unmethylated forms of modified DNA, makes the novel reaction as described here insensitive to a large degree by the extent of bisulfite modification. So long as partial modification does not discriminate between methylated and unmethylated forms of modified DNA, the efficiency of modification will be not critical. Efficiency of modification can be assessed by running of specific primers and probes in parallel on both lesional and non-neoplastic microdissected samples.

TABLE 22: PAX/PPARgamma GENE REARRANGEMENT STRATEGY FOR DETECTION IN FLUID SAMPLES AND MICRODISSECTED CELLULAR SPECIMENS

PAX8 Exon 6	
GATCAGGATAGCTGCCGACTAAGCATTGACTCACAGAGCAGCAGCAGC	GGACCCCGAAAGCACCTTCGCACGGATGCCTTCAGCCAGCACCACTC
GAGCCGCTCGAGTGCCCATTTGAGCGGCAGCACTACCCAGAGGCCTAT	GCCTCCCCAGCCACACCAAGGCGAGCAGGTGAGAAGCTGGGCCCT
GGGAGGTGAACAGGGTGGGCAAGGGCCAGAGAAGGTCTATTCTCCAAAT	
PAX8 Exon 7	
GCTGACTTCTCTTTGCTCCCCAGGCCTCTACCCGCTGCCCTTGCTCAACA	GCACCCTGGACGACGGGAAAGGCCACCCTGACCCCTTCCAACACGCCA
CTGGGGCGCAACCTCTCGACTCACCAGACCTACCCCGTGGTGGCAGGT	ACAACGCCCGGAGCCTCCC
TGGCAGGTGGAGGGCG	
PAX8 Exon 8	
TTGTCTCCCTCCCTCCGGCAGATCCTCACTCACCCCTTCGCCATAAAGCAG	GAAACCCCGGAGGTGTCCAGTTCTAGCTCCACCCCTTCTTTATCTA
GCTCCGCCTTTTTGGATCTGCAGCAAGTCGGCTCCGGGGTCCCGCCCT	TCAATGCCTTTCCCATGCTGCCTCCGTGTACGGGCAGTTCACGGGCC

AGGCCCTCCTCTCAG	GTACGACAGGGAGGTCCCGGGGCCATGCAG
<u>PAX8 Exon 9</u>	
GAGAGTGAGATGATACATATGGGAGCCCCCATGGTCCA	ACTGACGACTCTT
TGCTGTATTTTCCAG	<u>GGCGAGAGATGGTGGGGCCACGCTGCCCGGAT</u>
<u>ACCCACCCACATCC</u>	<u>CCACCAGCGGACAGGGCAGCTATGCCTCCTCTG</u>
<u>CCATCGCAGGCATGGTGGCAGGTAAGGAGAGG</u>	
<u>PPAR-GAMMA</u>	
TAGGACTTAACTTCACAGCTAGTCTATTTTTCCTTTCAGAAATGACCATGG	
<u>TTGACACAGAGATGCCATTCTGGCCCACTTTGGGATCAGCTCCG</u>	
<u>TGGATCTCTCCGTAA</u>	<u>TGGAAG</u>

Shown in Table 22 are exons 6-9 of the PAX8 gene and the first exon of the PPARgamma gene. The exon sequences of each domain are highlighted. Single nucleotide polymorphisms are indicated in bold and underlined. Exon sequences are in bold. The downstream primer in the PPAR gene is dotted underlined. Upstream primers can be created in each exon of the PAX8 gene for singleplex PCR reactions to test for the presence of gene rearrangement.

3.8.9 DNA Gene Rearrangement

Gene rearrangements can be detected using genomic DNA from fixative-treated specimens and fluid provided that the amplicon length is kept to a minimum. The longer the amplicon, the less likely the amplification will be successful. Positive detection of the rearrangement by positive nucleic acid amplification provides support for the existence of specific mutational alterations. An example is shown for the PAX/PPARgamma gene rearrangement present in certain forms of thyroid neoplasia (Table 22).

The most common approach to detect this rearrangement in aspirated thyroid fluid or microdissected thyroid cells is to convert the RNA to cDNA and then attempt to detect the rearrangement using primers designed to amplify a long segment of the rearrangement. This is less favored when dealing with fixative-treated, stained cytology, or free DNA from biological fluid samples, because the amplification may fail to work due to the poor quality of the DNA. Seeking to minimize the length of the amplified segment, separate primers are fashioned for each of the PAX exons that may be subject to translocation (Table 22). A similar approach can be used on other genes to detect gene rearrangements.

3.9 Analysis of Mutation Accumulation Sequence

Another aspect of the materials, methods and kits disclosed relates to the essential need to make greater use of the detected mutations. This aspect involves delineating a time course of detectable mutation accumulation in addition to simply noting the presence or absence of one or more mutations.

Mutation accumulation is also important to the diagnosis, prognosis, and treatment plan of a patient with an anomaly. The time course of mutation acquisition is currently not part of any system of pre-cancer or cancer diagnosis, classification, or characterization.

Currently, the emphasis is entirely upon cataloguing the presence of specific mutational alterations without regard for their temporal occurrence in relationship to each other. For example, DNA expression arrays and proteomic analysis of tumors make no attempt to align the detected alterations in terms of a temporal sequence of mutation acquisition. Time course considerations are overlooked, because there currently are no methods available for use with clinical tissue specimens that would permit such a delineation to be obtained.

Temporal sequence of mutation acquisition can, however, exert a dramatic effect upon cancer diagnosis and prognostication. This is exemplified by mutational acquisition in gliomas. For example, gliomas that acquire deletion of chromosome 1p and/or 19q manifest a greater sensitivity to chemotherapy with longer survival. The mechanism to account for this phenomenon is unclear, but may be related to mutational damage to DNA repair genes situated at 1p/19q that are involved in glioma tumorigenesis but are damaged with respect to nucleic acid repair. As a consequence of the damage, perhaps the chemotherapy is more effective. Thus, if the 1p/19q mutation is the first mutational change acquired, all subsequent glioma cells clonally derived from the affected cells will bear this mutational alteration. Although other mutational events may occur during the clonal expansion of tumor cells, the 1p/19q mutation will be essentially locked into all glioma cells. Thus, treatment targeting the 1p/19q mutation may be expected to dramatically affect the entire neoplasms. On the other hand, if 1p/19q deletion is acquired later in the time course of mutation acquisition, a significant component of the glioma will lack this alteration and may be able to resist chemotherapy. This may be the basis for recent observation that not all gliomas bearing a 1p/19q deletion show favorable treatment responsiveness.

Provided herein are methods and materials to determine the time course of mutation acquisition (FIG. 2). The approach is based on the established concept of clonal expansion of phenotypically more aggressive tumor cells. Clonal expansion is a unidirectional process replacing precursor neoplastic cells with a dominant tumor cell population of cells containing progressively more mutations (FIG. 2).

The first approach is carried out in tissue section using a microdissection genotyping technique. Thus, this aspect is carried out wherein a patient has provided both a biological fluid sample and a tissue sample from which both types of data can be provided. This approach may not be feasible in all fluid samples or in specimens composed of microdissected cells. Mutations present over a wider area of tumor are representative of having been acquired earlier in development of, for example, a breast cancer. Mutations present more focally are generally acquired more recently.

Another approach is depicted in FIGS. 3A and 3B. Quantitative determination of allelic imbalance mutation or other forms of mutation are arranged according to the degree of relative DNA involvement. Mutations with greater DNA involvement can be assigned

earlier temporal acquisition, while those with relatively less involvement can be defined as occurring later in time. By greater "DNA involvement" is meant a greater proportion of the DNA is subject to mutational change for a particular marker. Other marker mutations involving a greater proportion of the same DNA occur earlier in time, while those occurring later in time involve a lesser proportion of the target DNA.

Clinical validation for these concepts was demonstrated in gliomas classified according to the timing of 1p/19q deletion (Table 23). The latter has been associated with better treatment responsiveness, however this is not true for all patients. The subset of patient having acquired 1p/19q early with telomeric involvement of 1p by deletion showed significantly better treatment responsiveness. These experiments were performed on tissue specimens but the same analysis could be applied to fluid or cytology specimens.

TABLE 23: CLINICAL VALIDATION OF TG INTEGRATED MOLECULAR PATHOLOGY ANALYSIS OF GLIOMAS

MUTATIONAL PROFILE GROUP	n	PROPORTION WITH STABLE DISEASE	PROPORTION WITH COMPLETE RESPONSE	AVERAGE PROGRESSION FREE INTERVAL (months)
No detectable mutations (well-differentiated astrocytoma)	29	64%	20%	38 +/- 5
Early telomeric 1p deletion (oligodendroglial glioma)	42	82%	85%	54 +/- 6
Later telomeric 1p deletion (oligodendroglial glioma)	18	34%	24%	14 +/- 4
Interstitial 1p deletion (oligodendroglial glioma)	39	28%	48%	12 +/- 4
No 1p deletion; positive detectable mutations (high-grade astrocytic glioma)	63	14%	17%	11 +/- 3

Determination of the temporal sequence of mutation accumulation has not previously been effectively performed on fixative-treated clinical specimens or in a clinical context. Delineation of the time course of mutation acquisition is provided herein and is important to the diagnosis, prognosis, and treatment of a patient's disease. The identical profile of mutations between different individual patients with the same microscopic form of cancer may not be expected to behave in exactly the same manner. The order by which the mutations are accumulated may greatly influence the final biological behavior of the cancer.

This is especially true for response to treatment where earlier acquisition of treatment responsive mutations will be associated with greater therapeutic responsiveness in the treated patient due to the presence of those mutations.

Assuming a model of allelic loss with minimal non-neoplastic cell DNA inclusion, the percentage of mutated DNA may be determined for each marker. When two or more mutations, (e.g., K-ras-2 and/or allelic imbalance mutations), are detected, their time course of accumulation is inferred by the proportion of total DNA manifesting the alteration. In the case of K-ras-2 oncogene point mutation, involvement of 100% of cells is considered to be present when the intensity of the mutated base on sequencing is equal to or greater than the normal sequence base pair. This can be determined qualitatively by visual comparison and/or more precisely by quantitation. Alternatively, any method that provides relative amounts of mutated cell population can be used.

During the final evaluation of all the data from the methods, the presence of mutational change must first be analyzed with respect to the quantity and quality of DNA. It is important to do this analysis of mutational change, because the presence of low amounts of poor quality DNA can produce false-positive detection of deletion-type mutations due to a phenomenon in the PCR reaction called allelic drop-out.

These analytical methods may be used to diagnose cancer, pre-cancerous states, or non-neoplastic conditions. The analytical methods may also be used to determine the prognosis for a patient suffering from cancer, pre-cancerous state, or non-neoplastic condition. The analytical methods may be further used to determine a course of therapy or combination of therapeutic modalities for a patient suffering from cancer, a pre-cancerous condition, or an alternative neoplastic condition, as discussed herein. Because the methods allow for sensitive prediction of cancer and pre-cancerous conditions, a physician or other of skill in the art can use the methods to more accurately determine the prognosis of the condition. Furthermore, if a physician has advanced warning that cancer is present, treatment may be begun earlier and more carefully tailored to the condition (including surgery, radiation therapy, chemotherapy, and biologic therapy). To this end, the physician skilled in the art will be familiar with diagnosis, prognosis and treatment of breast anomalies. See, for example, DEVITA, JR. ET AL., CANCER: PRINCIPLES & PRACTICE ON ONCOLOGY, volume 2 (5th Ed., New York, 1997). Additionally, the methods described herein can be used to evaluate animals and animal models and their responsiveness to treatment modalities.

The insight gathered through temporal sequence analysis of acquired mutational damage was used to create a molecular classification of gliomas (FIG. 4). This classification has the advantage of providing more accurate information concerning the pattern of treatment responsiveness (FIG. 4). The classification also reconciles many of the

difficulties pathologists have had classifying gliomas based only on microscopic cellular appearance.

3.9.1 DETECTION OF FOCAL NEOPLASTIC PROGRESSION IN A HETEROGENEOUS SPECIMEN CONTAINING LOWER AND HIGHER GRADE NEOPLASIA

5

A common challenge for pathologists is to detect the presence of focal neoplastic progression in a heterogeneous lesion that contains a large component of lower grade neoplasia. DNA from a portion of the lesion of lower grade neoplasia, if present in a bathing fluid, may obscure the DNA emanating from the higher grade lesion. It is essential to detect the higher grade neoplasia, because its presence will determine how the disease will be treated and personalized for that subject.

10

The usual way this has been approached is by the use of a highly sensitive mutational detection assay, which selectively amplifies the mutation separate from the accompanying non-mutated DNA. This approach, while popular, has significant drawbacks in that even a minor degree of non-template based amplification of the desired amplicon will lead to a false-positive detection of mutation presence. This approach has a second very important drawback in that it is highly vulnerable to the presence of contaminating amplicons that may find their way into the PCR reaction. Once in the reaction, they will lead to a false positive results for the presence of mutation.

15

An alternative approach far better suited to address the needs for focal neoplastic progression is novel as it has not been reported. This consists of designing long-length amplicon production which can only be satisfied by the most intact and best quality DNA (Table 24). Tissues that have undergone *in vivo* degradation and/or have been exposed to chemical fixatives will fail to effectively amplify long-length amplicons. Progressive degradation of tissue will result in a corresponding failure to effectively amplify long-length products. By designing the PCR reaction to require that a long-length product be amplified, will result in selecting out those strands that are of sufficient length to accommodate the full amplification of the total pool of available DNA. This subset of DNA, in turn, will be derived from the most viable and actively proliferating cells. Conditions for PCR amplification, as well as other aspects of technical performance including mutation characterization by sequencing, are identical except for the requirement for long-length amplicon production. The appearance of mutation in the long-length amplicon only and not in the usual less than 200 base pair length PCR product is proof that there exists a focal area of neoplastic progression that has acquired this mutation. Short and long length PCR amplification of k-ras-2 exon 1 is displayed in Table 24. Boxed is codon 12 and 13 where point mutational damage occurs. The same upstream primer is used with varying downstream primers that increase the length of the amplicons. In this way, only high

20

25

30

35

quality DNA is used to achieve long-length amplification for mutational analysis by DNA sequencing. Suggested primers include those bolded and underlined below.

TABLE 24

TAGTGTATTAACTTATGTGTGACATGTTCTAATATAGTCACATTTTCATTATTTTATT
ATAAGGCCTGCTGAAAATGACTGAATATAAACTTGTGGTAGTTGGAGCT**GGTGCC**GTAGG
 CAAGAGTGCCTTGACGATACAGCTAATTCAGAATCATTTTGTGGACGAATATGATCCAAC
 AATAGAGGTAAATCTTGTTTAAATATG**CATATTACTGGTGACGAC**CATTCTTTGATACA
 GATAAAGGTTTCTCTGACCATTTTCATGAGTACTTATTACAAGATAATTATGCTGAAAGT
 TAAGTTAT**CTGAAATGTACCTTGGGTTTCA**AGTTATATGTAACCATTAAATATGGGAACCTT
 TACTTTCTTGGGAGTATGTCAGGGTCCATGATGTTCACTCTCTGTGCATTTTGATTGGA
 AGTGATTT**TCAGAGTTTCGTGAGAGGGT**AGAAATTTGTATCCTATCTGGACCTAAAAGAC
 AATCTTTTATTGTAACTTTTATTTTATGGGTTTCTTGGTATTGTGACATCATATGTAA
 AGGTTAGATTTAAATTGTACTAGTGAAATATAATTGTTTGATGGTTGATTTTTTTAACTT
 CATCAGCAGTATTTTCTATCTT**CTTCTCAACATTAGAGAACCTACAAC**TACCGGATAAA
 TTTTACAAAATGAATTATTTGCCTAAGGTGTGTTTATATAAAGGTACTATTACCAACTT

- 5 In a similar fashion, this approach can be configured in a novel way to detect allelic imbalance (loss of heterozygosity [LOH]) using polymorphic microsatellites (Table 25). A long-length PCR reaction is designed using suitable unlabeled primers, for example, as shown for three different microsatellites: D9S254, D17S1161, and D17S974. The PCR product is cleaned up using any commercially available kit or equivalent reagent system.
- 10 The isolated amplicon is then linearly copied using, for example, fluorescent labeled internal primers as exemplified in Table 25 (bolded and underlined). This linear amplification step is performed in the post-PCR area and is not vulnerable to contamination. Single strand product is then electrophoresed so as to quantify the relative content of each polymorphic allele. The presence of imbalance (LOH) in the long-length PCR reaction in the context of
- 15 normal allele balance in PCR reactions for products less than 200 base pair in length proves that there is a small focus of neoplastic progression in which this mutation has been acquired.

TABLE 25

D9S254
 CAAATCGTACTTAAGTGACT**TGCCCTACTTAAC**TATGGAGGAGTGTGAAGGGACTTACAAG
 TGGTCCCTAGGATTTCCATTTCCCTTAAGGACATTTAGATCAAAATTAGGACAATGGTCAT
 CACTTTGAGCCTTTGGATCTCTGCTTTGGGCCTTTTTTATAGAACAGAGGCAGATGTTTT
 ATTAAATTTCTTAAATGCGTGCCTATGTGTGATTTCAATTTTCATGTTTTGTGGCAGCAT
 TTTTGATCAATAATCCTTGGCTGCTATCTATTTTCTCTGTTCCCTTTTCTGACTTTTTTAAA
 TATCTCTAATTAATATATATAATATATATAAAACCTTATATATAATATATAATAAACT
 TTAGTTAACTAGTAGAGGATAAACCTGCTTCTGCTTCACTCAAAGATAACTTATATATA
 TGTATATATGGTTTATATATTACATAATTTGCATCAGTTATCCT**TGGGTAATAACTGCCGG**
AGAGATGGGAGTATATATATATGTGTGTGTATAGATAGGTAGATAGATAGATAGAT
AGATAGATAGATAGATAGATAATCTTGAGTGAAGCAGGTTTATCCTCTACTAGTTAATT

D17S1161
 GCACGACCAAAGCCCTGCTAATTCTGTTCATTCCAGGTGCCCTCCTGGTCCTCTCTGTCT

CCGGCTTCTAGAGGCCATCCCACCCTCAAACATCCTCTCTCTTTGGACAAACCCCCTGGC
CACCCTGACATTTGCAGAAGGCCAGCTCTTAGGCCAAGATAACCTTTATAGATCTTTTGGAT
TTTGTGTTACATATATCCCACAGCAATCCAGACACAAAACCGAGTTCTTCTTGTTTTT
CATAAAATA TGACTCACACCAAGAATTTACAATTTTAAATAGATTTCCCTGGCCTGGTG
CAGTGGCTCATGCCTGTAATCCCAACACTTTGGGAGGCTTAGGTGGGAGGATTGCTTGAG
GCCACAAGTTCGAGAACAGCCTGGGCAACATAGCAAGACCCCTGTCTCTACAAAAAATTT
AAAAAGTAGCAACACATGGTGGCACATGCCTGTAGTCCCAGCTACTCTGGAGGCTGAAAC
AGGAGGTACCTTGAGCCCACGAGGTTGAGGCTTCAGTTAGCCAAGATAATGCCATTGCA
CTCCAGCCTGGCAACAGAGCAAGACTGTCCAGATAGATAGATAGATGATGATAGATATA
TAGATATATAGATATATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGACTCT
CCTAAAGATTAGAGGGCACAGGGAGAAGTCCTTTATCCCCTTATGAGACACCTGGTCTTT

D17S974

TTCCACCCATTTGTCTGAGACAATTCAGTTATAGTTATAACATGTGACAG**GCATCATGCAG**
GGCCTTGAATTTGCTTCGTAGCAGAAATCTTTTAAGGCAGGTACTATAATTATTCACGTT
ACCGAGGAGGAAACTGAGGCTCAGAGAGGCGCTGTTATTCTTGCTGCAAGACTATGACAT
TCAGCTGGGTGCGATGGCTCACACCTGTAATTCAGCACTTTGGGAAGGTGACGGGGCAG
GATTTCTGGAGCCCAGGAGTTTGAGACCAGCCTGGGCAACATAGTGATACCTGTCTATA
CAAAATTTT TTTTAATTATCCAGACGTGGTGGCGAGTGCCTGTGGTCCCAGCTACTGG
GGAGGCTGAAGCAGGAGGAACACTTGAGCCAGGGAGGTCTAGGCTGCAGTGAGCTGAGAT
GGCGCCACTGCACCACTGCCTCCAGCCTGGGTGAGAGTGAGACCCTGTCTCAGATAGAT
GGATAGGTAGGCAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATA
GGTATGACATTCAGTTTGGGTGCCACCAACCTGTTAACAATGGCACCTCCTCACCTTAC

The sensitivity of this system to detect focal neoplastic progression is great, because it is not affected by inappropriate triggering of PCR, but rather by the absolute content of long-length template in the system. Detection thresholds of 1/1000th are readily attained by this approach; this means that one malignant cell in a background of, for example, 1000 adenomatous cells, can be detected as present and contributing significantly more intact DNA into the overall pool.

3.9.2 HANDLING OF THE LIQUID CYTOLOGY SPECIMEN

As described above, liquid cytology is increasingly being used for routine microscopic evaluation of specimens in cytology. It has reduced the rate of inadequate cytology preparations leading to better diagnosis. The approach encourages even dispersion of collected cells making for better microscopic interpretation. Any specimen can be adapted for liquid cytology analysis. Aspirated cells from solid specimens are ejected into liquid cytology fluid and then managed as a cellular containing liquid. Cellular containing liquid specimens can be mixed with fixative to result in the same type of liquid cytology specimen. Brushings can be vigorously stirred in the fixative and then process in the same fashion.

The presence of alcohol in the liquid cytology fixative facilitates the precipitation of free DNA during a cold centrifugation step. Molecular grade sodium acetate (e.g., Sigma Cat. No. 2899) can be added to a desired volume of liquid cytology fixative fluid containing

the cellular specimen for evaluation. The final concentration of sodium acetate is about 0.3 M. The contents are placed into a -20°C freezer for at least one hour, after which time the fluid/DNA/cellular mixture is centrifuged at high speed in order to create a pellet containing free precipitated DNA and pelleted cells. The supernatant is decanted, and the pelleted material is resuspended in 200 µL of 10 mM Tris-HCl, pH 7. A dense pellet can be resuspended in a large volume of Tris-HCl, pH 7. 200 µL of this mixture undergoes DNA extraction using for example a Qiagen kit or any other suitable method for extracting DNA from cells. The resuspended DNA is preferably placed in 10 mM Tris-HCl, pH 7, and is then ready for further genotyping.

10 **3.9.3. QUALITY CONTROL MEASURES**

Quality control measures to monitor the deletion expansion analysis are provided which consist of recommendations for replicate analysis at both the specimen level (sample replication), and replication of the individual marker results (marker replication) so that it may be most effectively applied to clinical specimens.

15 As in the case of all clinical analysis, it is important to have quality control measures in place to ensure that recorded results are accurate and reproducible. There are multiple layers of quality control to monitor analytic precision. First, the taking of multiple topographic samples provides a means for separate sample validation (FIG. 5). Quality control measures are shown for a cytology preparation. However, equivalent methods can also be applied to tissue specimens and to multiple samples of aspirated fluid or extracted DNA. Samples obtained close to or overlapping defined sites are more likely to be consistent than samples taken from widely separated regions of a tumor. The second level of quality control is replication of individual marker analyses (FIG. 5). The same marker performed in triplicate should yield equivalent quantitative results. The marker can be used 20 four or more times as well. The degree of variation may be used as a means to assess analytic precision with respect to marker replication. For determination of deletion expansion, the availability of additional polymorphic markers within the observed segment of expansion is a useful means to independently corroborate the results.

25 A preferred method for analytical validation can be achieved by comparing the molecular analysis from biopsy sampling of a particular tissue or organ to that obtained from analysis of microdissected samples from the resected specimen. A high degree of concordance should be evident.

35 **4. Data Acquisition, Compilation, and Weighting**

As data is obtained from the methods described herein, it can be further assessed via computer algorithms that weight the data obtained by the methods discussed with other risk factors. Fields of data included in such algorithms and data entry include, but are not

limited to: case number; accession date; collection date; name of physician; name of patient; social security number of patient or other patient identifier; information about the patient such as age or date-of-birth; any other person to be copied on the data report; patient history, including information from pathology report; ICD-9 code; pre-operative diagnosis; 5 post-operative diagnosis; information on any procedures performed; and the final diagnosis. Other fields considered by the algorithm include materials received relevant to the patient, such as stained and blank slides and the pathology report itself, as well as comments regarding tissue sections or fluid provided for analysis. Such comments may include information as the microdissection and comments from the pathologist, as well as to allelic 10 imbalance in each target area and mutation time course.

Using breast cancer as an example, the risk factors can be weighed and placed within a relational database. The data by the computer algorithms can include, but are not limited to patient age, reproductive history of the patient (*e.g.*, age at menarche, parity and age at first birth, age at menopause, and use of exogenous estrogens), family history (*e.g.*, 15 age at diagnosis, laterality of breast cancers, to be confirmed with pathology reports when possible, presence of other cancers, such as ovarian cancer in the case of breast cancer, and the number of unaffected relatives), prior breast biopsies (*e.g.*, number and histologic diagnoses of those biopsies), alcohol intake, smoking, diet, prior cancer diagnosis, cancer recurrence, and exposure to ionizing radiation.

20 Additionally, as data is acquired and a patient is staged according to a commonly accepted staging system, that data again can be added to the data obtained by the methods herein, and the compiled data weighted to provide the treating physician with a prognosis and treatment plan for the patient. Most staging for breast cancer involves analysis of the primary tumor (ranking of Tx, T0, Tis, and T1-T4), regional lymph node involvement 25 (which can involve clinical analysis (N) and/or pathologic analysis (pN)), distant metastases (M), and stage grouping (*i.e.*, TMN) of the prior factors.

EXAMPLES

EXAMPLE 1: BREAST DUCTAL LAVAGE ANALYSIS

30 The patient of this example was a female patient presenting with breast nipple discharge. Breast ductal lavage was performed and fluid samples were obtained for molecular analysis. Cytology evaluation of the same sample was considered insufficient for diagnosis.

200 μ L of breast ductal lavage fluid was treated and passed through a Qiagen spin 35 column (QIAamp DNA Mini Kit, Valencia, CA). The DNA was resuspended in 50 μ L of 10 mM Tris-EDTA buffer, pH 7.0. One μ L of the resuspended DNA was measured for DNA content and purity by optical density (NanoDrop Technologies, Wilmington, DE).

This was performed twice, consuming 2 μ L. The average of the two values for DNA content was used as the value for the sample. The resuspended DNA was then diluted to a value of 5 ng as the content of DNA was above that level. All subsequent analysis was performed on this normalized DNA content sample. A qPCR reaction for the first exon of the k-ras-2 gene was performed in duplicate to assess the amplifiability of the DNA. 2 μ L of DNA was used in each of these duplicate reactions (Icycler, BioRad, Hercules, CA). The Ct values for the duplicate reactions, which represent the quantity of amplifiable starting DNA, were averaged and used as a measure of DNA amplifiability. As an independent measure of DNA quality, duplicate 1 μ L aliquots of the normalized DNA was used in a PCR reaction to amplify a short (99 base pair, pseudogene) and long (154 base pair, true gene) segments of the glucocerebrosidase gene/pseudogene using the same pair of amplification primers.

At this point, mutational profiling was performed out for a broad set of cancer associated gene/genomic mutations using PCR (GeneAmp kit, Wellesley, MA) followed by capillary electrophoresis (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City CA). One μ L of normalized DNA was used in each singleplex reaction, with primers designed to amplify a specific genomic segment within or outside of a genes or genomic regions of interest. PCR was carried out according to manufacturer's instructions (GeneAmp kit, Wellesley, MA) with the following modifications. Desired oligonucleotide primers were added. A stock mixture buffer was then used to add *Taq* polymerase, deoxyribonucleotides and salts to the final amplification mix. The following was performed to enhance amplifiability. The magnesium concentration was significantly increased to a level of 8 mM in order to improve hybridization of the probe to the template DNA. Sucrose was also increased to a final concentration 12 g/100 mL (the function of which is not fully understood but may assist in the relaxation of double strands in turn encouraging primer annealing). For example, modifications may include adding additional 1.25 μ L magnesium chloride ($MgCl_2$) and 2.5 μ L of a 60% sucrose solution (Sucrose, Fisher Cat. No. AC41976-0010) in PCR Grade Distilled Water (PCR Grade Distilled Water DNase, RNase Free, 1 micron filtered; Fisher Scientific Cat. No. BP2470). 0.125 μ L of the upstream and downstream primers were then added to complete the mix. To this mix, 1 μ L of the crude lysate was added. Other substitutions using similar reagents can be used with equivalent effects to be expected. Other equivalent sugars may be substituted for sucrose. Further, other ions of equivalent valency, such as manganese, may be substituted for magnesium. The concentration of the reagents provided can be adjusted with correlative beneficial effects.

1 μ L of amplified DNA was mixed with 10-18 μ L of deionized formamide and 1 μ L of a 40% - 50% Rox Calibrator Solution. Samples were loaded onto the capillary

electrophoresis (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City, CA) and run through POP4 polymer. POP6 polymer from ABI is suitable, as well as any polymer which can discriminate the size of base pairs of interest. Samples were run through a 36 cm capillary array, although the 50 cm capillary array could be used. GeneScan software may be used, with the settings Dye set D and GS400HD Analysis Module. Injection times varied between 22 – 27 seconds, and run times varied between 1020 and 1500 seconds. Other software that allows the user to identify and measure peak height may be used, such as 3730xl DNA Analyzer software. Other capillary equipment from Beckman and Agilent may be used as well.

Amplification primers were designed with fluorescent labels attached to their 5' end so that they may be detected and quantified. Fluorescent labels include 5-HEX, TAMRA, 5,6-FAM (Integrated DNA Technologies, Inc., Coralville, IA) or (Synthegen, LLC, Houston, TX), and NED (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City, CA). As the fluorescently labeled DNA flows through the capillary electrophoresis, it is sorted by size, and as it passes a window, the laser fires to cause the labels to fluoresce. The fluorescence is captured by a camera as it passes the window. The time at which the fluorescence is captured determines the length of the patient's alleles, and the amount of fluorescence detected determines the peak height, which indicates the amount of PCR product that was amplified in the PCR reaction. The ratio of polymorphic peak heights or specific peak heights represents the amplification products, and was then used to detect and quantify mutational change that may be present. The software allows the user to choose peaks, and once a peak is chosen, it generates a row of data including the allele lengths and peak heights.

The data was copied into a Microsoft Excel® table. The Excel® table was then imported into a database which merges the capillary electrophoresis data with the sample information, so that results may be searched by patient's name or other identifiers, as well as used to generate a report of molecular results linked to each patient's target. Other spreadsheet software can also be used.

The first exon of the k-ras-2 oncogene was PCR amplified. The PCR product underwent a PCR clean up step using Microcon YM50 (Millipore Corp., Bedford, MA); instructions were followed to remove amplification primers. Cycle sequencing (BigDye Terminator Cycle Sequencing Kit V1.1, Applied Biosystems, Foster City, CA) by the dideoxy technique was performed following the kit instructions. Any method which allows for the separation and discarding of the primers from the PCR product may be used. DyeEx 2.0 Spin Kit (Qiagen, Hilden, Germany) was used for post-cycle sequencing, to remove excess unincorporated fluorescent dye and primers following their instructions. There are many kits on the market that can perform this task.

A 96-well standard PCR plate was used to load samples on the capillary electrophoresis equipment. Each well held 15 μ L of distilled or PCR Grade Distilled Water (PCR Grade Distilled Water DNase, RNase Free, 1 micron filtered, Fisher Scientific Cat. No. BP2470), and 5 μ L of the post cycle sequencing cleaned up sample. Samples were then entered into the Sequencing software program and run through capillary electrophoresis on the 3100 Genetic Analyzer (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City, CA) on the following settings Dye set E, mobility file DT3100POP4LR(BD)V1.mod, using the Run Module UltraSeq36_POP4DefaultModule and analyzed with BC-3100POP4UR_SeqOfftOff.saz module.

Fluorescent labeling was used and the sequence was read by fluorescent capillary electrophoresis. The iQ SYBR Green Supermix (Bio-Rad Cat. No. 170-8880, Hercules, CA) was used to make a 25 μ L volume reaction mix. Following manufacturer's instructions, 12.5 μ L of 2X SYBR Green Supermix was added, as well as 0.25 μ L of the upstream primer and 0.25 μ L of the downstream primer. In addition, 2.5 μ L of a 60% sucrose in PCR grade water (*i.e.*, the same water used for the other PCR reactions mixes) was added. The concentration of magnesium chloride was increased by adding 2 μ L of 25 mM magnesium chloride (Applied Biosystems, Foster City, CA). 5.75 μ L PCR grade water was added, to bring to the volume up to 23 μ L. 2 μ L of the DNA lysate (prepared from cytology or histology microdissected tissue or cells or buccal brush or other solid tissue as appropriate) or 2 μ L of DNA extracted (from fluid specimens) was added to the mix, using the Qiagen or equivalent kit or other DNA method of extracting DNA from cells. The iQ SYBR Green Supermix contains fluorescein which was used by a real-time PCR thermal cycler "iCycler," as a "well factor" background check. If another real-time PCR machine is used, it would be necessary to use another kit containing the appropriate fluorescent tag.

The mix may be prepared using the PCR mix described above with the addition of background fluorescence and the SYBR Green. There are also many detection methods known in the art, other than SYBR green, such as TaqMan® Probes. However, preferably 2.5 μ L of a 60% sucrose in PCR grade water should be added, and the concentration of magnesium chloride doubled, regardless of the detection method used.

The thermal cycling used is identical to that of PCR, except it is performed on a real-time thermal cycler such as the iCycler (Bio-Rad Cat. No. 170-8880, Hercules, CA). This real-time thermal cycler allows the immediate viewing of the increase in fluorescence, indicating an increase in PCR product. After PCR is complete, the real-time thermal cycler allows for the analysis of the amount of fluorescence detected (here, the detection of SYBR green). The software sets a threshold above background which indicates when the PCR reaction is in the exponential rate of growth phase. When each sample's fluorescence passes that threshold, the software assigns a "CT" or Cycle Threshold value. The CT is

calculated as the point in the PCR cycle when the curve passes the threshold. The lower the CT value, the more effective and more PCR product is being made; the higher the CT value, the less effective and less PCR product is made. If the sample does not pass the threshold, it is read as "NA" or not applicable, indicating that very little to no PCR product was made.

- 5 The CT infers how much and/or quality of the DNA was present at the starting point. DNA of poor quality (broken into small sections, cross-linked to proteins, etc.) or DNA present in very small amounts has a high to "NA" CT value. DNA of good quality and/or present in larger amounts has lower CT values. These results are printed out and used in the determination of the patient's diagnosis.

- 10 Genetic sequence information was obtained from the Genome Database (Internet:<URL:http://www.gdb.org>). Previously amplified DNA was labeled with radioactivity (^{35}S , ^{33}P , ^{32}P) and electrophoresed through 6-8% polyacrylamide bis-acrylamide gel.

- 15 The ratio of peaks was calculated by dividing the value for the shorter sized allele by that of the longer sized allele. Thresholds for significant allelic imbalance were determined beforehand in studies using normal (non-neoplastic) specimens representing each unique pairing of individual alleles for every marker used in the panel. Peak height ratios falling outside of two standard deviations beyond the mean for each polymorphic allele pairing were assessed as showing significant allelic imbalance. In each case, non-
- 20 neoplastic tissue served as a source of DNA to establish informativeness status and then used to determine the exact pattern of polymorphic marker alleles. Having established significant allelic imbalance, it was then possible to calculate the proportion of cellular DNA that was subject to hemizygous loss.

- 25 For example, a polymorphic marker pairing whose peak height ratio was ideally 1.00 in normal tissue with a 95% confidence interval from 0.78 to 1.62, could be inferred to have 50% of its cellular content affected by hemizygous loss if the peak height ratio was 0.5 or 2.0. This requires that a minimum of 50% of the DNA in a given sample be derived from cells possessing deletion of the specific microsatellite marker. The deviation from ideal normal ratio of 1.0 indicated which specific allele was affected. In a similar fashion, allele
- 30 ratios below 0.5 or above 2.0 could be mathematically correlated with the proportion of cells affected by genomic loss.

- Table 26 shows the various genomic targets that are being interrogated by oligonucleotide primer sets. Of importance is the delineation of the specific thermocycler profiles that are to be used to most effectively amplify these individual genomic targets
- 35 using the primers detailed in Table 34. The denaturing temperature in all cases is 95°C. The polymerization temperature in all cases is 72°C. The annealing temperature is shown by the first number and the total number of cycles needed is indicated by the second number.

Thus, for example, D17S974 was amplified with an annealing temperature of 55°C for 35 cycles. This provides the exact conditions by which each PCR is to be performed.

TABLE 26

The first number is the profile, the second number is the annealing temperature and the third is the # of cycles						
Panels/ Tests	3 M 55 30		2 55 35	5 53 35	6 53 30	ICYCLE R
9/14/2004		1 55 40				1 55 40
Major Analysis	D1s.407	D9s.254	D17S.1161	D9s.251	D3s.2303	
	D10s.1173	D5s.592	D1s.1193	D5s.615	D22s532	
	GAUCHER E9 FAM	D10s.520	D17s.974		D3s.1539	
		D21S.1244	D17s.1289			
Fluid Analysis	D1s.407	D9s.254	D17S.1161	D9s.251	D3s.2303	Kras2 E1
Pancreatic Fluid	D10s.1173	D5s.592	D1s.1193	D5s.615	D22s532	
	GAUCHER E9 FAM	D10s.520	D17s.974		D3s.1539	
		D21S.1244	D17s.1289			
Glioma	D1s.407	D9s.254	D17S.1161	D19s.400	D1s.1172	
	D10s.1173	D5s.592	D1s.1193	D9s.251		
	GAUCHER E9 FAM	D10s.520	D17s.974	D19s.559		
		Lmyc.5NT	D17s.1289	D5s.615		
ID QA Routine	D1s.407	AMELOG ENIN	D17s.1289	D19s.400	D3s.1539	
	D10s.1173	D5s.592				
		Lmyc.5NT				
MSI		D17s250	D2s123			
1st 5 markers			D5s346			
			BAT-25			
			BAT-26			

5

Results are presented below in Table 27.

TABLE 27

CYTOLOGY	Atypical cells present
DNA QUANTITY	30 ng/μL indicating high level of DNA
DNA QUALITY	- qPCR amplification of glucocerebrosidase at 26.5 cycles indicating good quality DNA - Ratio of short (pseudogene) to long (true gene) of exon 9 of the glucocerebrosidase gene is 0.85 indicating good

	quality DNA
MUTATIONAL PROFILING	Significant allelic imbalance detected for 1p36 and 17q21 acquired in that order
CONCLUSION	Cytology alone is indeterminate; integrated molecular pathology analysis is definitive for breast cancer
CONSEQUENCE	Patient undergoes surgical resection of early breast cancer with better outcome

EXAMPLE 2: BREAST DUCTAL LAVAGE ANALYSIS

The patient of this example was a female patient presenting with a breast abnormality detected by mammogram. Breast ductal lavage was performed and fluid samples were obtained for molecular analysis. Cytology evaluation of the same sample was considered insufficient for diagnosis.

200 μ L of breast ductal lavage fluid was treated and passed through a Qiagen spin column (QIAamp® DNA Mini Kit, Valencia, CA). The DNA was resuspended in 50 μ L of 10 mM Tris-EDTA buffer, pH 7.0. One μ L of the resuspended DNA was measured for DNA content and purity by optical density (NanoDrop Technologies, Wilmington, DE). This was performed twice consuming 2 μ L. The average of the two values for DNA content was used as the value for the sample. The resuspended DNA was then diluted to a value of 5 ng/ μ L as the content of DNA was above that level. All subsequent analysis was performed on this normalized DNA content sample. A qPCR reaction for the first exon of the k-ras-2 gene was performed in duplicate to assess the amplifiability of the DNA. 2 μ L of DNA was used in each of these duplicate reactions (Icycler, BioRad, Hercules, CA). The Ct values for the duplicate reactions, representing the quantity of amplifiable starting DNA, were averaged and used as a measure of DNA amplifiability. As an independent measure of DNA quality, duplicate 1 μ L aliquots of the normalized DNA were used in a PCR reaction to amplify short (99 base pair, pseudogene) and long (154 base pair, true gene) segments of the glucocerebrosidase gene/pseudogene using the same pair of amplification primers.

Mutational profiling was carried out for a broad set of cancer associated gene/genomic mutations using PCR (GeneAmp kit, Wellesley, MA) followed by capillary electrophoresis (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City CA). 1 μ L of normalized DNA was used in each singleplex reactions, with primers designed to amplify a specific genomic segment within or outside of a genes or genomic regions of interest. PCR was carried out according to manufacturer's instructions (GeneAmp kit, Wellesley, MA) with the following modifications. Desired oligonucleotide primers are added as desired. A stock mixture buffer is then used to add *Taq* polymerase, deoxyribonucleotides and salts to the final amplification mix.

The following was performed to enhance amplifiability. The magnesium concentration was significantly increased to a level of 8 mM in order to improve

hybridization of the probe to the template DNA. Also, the sucrose was added to bring the final concentration to 12 g/100 mL. For example, modifications may include adding additional 1.25 μ L magnesium chloride and 2.5 μ L of a 60% sucrose solution (Sucrose, Fisher cat# AC41976-0010 in PCR Grade Distilled Water (PCR Grade Distilled Water DNase, RNase Free 1 micron filtered, Fisher Scientific Cat. No. BP2470 or equivalent). 0.125 μ L of the upstream and downstream primers were added to complete the mix. To this mix, 1 μ L of the crude lysate was added. Other substitutions using similar reagents can be used with equivalent effects to be expected. Other equivalent sugars may be substituted for sucrose. Further, other ions of equivalent valency, such as manganese, may be substituted for magnesium. Also, the concentration of the reagents provided here may be adjusted with correlative beneficial effects.

The 1 μ L of amplified DNA was mixed with 10-18 μ L of deionized formamide and 1 μ L of a 40% - 50% Rox Calibrator Solution. Samples were loaded onto the capillary electrophoresis (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City, CA) and run through POP4 polymer. POP6 polymer from ABI is suitable, as well as any polymer which can discriminate the size of base pairs of interest. Samples were run through a 36 cm capillary array, although the 50 cm capillary array could be used. GeneScan software may be used, with the settings Dye set D and GS400HD Analysis Module. Injection times varied between 22 – 27 seconds and run times varied between 1020 and 1500 seconds. Other software that allows the user to identify and measure peak height may be used, such as 3730xl DNA Analyzer software. Other capillary equipment from Beckman and Agilent may be used as well.

Amplification primers were designed with fluorescent labels attached to their 5' end so that they may be detected and quantified. Fluorescent labels include 5-HEX, TAMRA, 5,6-FAM (Integrated DNA Technologies, Inc., Coralville, IA) or (Synthegen, LLC, Houston, TX), and NED (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City, CA). As the fluorescently labeled DNA flows through the capillary electrophoresis, it is sorted by size, and as it passes a window, the laser fires to cause the labels to fluorescence. The fluorescence is captured by a camera as it passes the window. The time in which the fluorescence is captured determines the length of the patient's alleles, and the amount of fluorescence determines the peak height, which indicates the amount of PCR product that was amplified in the PCR reaction. The ratio of polymorphic peak heights or specific peak heights representing defined amplification products was then used to detect and quantify mutational change(s) that may be present. The software allows the user to choose peaks, and once a peak is chosen, it generates a row of data including the allele lengths and peak heights.

The data was copied into a Microsoft Excel® table. The Excel® table was then imported into a database which merges the capillary electrophoresis data with the sample information, so that results may be searched by patient's name or other identifiers, as well as used to generate a report of molecular results linked to each patient's target. Other database software can be substituted.

The first exon of the k-ras-2 oncogene was PCR amplified. The PCR product underwent a PCR clean up step using Microcon YM50 (Millipore Corp., Bedford, MA) and following instructions to remove amplification primers. Cycle sequencing (BigDye Terminator Cycle Sequencing Kit V1.1, Applied Biosystems, Foster City, CA) by the dideoxy technique was performed following the kit instructions. Any method which allows for the separation and discarding of the primers from the PCR product may be used. DyeEx 2.0 Spin Kit (Qiagen, Hilden, Germany) was used for post-cycle sequencing, to remove excess unincorporated fluorescent dye and primers following their instructions. There are many kits on the market that can also be used.

A 96-well standard PCR plate was used to load samples on the capillary electrophoresis equipment. Each well held 15 µL of distilled or PCR Grade Distilled Water (PCR Grade Distilled Water DNase, RNase Free 1 micron filtered, Fisher Scientific Cat. No. BP2470 or equivalent) and 5 µL of the post cycle sequencing cleaned up sample. Samples were then entered into the Sequencing software program and run through capillary electrophoresis on the 3100 Genetic Analyzer (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City, CA) on the following settings Dye set E, mobility file DT3100POP4LR(BD)V1.mod, using the Run Module UltraSeq36_POP4DefaultModule and analyzed with BC-3100POP4UR_SeqOfftOff.saz module.

Fluorescent labeling was used and the sequence was read by fluorescent capillary electrophoresis. Genetic sequence information was obtained from the Genome Database (Internet:<URL:http://www.gdb.org>). Previously amplified DNA was labeled with radioisotopes (³⁵S, ³³P, ³²P) and electrophoresed through 6-8% polyacrylamide bis-acrylamide gel.

The ratio of peaks was calculated by dividing the value for the shorter sized allele by that of the longer sized allele. Thresholds for significant allelic imbalance were determined beforehand in studies using normal (non-neoplastic) specimens representing each unique pairing of individual alleles for every marker used in the panel. Peak height ratios falling outside of two standard deviations beyond the mean for each polymorphic allele pairing were assessed as showing significant allelic imbalance. In each case, non-neoplastic tissue served as a source of DNA to establish informativeness status and then to determine the exact pattern of polymorphic marker alleles. Having established significant

allelic imbalance, it was then possible to calculate the proportion of cellular DNA that was subject to hemizygous loss.

For example, a polymorphic marker pairing whose peak height ratio was ideally 1.00 in normal tissue with a 95% confidence intervals from 0.78 to 1.62, could be inferred to have 50% of its cellular content affected by hemizygous loss if the peak height ratio was 0.5 or 2.0. This requires that a minimum of 50% of the DNA in a given sample be derived from cells possessing deletion of the specific microsatellite marker. The deviation from an ideal normal ratio of 1.0 indicated which specific allele was affected. In a similar fashion, allele ratios below 0.5 or above 2.0 could be mathematically correlated with the proportion of cells affected by genomic loss.

Results are presented below in Table 28.

TABLE 28

CYTOLOGY	Insufficient for diagnosis
DNA QUANTITY	0.84 ng/μL indicating very low level of DNA
DNA QUALITY	- qPCR amplification of glucocerebrosidase at 37.0 cycles indicating very poor quality DNA - Ratio of short (pseudogene) to long (true gene) of exon 9 of the glucocerebrosidase gene is 0.13 indicating poor quality DNA
MUTATIONAL PROFILING	No definitive allelic imbalance detected. False positive allelic imbalance seen due to paucity of good quality DNA.
CONCLUSION	Cytology alone is not useful; integrated molecular pathology analysis is definitive for non-neoplastic breast disease
CONSEQUENCE	Patient spared further diagnostic biopsy and surgery to rule out malignancy

EXAMPLE 3: BREAST DUCTAL LAVAGE ANALYSIS: POSSIBLE LOCAL RECURRENCE OF PRIOR BREAST CANCER

The patient of this example was a female patient presenting with a strong family history of breast cancer, seen for routine examination. Given the family history, a decision was made to perform breast ductal lavage to improve the detection of cancer. Breast ductal lavage was performed and fluid samples were obtained for molecular analysis. Cytology evaluation of the same sample was considered insufficient for diagnosis.

200 μ L of breast ductal lavage fluid was treated and passed through a Qiagen spin column (QIAamp DNA Mini Kit, Valencia, CA) according to manufacturer's instructions designed to extract DNA. The DNA was resuspended in 50 μ L of 10 mM Tris-EDTA buffer, pH 7.0. One μ L of the resuspended DNA was measured for DNA content and purity by optical density (NanoDrop Technologies, Wilmington, DE). This was performed twice consuming 2 μ L. The average of the two values for DNA content was used as the value for

the sample. The resuspended DNA was then diluted to a value of 5 ng/ μ L as the content of DNA was above that level. All subsequent analysis was performed on this normalized DNA content sample. A qPCR reaction for the first exon of the k-ras-2 gene was performed in duplicate to assess the amplifiability of the DNA. Two μ L of DNA was used in each of these duplicate reactions (Icycler, BioRad, Hercules, CA). The Ct values for the duplicate reactions, representing the quantity of amplifiable starting DNA, were averaged and used as a measure of DNA amplifiability. As an independent measure of DNA quality, duplicate 1 μ L aliquots of the normalized DNA were used in a PCR reaction to amplify a short (99 base pair, pseudogene) and long (154 base pair, true gene) segments of the glucocerebrosidase gene/pseudogene using the same pair of amplification primers.

At this point, mutational profiling was carried out for a broad set of cancer associated gene/genomic mutations using PCR (GeneAmp kit, Wellesley, MA) followed by capillary electrophoresis (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City CA). One μ L of normalized DNA was used in each singleplex reactions with primers designed to amplify a specific genomic segment within or outside of a genes or genomic regions of interest. PCR was carried out according to manufacturer's instructions (GeneAmp kit, Wellesley, MA) with the following modifications. Desired oligonucleotide primers were added as desired. A stock mixture buffer was then used to add *Taq* polymerase, deoxyribonucleotides and salts to the final amplification mix. The following were performed to enhance amplifiability. The magnesium concentration is significantly increased to a level of 8 mM final concentration in order to improve hybridization of the probe to the template DNA. Also critical was the addition of sucrose to a final concentration 12 g/100 mL. For example, modifications may include adding additional 1.25 μ L magnesium chloride and 2.5 μ L of a 60% sucrose solution (Sucrose, Fisher cat# AC41976-0010 in PCR Grade Distilled Water (PCR Grade Distilled Water DNase, RNase Free 1 micron filtered, Fisher Scientific Cat. No. BP2470 or equivalent). 0.125 μ L of the upstream and downstream primers are then added to complete the mix. To this mix, 1 μ L of the crude lysate can be added. Other substitutions using similar reagents may be used with equivalent effects to be expected. Other equivalent sugars may be substituted for sucrose. Further, other ions of equivalent valency, such as manganese, may be substituted for magnesium. Also, the concentration of the reagents provided here may be adjusted with correlative beneficial effects.

The 1 μ L of amplified DNA was mixed with 10-18 μ L of deionized formamide and 1 μ L of a 40% - 50% Rox Calibrator Solution. Samples were loaded onto the capillary electrophoresis (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City, CA) and ran through POP4 polymer. POP6 polymer from ABI is suitable as well as any polymer which can discriminate the size of base pairs of interest. Samples were run through a 36 cm

capillary array, although the 50 cm capillary array can also be used. GeneScan software may be used, with the settings Dye set D and GS400HD Analysis Module. Injection times varied between 22 – 27 seconds, and run times varied between 1020 and 1500 seconds. Other software that allows the user to identify and measure peak height may be used, such as 3730xl DNA Analyzer software. Other capillary equipment from Beckman and Agilent may be used as well.

Amplification primers were designed with fluorescent labels attached to their 5' end so that they may be detected and quantified. Fluorescent labels include 5-HEX, TAMRA, 5,6-FAM (Integrated DNA Technologies, Inc., Coralville, IA) or (Synthegen, LLC, Houston, TX), and NED (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City, CA). As the fluorescently labeled DNA flows through the capillary electrophoresis, it is sorted by size and as it passes a window, the laser fires to cause the labels to fluoresce. The fluorescence is captured by a camera as it passes the window. The time at which the fluorescence is captured determined the length of the patient's alleles, and the amount of fluorescence determines the peak height which indicates the amount of PCR product that was amplified in the PCR reaction. The ratio of polymorphic peak heights or specific peak heights representing defined amplification products was then used to detect and quantify mutational change that may be present. The software allows the user to choose peaks, and once a peak is chosen, it generates a row of data including the allele lengths and peak heights.

The data was copied into a Microsoft Excel® table. The Excel® table was then imported into a database which merges the capillary electrophoresis data with the sample information, so that results may be searched by patient's name or other identifiers, as well as used to generate a report of molecular results linked to each patient's target. Again, other database software can be substituted.

The first exon of the k-ras-2 oncogene was PCR amplified. The PCR product underwent a PCR clean up step using Microcon YM50 (Millipore Corp., Bedford, MA) and following instructions to remove amplification primers. Cycle sequencing (BigDye Terminator Cycle Sequencing Kit V1.1, Applied Biosystems, Foster City, CA) by the dideoxy technique was performed following the kit instructions. Any method which allows for the separation and discarding of the primers from the PCR product may be used. DyeEx 2.0 Spin Kit (Qiagen, Hilden, Germany) was used for post-cycle sequencing, to remove excess unincorporated fluorescent dye and primers following their instructions. There are many kits on the market that can perform this task.

A 96-well standard PCR plate was used to load samples on the capillary electrophoresis equipment. Each well held 15 µL of distilled or PCR grade distilled water (PCR Grade Distilled Water DNase, RNase Free 1 micron filtered, Fisher Scientific Cat. #

BP2470 or equivalent) and 5 μ L of the post cycle sequencing cleaned up sample. Samples were then entered into the Sequencing software program and run through capillary electrophoresis on the 3100 Genetic Analyzer (ABI 3100 Genetic Analyzer, Applied Biosystems, Foster City, CA) on the following settings Dye set E, mobility file

- 5 DT3100POP4LR(BD)V1.mod, using the Run Module UltraSeq36_POP4DefaultModule and analyzed with BC-3100POP4UR_SeqOfftOff.saz module.

Fluorescent labeling was used and the sequence was read by fluorescent capillary electrophoresis. Genetic sequence information was obtained from the Genome Database (Internet:<URL:http://www.gdb.org>). Previously amplified DNA was labeled with
10 radioactivity (^{35}S , ^{33}P , ^{32}P) and electrophoresed through a 6-8% polyacrylamide bis-acrylamide gel.

The ratio of peaks was calculated by dividing the value for the shorter sized allele by that of the longer sized allele. Thresholds for significant allelic imbalance were determined beforehand in extensive studies using normal (non-neoplastic) specimens
15 representing each unique pairing of individual alleles for every marker used in the panel. Peak height ratios falling outside of two standard deviations beyond the mean for each polymorphic allele pairing were assessed as showing significant allelic imbalance. In each case, non-neoplastic tissue served as a source of DNA to establish informativeness status and then to determine the exact pattern of polymorphic marker alleles. Having established
20 significant allelic imbalance, it was then possible to calculate the proportion of cellular DNA that was subject to hemizygous loss.

For example, a polymorphic marker pairing whose peak height ratio was ideally 1.00 in normal tissue with a 95% confidence intervals from 0.78 to 1.62, could be inferred to have 50% of its cellular content affected by hemizygous loss if the peak height ratio was 0.5
25 or 2.0. This requires that a minimum of 50% of the DNA in a given sample be derived from cells possessing deletion of the specific microsatellite marker. The deviation from an ideal normal ratio of 1.0 indicated which specific allele was affected. In a similar fashion, allele ratios below 0.5 or above 2.0 could be mathematically correlated with the proportion of cells affected by genomic loss.

30 Results are presented below in Table 29.

TABLE 29

CYTOLOGY	Atypical cells present, cannot rule out malignancy.
DNA QUANTITY	18 ng/μL indicating high level of DNA.
DNA QUALITY	- qPCR amplification of glucocerebrosidase at 27.3 cycles indicating good quality DNA. - Ratio of short (pseudogene) to long (true gene) of exon 9 of the glucocerebrosidase gene is 0.75 indicating good quality DNA.
MUTATIONAL	Significant allelic imbalance detected for 5q23, 9p21, and

PROFILING	17p13 acquired in that temporal sequence of mutation accumulation order. Mutational profile and temporal sequence of mutation acquisition matches prior primary breast cancer.
CONCLUSION	• Cytology alone is indeterminate; integrated molecular pathology analysis is definitive for breast cancer and recurrence of prior breast cancer established as dynamic mutational profiles match.
CONSEQUENCE	• Patient proceeds to treatment for local recurrence of breast cancer. • No need for further diagnostic measures. Issue of recurrent versus <i>de novo</i> breast cancer completely resolved.

EXAMPLE 4: PANCREATIC CYST ASPIRATION ANALYSIS

A study was undertaken spanning 19 months. All patients presenting to Presbyterian University Hospital with a pancreatic cyst were eligible for inclusion.

5 Endoscopic Ultrasound-Guided (EUS) drainage of the cyst(s) was performed with a Pentax CLA echo-endoscope. Pancreatic cyst aspiration was performed according to clinical need (Wilson Cook N1 (22 gauge or 25 gauge needles)). Collected fluid from the pancreatic cysts was sent for cytological evaluation and CEA (carcinoembryonic antigen) tumor marker analysis. Intravenous (IV) antibiotic (Levofloxacin, 500 mg) was administered at

10 the time of cyst aspiration.

Cyst aspirate cytology exams were performed in a routine fashion, familiar to the skilled artisan in the medical profession. The primary tissue and fluid targets were from the head of the pancreas. Cytopathologic criteria for malignancy included nuclear enlargement, pleomorphism (minimum of 3 to 4 fold variation in nuclear size), elevated nucleus

15 cytoplasmic (N/C) ratio, nuclear membrane irregularity, and coarse chromatin (Solcia, et al., "Tumors of the Pancreas," IN ATLAS OF TUMOR PATHOLOGY, ARMED FORCES INSTITUTE OF PATHOLOGY (American Registry of Pathology, 1995). Cases diagnosed as inconclusive fulfilled some, but not all criteria for malignancy. Less than 1 mL (0.4 mL) of the aspirate was used for the study. DNA analysis was performed as detailed below.

20 DNA was extracted from the pancreatic fluid by spin column separation column (QIAamp DNA Mini Kit, Valencia, CA) (Qiagen, Valencia, CA). The spin columns are commercially available and operate on the basis for size discrimination and were used according to manufacturer's instructions. The fluid was centrifuged at the recommended speed. The centrifugal forces cause the fluid containing DNA to cross the membrane which

25 retards and holds back DNA that is beyond a certain minimal base pair length, usually about 70 bases. The DNA captured on the membrane was eluted off and collected in distilled, deionized water. The optical density (OD) was measured at 260 and 280 nanometer wavelengths. A higher OD value indicates higher concentration of DNA.

Quantitative PCR was performed using unlabeled oligonucleotide primers with sybr green signal to determine production of double-stranded DNA during cycling. Known quantitative controls and replicate microdissection of samples were used to standardize amplification reactions. As a further measure to monitor the effect of allelic imbalance, analyses were performed in replicate genotyping according to the availability of DNA. PCR was carried out in a standard fashion, but can be modified without affecting the application as discussed. Oligonucleotides flanking the microsatellite, single nucleotide polymorphism, or sequence of interest, were prepared based upon known sequence as can be obtained from GenBank. Other sources for genetic sequence information can also be used. The cycling temperatures used were a denaturing step at 95°C for 30 seconds, followed by annealing at 55°C for 30 seconds, followed by polymerization at 72°C for 60 seconds. Other temperature profiles can be used with equal effectiveness. *Taq* Polymerase Gold was used for DNA polymerization; however, other *Taq* or similar polymerase enzymes can be substituted.

One-microliter aliquots of pancreatic cyst aspirate were used in a polymerase chain amplification reaction (PCR) as described above for a broad panel of tumor suppressor genes commonly involved in human pancreatic carcinogenesis. Tables 30 and 31 provide the panel of mutations used in the study. *K-ras-2* gene and details on tumor suppressor genes (with associated markers) and chromosomal location and mutation type are provided (Table 34, provides sequences of the primers used herein). Other microsatellite markers or gene/genomic targets can be substituted with equal effectiveness.

TABLE 30

MARKER	CYTOGENETIC LOCATION	PROXIMITY TUMOR SUPPRESSOR GENE
D1S407	1p36	UNKNOWN
D1S1193	1p36	UNKNOWN
LMYC	1p34	UNKNOWN
D1S1172	1p22	UNKNOWN
D3S1539	3p25	VHL
D3S2303	3P26	OGG1
D5S592	5q23	APC

Tumor suppressor gene LOH was determined by analysis of tightly linked, informative polymorphic microsatellites. Use of two markers within each locus was used to increase the likelihood that at least one of the markers would be polymorphic within a subject, and thus informative for LOH analysis.

PCR amplification was designed to generate amplicons of less than 200 base pairs using synthetic oligonucleotide primers flanking each microsatellite. Oligonucleotide primers were created with 5' fluorescent moieties (*i.e.*, FAM, HEX, NED) suitable for

automated fragment analysis. The PCR products were analyzed by capillary electrophoresis using an ABI 3100 system according to manufacturer's instructions (Applied Biosystems, Foster City, CA). Allele peak heights and lengths were used to define the presence or absence of allelic imbalance (*i.e.*, LOH) for a given sample. Allelic imbalance was reported when the ratio of polymorphic allelic bands for a particular marker was beyond 95% confidence limits for the variation in peak heights for individual allele pairings derived from analysis using non-neoplastic specimen samples. In general, this value was below 0.5 or above 2.0 (Rolston, et. al., 2001 *J. Mol. Diagn.* 3: 129-32). DNA sequencing of K-ras-2 exon 1 PCR amplified DNA was used to search for and characterize point mutations in codons 12 and 13.

Allelic imbalance mutations were treated as genomic deletions associated with tumor suppressor genes. The ratio of allele peak heights is a measure of the admixture of mutated and non-mutated cells or DNA, and varies according to the individual pairing of specific microsatellite marker alleles. Allele ratios of about 2.0 or about 0.5 were said to be present when 50% of the total DNA was derived from cells possessing the loss. The deviation from an ideal normal ratio of about 1.0 indicated which specific allele was affected. Similarly, allele ratios below about 0.5 or above about 2.0 could be mathematically correlated with the proportion of cells affected by a genomic loss. The order of mutation acquisition is then arranged in a temporal sequence of mutation accumulation reflecting the proportion of cells affected by specific microsatellite marker loss. Markers displaying more extreme ratios (below 0.5 or above 2.0) are considered to have been acquired earlier in the development of the tumor. Assuming a model of allelic loss with minimal non-neoplastic cell DNA inclusion, the percentage of mutated DNA was determined for each marker. When two or more mutations (*i.e.*, K-ras-2 and/or allelic imbalance mutations) were detected, their time course of accumulation was inferred by the proportion of total DNA manifesting the alteration. In the case of K-ras-2 oncogene point mutation, involvement of 100% of cells was considered to be present when the density of the mutated base using sequencing autoradiography was equal to or greater than the normal sequence base pair. This was determined qualitatively by visual comparison. This can also be performed quantitatively in the same manner using automated fluorescent capillary gel electrophoresis, for example.

Continuous variables were presented as mean $\pm$ standard deviation (SD). Differences across groups were compared using a one-way analysis of variance. Multiple comparisons were adjusted using the *post-hoc* Bonferroni test. In cases where the variance across groups was not homogenous, as assessed by the Levene statistic, data was analyzed using a non-parametric analysis of variance or the Kruskal Wallis test. If this was significant, further, hypothesis driven two-group, non-parametric comparisons were

undertaken using the Mann-Whitney U test. A two-tailed p-value of <0.05 was considered significant. Sensitivity and specificity were calculated as appropriate. Data were analyzed using the statistical package, SPSS version 12.0 for Windows (SPSS Inc., Chicago IL).

Results. Thirty-three patients with pancreatic cysts were eligible for analysis. This was based on the presence of final surgical pathology (27 cases) or cytological proven cancer from the fine needle aspirate (FNA) sample (5 cases). Eleven cystic lesions were malignant (8 invasive cancer, 3 carcinoma-in-situ), 14 were premalignant (10 focal borderline histology, 5 no dysplasia), and 8 cystic lesions were benign (6 pseudocysts, 1 lymphoepithelial cyst, 1 mesothelial cyst). The results are provided in Table 32.

Five malignant cysts were diagnosed as malignant due to the presence of unequivocally malignant cells upon cytological evaluation of the FNA of an associated solid component in the cyst. These patients were not deemed surgical candidates, and therefore pathological confirmation is not available. Three patients having cysts with invasive cancer and 3 patients with carcinoma-in-situ underwent surgery. Of these, only 1 cyst was diagnosed as malignant on the basis of cytology. This group of 6 cases was composed of 4 IPMN and 2 mucinous cystadenocarcinomas.

There were 10 IPMN and 4 mucinous cystadenomas in the premalignant category. FNA cytology evaluation of only one premalignant cyst was diagnostic for a low-grade mucinous cystic neoplasm. Two premalignant cysts had inconclusive cytology and the others (11 cysts) had negative cytology. The 8 benign cysts had inconclusive cytology in one and a negative cytology evaluation in 7 cysts.

Cyst CEA level was available in 24 cases (9 malignant, 9 premalignant, and 6 benign cysts). Cyst fluid CEA analysis could not be performed in the other 9 cases due to insufficient fluid to run the test. Mean CEA level for the benign group is 17.7 (STD +/- 21.7) [95% CI 5.0 to 40.5], for the premalignant group is 5774 (STD +/- 10984) [95% CI 2668 to 14217], and for the malignant group was 108360 (STD +/- 251860) [95% CI 85236 to 301957]. This difference was significant between the benign and premalignant groups ($p = 0.001$) and the premalignant and malignant groups ($p < 0.05$). Due to considerable overlap and extreme values, receiver-operator-curves were generated for the log value of the CEA. A CEA level of 42 corresponded to a sensitivity and specificity of 94% to differentiate between benign cysts and MCN. A CEA level of 2320 yielded a sensitivity and specificity of 78% to differentiate between malignant and premalignant cysts.

The OD was available on 32 cyst aspirates. Mean OD for the 3 groups is as follows: benign 6.5 (STD +/- 5.9) [95% CI 1.5 to 11.4], premalignant 3.7 (STD +/- 2.6) [95% CI 2.2 to 5.2], and malignant group 16.5 (STD +/- 15.7) [95% CI 5.2 to 27.8]. The difference between the premalignant and malignant group was significant ($p = 0.008$). An

OD value of 7 yielded 80% sensitivity and specificity respectively for the presence of malignancy in a mucinous cystic neoplasm.

The test of DNA quality or amplifiable DNA (quantitative PCR) as presented by the mean qPCR cycle number for the 3 groups was available for 27 cysts. In 5 cysts (3 premalignant and 2 benign), there was no amplifiable DNA even after 40 cycles on the qPCR. Hence, for the sake of analysis 40 was assumed to be the value. One malignant sample did not undergo qPCR.

The results are as follows: benign 33 (STD +/- 4.9) [95% confidence intervals (CI) of 28.9 to 37], premalignant 31.2 (STD +/- 5) [95% CI 28.3 to 34.1], and malignant group 24.5 (STD +/- 4.5) [95% CI 21.3 to 27.8]. This difference was significant for the benign and malignant group (Bonferroni $p = 0.003$) and the premalignant and malignant group (Bonferroni $p = 0.007$). A qPCR cycle number of 27 yielded a sensitivity and specificity of 80% for the presence of malignancy in a mucinous cystic neoplasm.

Mutational analysis (LOH and K-ras point mutation) was available on all cyst aspirates. No mutations were detected in any of the 8 cysts in the benign group. Six of 14 cysts in the premalignant group carried mutations. One patient with 2 cysts (IPMN and counted as 1 case for the analysis) had an identical K-ras mutation followed by separate allelic losses. The mean number of mutations for the premalignant group is 0.9 (STD +/- 1.2) [95% CI 0.2 to 1.6]. Ten of the 11 malignant cysts carried multiple mutations. No mutations were detected in one cyst with HGD. The mean number of mutations in the malignant group is 2.8 (STD +/- 1.3) [95% CI 1.9 to 3.7]. The number of mutations differed significantly between the malignant and premalignant ($p = 0.003$), malignant and benign ($p < 0.001$), and premalignant and benign categories ($p = 0.036$).

The sequence of mutation accumulation was significantly different between the premalignant and malignant categories. All cysts with invasive cancer and 2 of 3 cysts with high grade dysplasia had first acquired a K-ras mutation followed by allelic loss. In contrast, 2 of 14 premalignant cysts acquired a K-ras mutation as the first step of DNA damage. Of these, only 1 cyst had an additional allelic loss. The occurrence of K-ras point mutation first with or without subsequent allelic loss was significantly associated with a malignant cyst ($p < 0.001$). The sensitivity and specificity of a K-ras mutation occurring as the first hit was 91% and 86% respectively, for the presence of malignancy in a mucinous cystic neoplasm. The presence of allelic loss following K-ras mutation yielded a sensitivity of 91% and specificity of 93%, for the presence of malignancy in a mucinous cystic neoplasm.

A cyst aspirate CEA level of 42 yielded a much higher sensitivity and specificity value on the receiver operator curves (94% compared to 79%). This maybe due to lack of uniformity across laboratories measuring CEA levels with different assays. Cyst aspirate

CEA level also appeared to predict the presence of early malignancy with moderate accuracy (sensitivity and specificity less than 80% for a CEA level of 2320), although confounded by extreme values. Nevertheless, the overwhelming evidence speaks for the presence of low or borderline amounts of intact DNA in low-grades forms of DNA.

- 5 Cytological analysis of pancreatic cysts in the absence of a solid component remained insensitive, reflecting the sparse number of cells in cyst aspirates and the subjectivity of the evaluation.

One aspect of the materials, methods, and kits involves the initial evaluation of the presence and quality of DNA followed by DNA mutational analysis. Benign pancreatic
 10 cysts (*e.g.*, pseudocysts) and low-grade MCN have a low rate of cellular turnover and by extension scant cyst fluid DNA. In contrast, malignant PCN should have uncontrolled cell growth and constant release of high quality, albeit mutated, DNA into the cyst fluid bathing the malignant lining. This is apparent from the results showing significantly different cyst aspirate OD and qPCR cycle values for the malignant and non malignant categories.

- 15 After the presence and cumulative amount of mutational damage in pancreatic cyst fluid was defined, the temporal sequence of individual mutation acquisition was inferred. The time course derived from analysis of the cyst fluid was compared to that derived from genotyping microdissected tissue samples from the pancreatic resection specimen. Table 33 provides the mutation acquisition pattern of four selected malignant specimens. The middle
 20 column provides the mutation sequence in microdissected tissue from the surgical specimen (first 3 cases) or from a positive cytology slide prepared from FNA of a solid component of a malignant cyst. The right column of Table 33 provides the mutation sequence from the corresponding cyst fluid aspirate. A near perfect correlation with respect to time course of mutational acquisition was found for the earlier occurring mutations, strongly supporting the
 25 validity of these methods used to detect and characterize mutational change. K-ras-2 point mutational change was noted as a first event followed by allelic loss in all patients but one ultimately demonstrated to have a malignant MCN. Only one patient in the premalignant group exhibited this pattern. This pattern of mutational damage was found to be very sensitive and specific for the presence of malignancy in MCN (91% and 93% respectively).
 30 Prior studies in brushings of malignant biliary strictures (GUT) also report the presence of this degree of mutational damage to be supportive of the presence of cancer.

TABLE 31

<u>Gene</u>	<u>Mutation Type</u>	<u>Locus</u>	<u>Microsatellite Marker</u>	<u>GeneBank Reference</u>
K-ras	Point mutation	12p12		
CMM/RIZ	Deletion	1p36 - 1p34	D1S407 L-Myc	L18040 M19720
VHL	Deletion	3p26 - 3p25	D3S1539 D3S2303	L16393 L17972
APC	Deletion	5q23 - 5q23	D5S592 D5S615	L16423 L18737
P16	Deletion	9p21 - 9p23	D9S251 D9S254	L18726 L18050
PTEN	Deletion	10q23 - 10q23	D10S520 D10S1173	L16357 L30341
P53	Deletion	17p13 - 17p13	D17S974 D17S1289	G07961 G09615

TABLE 32

<u>OBSERVATION</u>	<u>PATHOLOGY</u>	<u>CYTOLOGY</u>	<u>CEA</u>	<u>OD</u>	<u>qPCR cycle#</u>	<u>MUTATIONS</u>
1	MCAC	Positive	58960		25.2	kras-9p,9p-17p,17p
2	Malignant	Positive	77600	20.95		kras-1p-1p
3	IPMN/CA	Inconclusive	3.8	1.9	27.1	kras-10q-9p-5q
4	Malignant	Positive	23000	27.7	25.3	kras-9p-1p
5	Malignant	Positive	4080	16.5	25.7	kras-17p-1p-5q
6	Malignant	Positive		2.25	13.9	kras-5q
7	Malignant	Inconclusive	776000	18.9	23.3	kras-9p-17p
8	MCAC	Inconclusive	30150	9.3	22.9	kras-1p
9	IPMN/HGD	Negative	3871	7.45	31.4	kras-9q-5q
10	IPMN/HGD	Inconclusive		54.3	23.1	No mutation
11	IPMN/HGD	Negative	1580	6.1	27.7	K-ras-10q
12	IPMN	Negative	2550	5.8	27.2	K-ras-17p/3p
13	IPMN	Positive	98	6.6	29	K-ras
14	MCA	Inconclusive	15492	5.4	30.9	No mutation
15	MCA	Negative	109	7.55	25.9	5q-kras-1p
16	MCA	Negative	31900	3.1	30.5	3p-kras
17	IPMN	Inconclusive	879	0.4	40	No mutation
18	IPMN	Negative	203	1.8	31.8	No mutation
19	MCA	Negative	614	2.4	27.3	3p-1p
20	IPMN	Negative	126	3.05	29.3	No mutation
21	IPMN	Negative		2.6	27.6	No mutation
22	IPMN	Negative		7.7	27.2	No mutation
23	IPMN	Negative		4.55	30.7	5q-1p-10q
24	IPMN			0.8		No mutation
25	IPMN					No mutation
26	Lymphoepithelial cyst	Negative		14.4	29.4	No mutation
27	Pseudocyst	Negative	56.8	14.8	27.2	No mutation
28	Pseudocyst	Negative		2.5	30.1	No mutation
29	Pseudocyst	Negative	5.5	1.8	29.8	No mutation
30	Pseudocyst	Negative	2.6	1.2	40	No mutation
31	Pseudocyst	Negative	12.1	1.2	32.4	No mutation
32	Pseudocyst	Negative	28.8	10.7	40	No mutation
33	Mesothelial cyst	Inconclusive	0.5	5.4	40	No mutation

100

TABLE 33

<u>PATHOLOGY</u>	<u>PATH MUTATION- ORDER</u>	<u>ASPIRATE MUTATION- ORDER</u>
Mucinous Cystadenocarcinoma	kras-9p,9p-17p,17p-9q-1p	kras-9p,9p-17p,17p
Mucinous Cystadenocarcinoma	kras-9p-17p-3p	kras-9p-17p
IPMN with high-grade dysplasia	kras-9q-9p-5q-17p	kras-9q-5q
Mucinous Cystadenocarcinoma	kras-1p-10q-5q	kras-1p

**EXAMPLE 5: INTEGRATED MOLECULAR PATHOLOGY ANALYSIS OF FINE
NEEDLE ASPIRATION BIOPSY SPECIMENS FOR DIAGNOSIS AND TUMOR
CHARACTERIZATION**

Methods:

Residual, unused fine needle aspiration biopsy (FNAB) material was collected *ex vivo* from 17 patients with solid organ cancers (lung, esophagus, stomach, colon, breast, kidney, and soft tissue), together with corresponding tissue sections of the resected organ for correlative molecular pathology analysis. The corresponding tissue sections served as the standard for mutational profiling. The tissues samples were treated as discussed in Examples 1-4, *supra*.

Three types of representative FNAB samples were obtained from each patient, consisting of 1) 50-100 μ L of original cellular fluid (OCF), 2) 15-30 mLs of residual ThinPrep fixative (TPF), and 3) cytology slides with cover slips removed for microdissection of individual cell clusters (MCC). The DNA from each sample was extracted and PCR amplified for a broad panel of markers targeting 1p, 3p, 5q, 9p, 10q, 17p, and k-ras-2. Quantitative genotyping for LOH and point mutational analysis determined the cumulative amount and temporal sequence of the mutation acquisition.

Results:

A total of 221 individual genotyping reactions were performed resulting in 155 informative results and 83 detectable mutations. The concordance rate for mutational change was OCF 92%, TPF 94%, MCC 92%, and all three methods 99%. The time course of mutation acquisition was established by each method, and was similar to that defined in resected tissue in all cases. Molecular analysis did not interfere with traditional cytology practice for evaluating these specimens.

Thus, reliance solely upon microscopic alterations can be limiting for interpretation of FNAB from certain organs and sites. Integrated mutational analysis before, during and following cytology evaluation, using residual samples not otherwise required for practice, can provide highly discriminating information to diagnose, classify and prognosticate cancer and related lesions.

The primers in Table 33 are all listed in a 5' to 3' direction.

TABLE 34: Primers

Primer Number	Primer Name	Sequence
FORWARD	D10S1173	FAM-TCATGCCAAGACTGAAACTCC
REVERSE	D10S1173	GCTGGCCATGACTGTTTTAC
REVERSE	D10S520	GTCCTTGTGAGAACTGGATGC
FORWARD	D10S520	HEX-CAGCCTATGCAACAGAACAAG
FORWARD	D17S1161	HEX-AACAGAGCAAGACTGTCCAG
REVERSE	D17S1161	GTCCCTCTAACCTTTAGGAGAG
REVERSE	D17S1289	FAM-CTGCCTCTAAGCAGTCATTTAGA
FORWARD	D17S1289	GCATGGTCTTTTTCCATTCC
FORWARD	D17S974	NED-AGCCTGGGTGAGAGTGAGAC
REVERSE	D17S974	GCCATTGTTAACAGGTTGGTG
REVERSE	D18S814	TET-CCCACTATATGTATGTTCCACC
FORWARD	D18S814	CTCTCTGCCTCTCCCACC
REVERSE	D19S400	NED-CAGGGTCTTATTTCTGTCTGTC
FORWARD	D19S400	GCCTCTATAAATAAATAAAGAC
REVERSE	D19S559	TET-TTGAGGTATCTATGTGGATATC
FORWARD	D19S559	GAGTGAGACCCTGTCTTTAC
FORWARD	D1S1193	TAMRA-TCGGCGACATAGCCAGAG
REVERSE	D1S1193	CTTTGATCTAAGGATTACCTAC
REVERSE	D1S407	TGGGCGGGGGATAGAAGG
FORWARD	D1S407	FAM-TGCTAACCACATGGAGAGG
FORWARD	D21S1244	HEX-TCTTCTATCTCATATGTGTATC
REVERSE	D21S1244	GGAGGAACTTGAGGATGTG
FORWARD	D22S532	FAM-CCTGGGCAACAGAGCGAG
REVERSE	D22S532	GTCTGAGAAGATACTTGATATAG
FORWARD	D3S1539	NED-CTCTTTCCATTACTCTCTCC
REVERSE	D3S1539	GTTCTCCATCTATCTTTCTCTC
REVERSE	D3S2303	TET-CTCCAGAGCTTTGTTTTCAAC
FORWARD	D3S2303	GTGCCTACATGTTAGTATCC
REVERSE	D5S592	TAMRA-TGGATACATATTGTTTTCTGCTG
FORWARD	D5S592	GGTGTCAACAAAGTAATGTAAAG
REVERSE	D5S615	FAM-TCCACAGTGGTAAGAACCAG
FORWARD	D5S615	GAGATAGGTAGGTAGGTAGG
REVERSE	D7S1530	FAM-ACACAGTCTTCAGCCCTAC
FORWARD	D7S1530	ACAGAGCTAAACTCTGTCTC
REVERSE	D8S373	HEX-TGTATGCATTATTTCTTCAATC
FORWARD	D8S373	GTAGGCAGGTGGGTGGGC
REVERSE	D9S251	HEX-CAATACTTTTTAAGGCTTTGTAGG
FORWARD	D9S251	GTTTTATGTGCACTAACTAATGG
FORWARD	D9S252	NED-TTGTCAACTCCTAATATGGAC
REVERSE	D9S252	GATATCCCCAAGTTCTCATAC
REVERSE	D9S254	GAGGATAAACCTGCTTCACTCAA
FORWARD	D9S254	FAM-TGGGTAATAACTGCCGGAGA
FORWARD	L-MYC	FAM-TGAACCGTAGCCTGGCGAG
REVERSE	L-MYC	GCTGTTCTTCCTTTTAAGCTG
FORWARD	D1S1172	GAACAGAACCTGGTACCTTC
REVERSE	D1S1172	FAM-CTG CAC CCG GCT GAT GTTC
REVERSE	KRAS 2 Exon1	TCCTGCACCAGTAATATGCA
FORWARD	KRAS 2 Exon1	TAAGGCCTGCTGAAAATGACTGA
FORWARD	D5S346	HEX-ACTCACTCTAGTGATAAATCGGG

Primer Number	Primer Name	Sequence
REVERSE	D5S346	AGCAGATAAGACAGTATTACTAGTT
FORWARD	D5S1170	FAM-AAATTAAGTGTAAACCCTAGTGTG
REVERSE	D5S1170	TCTACCCTGAAGTAGCTCCAAAC
FORWARD	D5S2027	FAM-ACCTGGCAGATTTTCCACTC
REVERSE	D5S2027	CACCTCATTGACTGGGAC
FORWARD	D5S1965	FAM-TGTCCCGTTGATAAAAATTACTGCG
REVERSE	D5S1965	GTGTCTGGGATTTCTACGCAATG
FORWARD	D9S916	HEX-ATCGATCAGGTGCCAGAG
REVERSE	D9S916	AAAGGGAGAATATATTACTTGTGCAG
FORWARD	D9S736	HEX-TTCTAGACCTCTCAGCAGAC
REVERSE	D9S736	GATAGTGTGGAGACACCAG
FORWARD	D9S1814	FAM-CATGGTCTTCTACTCAGG
REVERSE	D9S1814	CAGAATGCTGGYYCTTAGCCTAGG
FORWARD	D9S966	GATTGTACCAACAGCACTGAG
REVERSE	D9S966	FAM-CAGGCTTGATCAGCTTCTG
FORWARD	D10S541	HEX-AGAAGTGAATAGAAGGAAC
REVERSE	D10S541	AGGGATGGGTACAGAACTC
FORWARD	D10S2339	FAM-CTCTAGGCTCATTTCTCTTCC
REVERSE	D10S2339	CACAAATGCAGATAAAATAAGACA
FORWARD	D10S579	HEX-ACTAGGAAGGTTTCATATCC
REVERSE	D10S579	TCATCCCCTGAAATACACC
FORWARD	D10S2394	HEX-CCAACGAGATGATGTTGCC
REVERSE	D10S2394	AAGAGCTGGGTGGGGTACT
FORWARD	D17S655	FAM-TGAACCCGGGAGGCAGAGC
REVERSE	D17S655	GTGTGTGAGGCTGGTAGAACATG
FORWARD	D17S1796	GCTGATGTCAGACTGAGC
REVERSE	D17S1796	FAM-CACAGATTAATAGAGATATTTTCCC
FORWARD	D17S2159	HEX-ACATTACCCCATGTCTTCATCC
REVERSE	D17S2159	TCTATTGTTTGTGTTAGCCTTACCC
FORWARD	D17S1783	HEX-ATGTGTACAGAGCCGCAAAG
REVERSE	D17S1783	CATGATTCAGCCACGGACTC

**EXAMPLE 6: INTEGRATED MOLECULAR PATHOLOGY ANALYSIS OF
UTERINE CERVIX LIQUID CYTOLOGY ANALYSIS FOR CUMULATIVE
ACQUIRED MUTATIONAL CHANGES**

Methods:

Residual liquid cytology samples were taken from 31 women with varying degrees of uterine cervix dysplasia and were evaluated using the advances described in this application. In each instance, the microscopic diagnosis obtained by reviewing the cytology constituted one form of diagnosis. This diagnosis however is subjective and may not accurately reflect the true status of the presence and degree of dysplasia in any particular patient. The residual liquid cytology fluid containing cells from the cervix scraping was processed as described above. The extracted DNA was evaluated for cumulative mutational change. The presence of detectable mutations provided support for the existence of true dysplasia. The higher the cumulative amount of detectable mutation provided a measure of the degree of dysplasia.

Results:

Patients were classified into seven groups according to microscopic appearance of the cytology smears (Table 35). Six patients with no microscopic evidence of dysplasia were found not to manifest any detectable mutations using a broad panel of allelic imbalance makers associated with cervical dysplasia. One of five patients with atypical squamous cells of undetermined significance (ASC-US), were determined to have two detectable mutations confirming the presence of true dysplasia. Three of four patients with atypical squamous cells, cannot exclude high grade dysplasia displayed acquired mutations. Two of the three patients showed a high level of acquired mutations supporting true high grade dysplasia. The remaining patient showed a low level of acquired mutations in keeping with true low grade dysplasia. One patient lacked evidence of detectable mutations which supported that the morphologic changes likely did not reflect true dysplasia.

As shown in Table 35, the remaining morphologic categories of low grade dysplasia (LSIL), high grade dysplasia (HSIL), and squamous cell carcinoma (SCC) demonstrated acquired mutations which in general correlated with microscopic classification. However, except for squamous cell carcinoma, the dysplasia groups contained individual patients that varied in cumulative mutations. In the low grade category, there were patients without mutations as well as those with abundant mutations. In the high grade category, there were patients with mutation accumulation more in keeping with low grade dysplasia.

The extent of mutation accumulation could be quantified using two indices termed FAL_{MARKER} and FAL_{LOCUS} . The FAL_{MARKER} represents the total number of mutated markers divided by the total number of informative markers. This index assumes that each marker carries equal weight and that no two markers are linked to each other in mutational change. The FAL_{LOCUS} represents the total number of mutated genomic loci divided by the total number of informative genomic loci. As there are two markers for each genomic loci, the likelihood of informativeness for each locus is higher than for each marker alone. This index assumes that both markers for each locus are linked and report similar mutational data. Both indices provide support that morphologic classification likely includes varying states of dysplasia, and that classification based on cumulative amount of mutational change makes for a better system of assessment of dysplasia status. The analysis supports the efficacy of genotyping liquid cytology specimens (Table 34).

Pr#	MICRO DX	1p36	3p14	3p14	3p25	4p15	4p15	4q34	4q34	6p21	6p21	9p21	9p21	10q23	10q23	11q23	11q23	17p13	17p13	18q25	18q25	FAL MARKER	FAL LODSCORE
1	neg	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NI	NO LOH	0 18	0 11
4		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	0 17	0 11
6		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	0 15	0 11
13		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	0 16	0 11
21		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	0 16	0 11
27		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	0 16	0 11
3	ASC- US	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	0 15	0 11
9		NI LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	0 17	0 11
10		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	0 17	0 10
17		NI LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	2 17	2 11
24		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NI	NO LOH	0 15	0 10
8	ASC-H	NI LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	4 16	3 11
18		NI LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	2 16	2 11
23		NI LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	3 15	3 10
30		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	1 19	1 11
2	LSIL CIN-1	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	3 18	2 11
12		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	0 18	0 11
19		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	6 19	5 11
22		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	0 18	0 11
29		NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	2 19	2 10

PH#	MICRO DX	1p36	1p36	3p14	3p14	3p25	3p25	4p15	4p15	4q34	4q34	6p21	6p21	9p21	9p21	10q23	10q23	11q23	11q23	17p13	17p13	18q25	18q25	FAL MARKER	FAL=
11	HSIL CIN-2	NI	NO	NI	NI	45%	NI	NO	NO	NO	NO	LOH	LOH	NO	LOH	NO	LOH	77%	NI	NO	LOH	NO	LOH	7 13	6 10
16		NO	LOH	NI	LOH	LOH	NO	LOH	LOH	LOH	LOH	LOH	LOH	NO	LOH	NI	LOH	NI	LOH	NI	LOH	NO	LOH	0 17	0 11
25		61%	NO	NO	LOH	LOH	NO	LOH	LOH	LOH	LOH	LOH	LOH	NO	LOH	NO	LOH	95%	LOH	LOH	LOH	LOH	LOH	8 18	5 9
28		NI	NO	LOH	LOH	NI	LOH	LOH	LOH	LOH	LOH	LOH	LOH	NO	LOH	NI	LOH	NI	LOH	NI	LOH	NO	LOH	2 16	2 11
5	HSIL CIN-3	94%	NO	LOH	NI	NI	NO	LOH	LOH	LOH	LOH	LOH	LOH	NO	LOH	NO	LOH	99%	NI	NO	LOH	NO	LOH	7 16	4 11
14		NI	NO	LOH	NI	LOH	NO	LOH	LOH	LOH	LOH	LOH	LOH	NO	LOH	NO	LOH	NI	LOH	LOH	LOH	NO	LOH	4 13	4 10
20		NI	NO	LOH	NI	LOH	NO	LOH	LOH	LOH	LOH	LOH	LOH	NO	LOH	NO	LOH	NI	LOH	LOH	LOH	NO	LOH	6 16	4 11
26		48%	NI	LOH	NO	LOH	NI	LOH	LOH	LOH	LOH	LOH	LOH	NO	LOH	NO	LOH	93%	LOH	NI	LOH	NO	LOH	5 16	4 10
7	SCCA	NI	47%	LOH	NI	LOH	NO	LOH	LOH	LOH	LOH	LOH	LOH	NO	LOH	NO	LOH	62%	NI	NO	LOH	NI	LOH	9 15	7 11
15		NI	NO	LOH	NI	LOH	NO	LOH	LOH	LOH	LOH	LOH	LOH	NO	LOH	NO	LOH	99%	LOH	LOH	LOH	LOH	LOH	7 17	5 11
31		NI	LOH	NI	LOH	NI	LOH	LOH	LOH	LOH	LOH	LOH	LOH	NO	LOH	NO	LOH	83%	LOH	LOH	LOH	LOH	LOH	6 13	5 11

The two allele copies are distinguished by dark and light shading representing the relatively deficient allele.

Proportion of mutated cells is calculated from the degree of allele imbalance (ratio of allele peaks) and expressed as a percentage when significant allelic imbalance is present.

EXAMPLE 7: ANALYSIS OF PANCREATIC CYST FLUID FOR THE PRESENCE OF FOCAL MALIGNANT TRANSFORMATION

Methods: The aspirated pancreatic cyst fluid from a 60 year old woman was submitted for diagnosis. The lesion consisted in part of a cystic abnormality that was in direct continuity with a solid, mass-like component. At issue was what the nature of the cystic and solid abnormalities were and what relationship, in any, was their between the two alterations in the pancreas.

Cytology evaluation of the cyst fluid led to a diagnosis of "insufficient cellular material for diagnosis; no cytology diagnosis possible". The cyst fluid was handled for detection of mutational changes in the manner described above using a marker panel consisting of: D1S1193, D1S407, D3S2303, D3S1539, D5S592, D5S615, D9S254, D9S251, D10S1173, D10S520, D17S974, D17S1289, D17S1161, D18S814, D21S1244, D22S532, and K-RAS-2. DNA was extracted as described above, and the amount found to be elevated at a level of 25 ng/μL. The quality was good with detectable product seen at 24 cycles for the k-ras-2 first exon. The results for allelic imbalance are described below.

Due to the presence of a solid component in direct contact with the cyst fluid, long length PCR was performed to find evidence for or against the presence of further neoplastic progression with malignant transformation. Two k-ras-2 PCR reactions were performed designed to produce a product that was approximately 650 bases and 450 bases in length. This was performed in order to selectively enhance the amplification of the most intact, best quality DNA. The product of the reaction was designed to be sequenced with an internal primer so that the presence or absence of mutation acquisition could be determined (FIG. 8). The reaction was performed quantitatively (qPCR, Icycler, Biorad).

Results. Two allelic imbalance mutations were detected involving D9S251 at 9p21 and D17S974 at 17p13. The ratio of peak heights was calculated for each allelic imbalance marker and the value compared to the average set of values generated from normal specimens for each unique combination of markers. An example of this data in tabular form is shown in Tables 9-12. When specimen values fell outside the normal range defined as two standard deviations above or below the mean, the sample was considered mutated for that markers. The percentage of mutated cells based on the ratio of peak heights was calculated using the conversion formulae:

Percentage % mutated cells = $[1 \text{ minus } (N_{\text{AVERAGE}} / \text{Value})] \times 100\%$ if the ratio of peak heights for the sample of interest is greater than N_{AVERAGE}

Percentage % mutated cells = $[\text{Value} / N_{\text{AVERAGE}}] \times 100\%$ if the ratio of peak heights for the sample of interest is less than N_{AVERAGE}

Where N_{AVERAGE} represents the average for all normal sample ratio of polymorphic peak heights for that individual allele pairing. The percentage of mutated cells based on the ratio of peak heights for these two mutated markers was 83% and 69% respectively using the conversion formulae.

The results for long length PCR amplification of k-ras-2 is shown in FIG. 8 together with several control reactions. The lesional sample in this case for the shorter 450 base pair product and the longer 650 base pair product was samples EO1 and DO1 respectively (FIG. 8, bottom tabular data).

Note that more starting template was present for the 450 base pair length product (Ct of 26 cycles) than for the 650 base pair product (Ct of 30.8 cycles). Both of these products were sequenced and showed evidence of codon 12 k-ras-2 point mutation in which aspartic acid substituted for the normal glycine in one allele. Given that no mutation was detected when the amplicon was approximately 150 bases long (Table 24), this provides support that the solid component in fact represents malignant transformation. Control reactions were run at the same time (FIG. 8). A sample of poor quality DNA was run and showed no amplification of the 450 and 650 base pair length product (AO1 and FO1 respectively).

Although the materials, methods, and kits have been described in detail with reference to examples above, it is understood that various modifications can be made without departing from the spirit of the disclosed embodiments, and would be readily known to the skilled artisan.

The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the embodiments are not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates, which may need to be independently confirmed.

This application claims priority to the following U.S. Provisional Application Nos. 60/620,926 filed October 22, 2004; 60/631,240 filed November 29, 2004; 60/644,568 filed January 19, 2005; 60/679,968 filed May 12, 2005; and 60/679,969 filed May 12, 2005, all of which are herein incorporated by reference in their entirety for all purposes. Applicants also incorporated herein in their entirety the U.S. Application entitled "Enhanced Amplifiability of Minute Fixative-Treated Tissue Samples, Minute Stained Cytology Samples, and Other Minute Sources of DNA" filed October 24, 2005, and "Dynamic Genomic Deletion Expansion and Formulation of Molecular Marker Panels for Integrated Molecular Pathology

Diagnosis and Characterization of Tissue, Cellular Fluid and Pure Fluid Specimens” filed October 24, 2005.

All references cited above are incorporated herein in their entirety for all purposes.

I claim:

1. A method for characterizing a biological fluid sample from a patient for evidence of cancer comprising:

(a) performing molecular analysis of DNA from the biological fluid sample from a patient comprising:

(i) performing an optical density analysis of the biological fluid sample to determine DNA quantity;

(ii) performing a quantitative PCR analysis of the biological fluid sample to determine DNA quality;

(iii) performing competitive template PCR to determine DNA quality of DNA in the biological fluid sample; and

(b) performing mutation analysis of the DNA of the biological fluid sample comprising:

(i) determining the presence of mutations in a tumor suppressor gene, and/or the presence of a cancer related genetic marker;

(ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy;

(iii) determining point mutations in the K-ras oncogene and/or point mutations in at least one other cancer-associated gene;

(iv) determining other structural alterations in DNA;

(v) determining the percentage of mutated DNA from steps (b) (ii) and (b) (iii); and

(vi) determining the specific temporal sequence of mutation accumulation based on step (v), and

(c) determining whether the biological fluid sample has evidence of a cancer, a dysplasia, a pre-cancerous state, or a non-neoplastic condition based on steps (a) to (b).

2. The method of claim 1, further determining whether a focal site of neoplastic progression of a neoplastic lesion is more advanced than other portions of the lesion.

3. The method of claim 1, further comprising confirming the results of steps (a) and (b) by comparing said results with a pathologic analysis of resected tissue obtained from the patient.
4. The method of claim 1, further comprising analyzing the carcinoembryonic antigen (CEA) level of the biological fluid sample.
5. The method of claim 1, wherein the DNA in the biological fluid sample is free-floating, or free and adherent to the surface of cells or tissue constituents of the source of the biological fluid sample.
6. The method of claim 5, wherein the source is a cyst.
7. The method of claim 1, wherein the biological fluid sample is a liquid cytology sample.
8. The method of claim 1, wherein the cycles of quantitative PCR performed in step (b) (ii) is greater than a threshold unique for that specific type of DNA.
9. The method of claim 1, wherein the patient is a mammal.
10. The method of claim 9, wherein the patient is a human.
11. The method of claim 1, wherein the cancer is selected from the group consisting of carcinoma, an epithelial malignancy, a sarcoma, a mesenchymal malignancy, a breast cancer, a pancreatic cancer, an urinary tract cancer, a cervical cancer, a lymphohematopoietic cancer, a neuroepithelial cancer, a central nervous system cancer.
12. The method of claim 11, wherein the central nervous system cancer is a glioma.
13. The method of claim 1, wherein the pre-cancerous state is a mucinous cystadenoma, leukoplakia, a serous cystadenoma, a colon polyp, a mesenchymal precancerous lesions, a neuroglial precancerous lesion, or a lymphohematopoietic precancerous condition.

14. The method of claim 1, wherein the non-neoplastic condition is pancreatitis, a pancreatic pseudocyst, a mesothelial cyst of the pancreas, a lymphoepithelial cyst of the pancreas, an ischemic necrosis of the pancreas, mastitis, a non-neoplastic mesenchymal lesion, a neuroglial non-neoplastic lesion, or a lymphohematopoietic condition.

15. The method of claim 1, wherein the other structural alterations in DNA are selected from the group consisting of: gene amplification, gene translocation, gene rearrangement, and epigenetic modification of DNA by DNA methylation.

16. A method for diagnosing and/or determining the prognosis of a cancer, a dysplasia, a pre-cancerous state, or a non-neoplastic condition in a patient comprising:

(a) performing molecular analysis of DNA from a biological fluid sample of the patient comprising:

(i) performing an optical density analysis of the DNA in the biological fluid sample to determine DNA quantity;

(ii) performing a quantitative PCR analysis of the DNA in the biological fluid sample to determine DNA quality; and

(iii) performing competitive template PCR of the DNA in the biological fluid sample to determine DNA quality; and

(b) performing mutation analysis of the DNA in the biological fluid sample comprising:

(i) determining the presence of one or more mutations in a tumor suppressor gene, and/or the presence of a cancer related genetic marker;

(ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy;

(iii) determining point mutations in the K-ras oncogene and/or point mutations in at least one other cancer-associated gene;

(iv) determining other structural alterations in DNA;

(v) determining the percentage of mutated DNA from steps (b) (ii) and (b) (iii); and

(vi) determining a temporal sequence of mutation accumulation based on step (v); and

(c) diagnosing and/or determining the prognosis of the cancer or dysplasia, the pre-cancerous state, or the non-neoplastic condition in the patient in need thereof based on the results of steps (a) and (b).

17. The method of claim 16, further determining whether a focal site of neoplastic progression of a neoplastic lesion is more advanced than other portions of the lesion.

18. The method of claim 16, wherein the cancer or dysplasia is selected from the group consisting of: a carcinoma, an epithelial malignancy, a sarcoma, a mesenchymal malignancy, a breast cancer, a pancreatic cancer, a cervical cancer, an urinary tract cancer, a lymphohematopoietic cancer, a neuroepithelial cancer, and a central nervous system cancer.

19. The method of claim 16, wherein the pre-cancerous state is an epithelial precancerous state selected from the group consisting of a mucinous cystadenoma, leukoplakia, a serous cystadenoma, a colon polyp, a mesenchymal precancerous lesion, a neuroglial precancerous lesion, or a lymphohematopoietic precancerous condition.

20. The method of claim 16, wherein the non-neoplastic condition is pancreatitis, a pancreatic pseudocyst, a mesothelial cyst of the pancreas, a lymphoepithelial cyst of the pancreas, an ischemic necrosis of the pancreas, mastitis, a non-neoplastic mesenchymal lesion, a non-neoplastic neuroglial lesion, or a lymphohematopoietic non-neoplastic condition.

21. A method for determining a course of treatment for a patient having a cancer or dysplasia, a pre-cancerous state, or a non-neoplastic condition comprising:

(a) performing molecular analysis of DNA from a biological fluid sample from the patient comprising:

(i) performing an optical density analysis of the DNA in the biological fluid sample to determine DNA quantity;

(ii) performing a quantitative PCR analysis of the DNA in the biological fluid sample to determine DNA quality; and

(iii) performing competitive template PCR of the DNA in the biological fluid sample to determine DNA quality; and

(b) performing mutation analysis of the DNA in the biological fluid sample comprising:

- (i) determining the presence of mutations in a tumor suppressor gene and/or the presence of a cancer related genetic marker;
- (ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy;
- (iii) determining point mutations in the K-ras oncogene and/or point mutations in at least one other cancer-associated gene;
- (iv) determining other structural alterations in DNA;
- (v) determining the percentage of mutated DNA from steps (b) (ii) and (b) (iii); and
- (vi) determining the specific temporal sequence of mutation accumulation based on step (v); and
- (c) determining whether the patient has cancer, dysplasia, a pre-cancerous state, or a non-neoplastic condition, and determining a course of treatment for said cancer, dysplasia, precancerous state, or non-neoplastic condition.

22. The method of claim 21, further determining whether a focal site of neoplastic progression of a neoplastic lesion is more advanced than other portions of the lesion.

23. The method of claim 21, wherein the cancer or dysplasia is a carcinoma, an epithelial malignancy, a sarcoma, a mesenchymal malignancy, a breast cancer, a pancreatic cancer, a lymphohematopoietic cancer, a neuroepithelial cancer, or a central nervous system cancer.

24. The method of claim 21, wherein the pre-cancerous state is an epithelial precancerous state selected from the group consisting of: a mucinous cystadenoma, a leukoplakia, a serous cystadenoma, a colon polyp, a mesenchymal precancerous lesion, a neuroglial precancerous lesion, or a lymphohematopoietic condition.

25. The method of claim 21, wherein the non-neoplastic condition is an epithelial non-neoplastic lesion selected from the group consisting of: pancreatitis, a pancreatic pseudocyst, a mesothelial cyst of the pancreas, a lymphoepithelial cyst of the pancreas, an ischemic necrosis of the pancreas, mastitis, a non-neoplastic mesenchymal

lesion, a neuroglial and non-neoplastic lesion, and a non-neoplastic lymphohematopoietic condition.

26. A kit for determining the diagnosis and/or prognosis of a patient comprising:

(a) a device for performing molecular analysis of DNA in a biological fluid sample from the patient;

(b) a device for performing mutation analysis of the DNA in the biological fluid sample;

(c) reagents for performing the molecular analysis and the mutation analysis; and

(d) optionally a device for recording data obtained from the devices of (a) and (b).

27. The kit of claim 26, wherein the device for recording data records patient age, patient sex, patient medical history, prior cancer history of patient, patient weight, patient family medical history, diagnosis determined from the mutation analysis of the DNA and molecular analysis of the DNA, temporal sequence of mutation accumulation, and proposed treatment based on determined diagnosis.

28. A method for determining and characterizing a breast anomaly in a patient comprising:

(a) performing a molecular analysis of DNA from a biological sample of patient breast tissue comprising:

(i) performing optical density analysis of the biological sample to determine DNA quantity;

(ii) performing quantitative PCR analysis of the aspirate to determine DNA quality;

(iii) performing competitive template PCR to determine DNA quality; and

(b) performing mutation analysis of the DNA of the biological sample comprising:

(i) determining the presence of mutations in a tumor suppressor gene and/or the presence of a cancer related genetic marker in the DNA of the biological sample;

(ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy;

- (iii) determining point mutations in the K-ras oncogene and/or point mutations in at least one other cancer-associated gene;
 - (iv) determining copy number alterations such as homozygous deletion and oncogene amplification;
 - (v) determining other structural alterations in DNA;
 - (vi) determining the percentage of mutated DNA from steps (b)(ii) and (b)(iii); and
 - (vii) determining a temporal sequence of mutation accumulation based on step (v); and
- (c) assessing the data from steps (a) and (b) to determine the presence of a breast anomaly and characterize the type of breast anomaly.

29. The method of claim 28, further determining whether a focal site of neoplastic progression of a neoplastic lesion is more advanced than other portions of the lesion.

30. The method of claim 1, wherein the biological sample is an aspirate from a breast cyst.

31. The method of claim 1, wherein the breast anomaly is a breast cancer, a breast dysplasia, a pre-cancerous breast condition, or a non-neoplastic condition.

32. The method of claim 31, wherein the breast cancer is selected from the group consisting of invasive ductal carcinoma, invasive lobular carcinoma, breast sarcoma, ductal adenocarcinoma, breast acinar cell carcinoma, metastatic cancer involving the breast, recurrent breast cancer, cystosarcoma phyllodes, Paget's disease, and metastatic breast cancer.

33. The method of claim 32, wherein the metastatic breast cancer is an axillary metastasis with no apparent primary tumor or an axillary adenopathy of a metastatic adenocarcinoma.

34. The method of claim 31, wherein the breast dysplasia is atypical ductal hyperplasia or atypical lobular hyperplasia.

35. The method of claim 31, wherein the pre-cancerous breast condition is selected from the group consisting of mucinous cystadenoma, serous cystadenoma,

mucinous duct ectasia, intraductal papillary mucinous neoplasm, breast intraepithelial neoplasia, and a papilloma of the breast.

36. The method of claim 31, wherein the non-neoplastic condition is selected from the group consisting of breast pseudocyst, fibrocystic disease of the breast, fibroadenoma, inflammatory mastitis, soft tissue trauma, lymphedema, and mesothelial cyst.

37. A method of analyzing a liquid cytology sample for the presence of a mutational change in a patient comprising:

(a) performing a molecular analysis of DNA from a liquid cytology sample of the patient comprising:

(i) performing optical density analysis of the liquid cytology sample to determine DNA quantity;

(ii) performing quantitative PCR analysis of the liquid cytology sample to determine DNA quality;

(iii) performing competitive template PCR to determine DNA quality; and

(b) performing mutation analysis of the DNA of the liquid cytology sample comprising:

(i) determining the presence of mutations in a tumor suppressor gene and/or the presence of a cancer related genetic marker in the DNA of the liquid cytology sample;

(ii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy;

(iii) determining point mutations in a K-ras oncogene and/or point mutations in at least one other cancer-associated gene;

(iv) determining copy number alterations such as homozygous deletion and oncogene amplification;

(v) determining other structural alterations in DNA;

(vi) determining the percentage of mutated DNA from steps (c)(ii) and (c)(iii); and

(vii) determining a temporal sequence of mutation accumulation based on step (v); and

(c) assessing the data from steps (a) and (b) to determine the presence of a mutational change in the patient.

38. The method of claim 37, further determining whether a focal site of neoplastic progression of a neoplastic lesion is more advanced than other portions of the lesion.

39. A method is for analyzing a fluid biological sample, with or without a cell, for detecting focal site of neoplasia progression in a patient with or without a large component of neoplastic or normal tissue comprising:

(a) performing a molecular analysis of DNA on a liquid biological sample of the patient comprising:

(i) performing optical density analysis of the biological sample to determine DNA quantity;

(ii) performing quantitative PCR analysis of the aspirate to determine DNA quality;

(iii) performing competitive template PCR to determine DNA quality; and

(b) performing mutation analysis of the DNA of the liquid biological sample comprising:

(i) using DNA primers to amplify an amplicon of at least 300 nucleotides to 1500 nucleotides, thereby amplifying DNA from more actively replicating tissue sites and excluding degraded DNA from amplification;

(ii) determining mutation presence in a tumor suppressor gene and/or mutation presence in a cancer related genetic marker in the DNA of the liquid biological sample;

(iii) determining tumor suppressor gene loss of heterozygosity (LOH) by analyzing polymorphic microsatellites or other polymorphic markers linked to tumor suppressor genes with respect to their allelic balances, wherein both alleles of each polymorphic microsatellite or other polymorphic marker can be distinguished, thereby distinguishing mutational and/or structural alterations of each allelic copy;

(iv) determining point mutations in the K-ras oncogene and/or point mutations in at least one other cancer-associated gene;

(v) determining copy number alterations such as homozygous deletion and oncogene amplification;

(vi) determining other structural alterations in the DNA of the liquid biological sample;

(vii) determining the percentage of mutated DNA from steps (b)(ii) and (b)(iii); and

- (viii) determining a temporal sequence of mutation accumulation based on step (v); and
- (c) determining the focal site of neoplasia progression in the patient by the steps of (a) to (b).

40. A method of detecting an allelic imbalance in a patient without an internal normal, non-neoplastic specimen from the patient comprising:

- (a) performing a molecular analysis of DNA from a biological sample of the patient comprising:
 - (i) performing optical density analysis of the biological sample to determine DNA quantity;
 - (ii) performing quantitative PCR analysis of the DNA quality;
 - (iii) performing competitive template PCR to determine DNA quality; and
- (b) performing mutation analysis of the DNA of the biological sample comprising:
 - (i) using DNA primers to amplify a polymorphic region of DNA;
 - (ii) defining a combination of specific polymorphic alleles in specimens from the general population;
 - (iii) calculating quantitatively content ratio of polymorphic alleles in specimens from the general population;
 - (iv) defining a statistical average and normal distribution of the content ratio of each combination of polymorphic alleles in specimens from the general population;
 - (v) determining a threshold for designation as within normal limits;
 - (vi) defining content ratios exceeding threshold values as outside the range of content ratio variation in specimens from the general population; and
- (c) calculating a percentage of cellular DNA subject to allelic imbalance in step (vi) using the average content ratio as a normalizing factor thereby determining the allelic imbalance of the patient without an internal normal, non-neoplastic specimen.

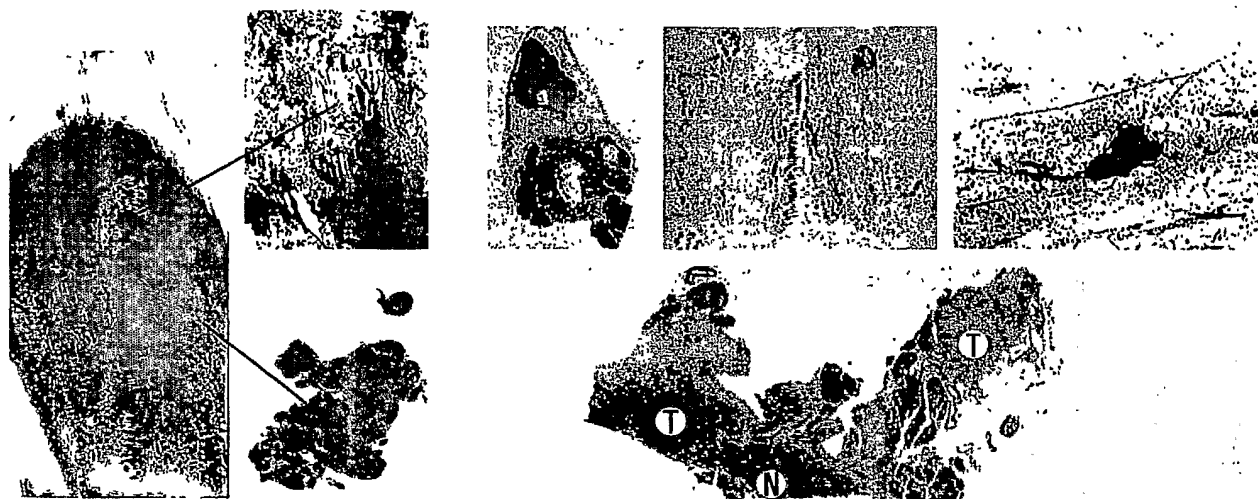
41. The method of claim 39, wherein the amplicon is at least 500 to 1500 nucleotides long.

42. The method of claim 40, wherein the threshold is at least about 95%.

43. The method of claim 40, wherein the polymorphic region of DNA amplified is a microsatellite, a minisatellite, or a single nucleotide polymorphism.

1/8

TOPOGRAPHIC GENOTYPING OF PANCREATIC CYTOLOGY



		NORMAL	TUMOR	TUMOR			
		PANCREAS	PANCREAS	PANCREAS	PANCREAS	PANCREAS	PANCREAS
		RESECTION	RESECTION	RESECTION	CYTOLOGY	CYTOLOGY	CYTOLOGY
1p36	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH
1p36	NO LOH	NO LOH	55%	59%	60%	48%	53%
3p26	NO LOH	NI	NI	NI	NI	NI	NI
3p26	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH
5q21	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH
5q21	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH
9p21	NO LOH	NI	NI	NI	NI	NI	NI
9p21	NO LOH	NO LOH	80%	85%	78%	76%	85%
10q23	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	45%	
10q23	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH
17p13	NO LOH	NI	NI	NI	NI	NI	NI
17p13	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH
17q21	NO LOH	NO LOH	76%	76%	NO LOH	76%	76%
18q25	NO LOH	NO LOH	60%	60%	60%	60%	60%
21q23	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH
KRAS	NO MUT	NO MUT	12G-R(99%)	12G-R(99%)	12G-R(92%)	12G-R(94%)	12G-R(97%)

FIG. 1

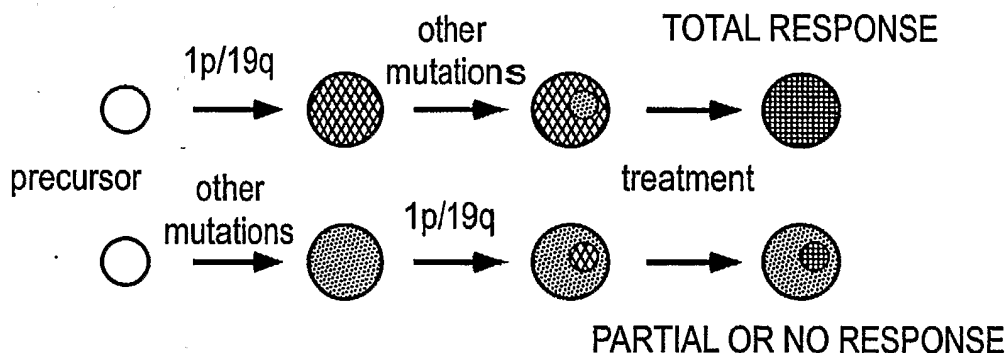
2/8

FIG. 2

TIME COURSE OF MUTATION ACQUISITION

IMPORTANCE OF TIMING OF MUTATION ACQUISITION

- The timing of mutation acquisition plays a role in treatment responsiveness (clonal expansion)
- Early acquisition of 1p/19q loss will delineate the treatment responsive subset since it will result in all glioma cells possessing this potentially favorable alteration
- Later acquisition of 1p/19q would not be as favorable since precursor cells without 1p/19q loss will be present to resist and grow in the face of treatment



3/8

FIG. 3A

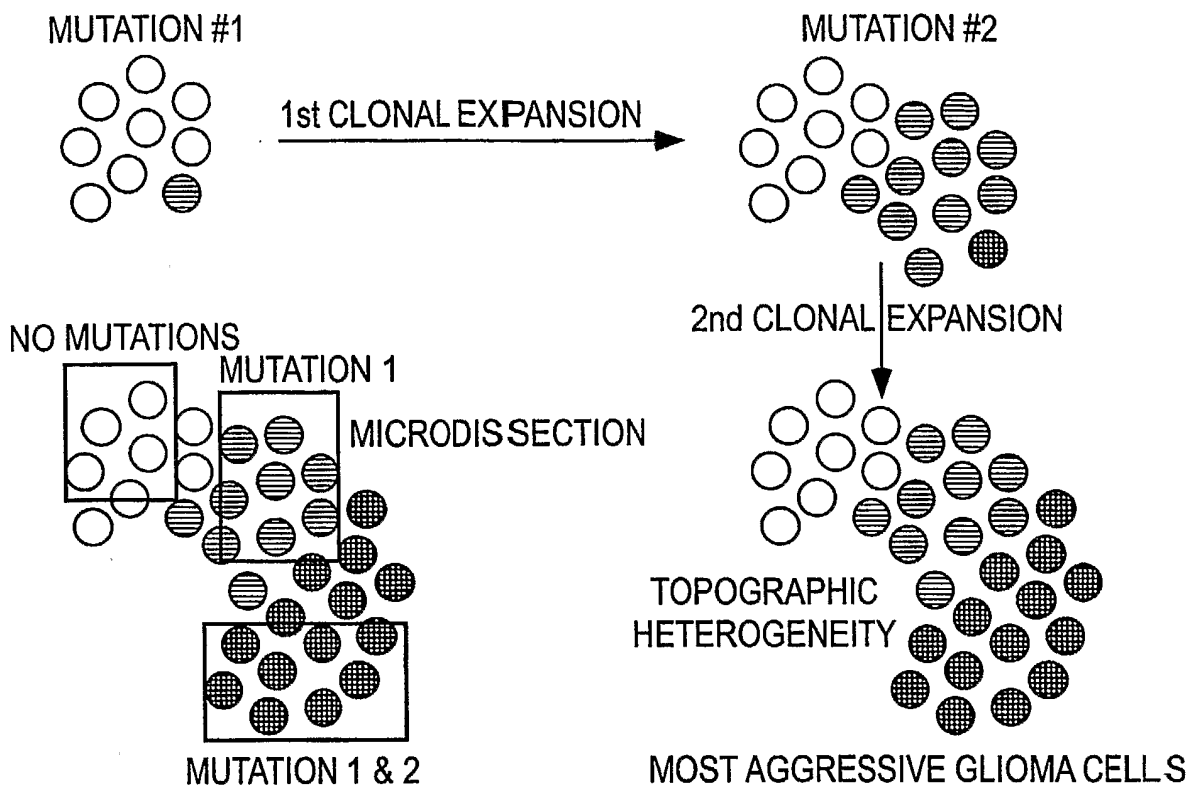
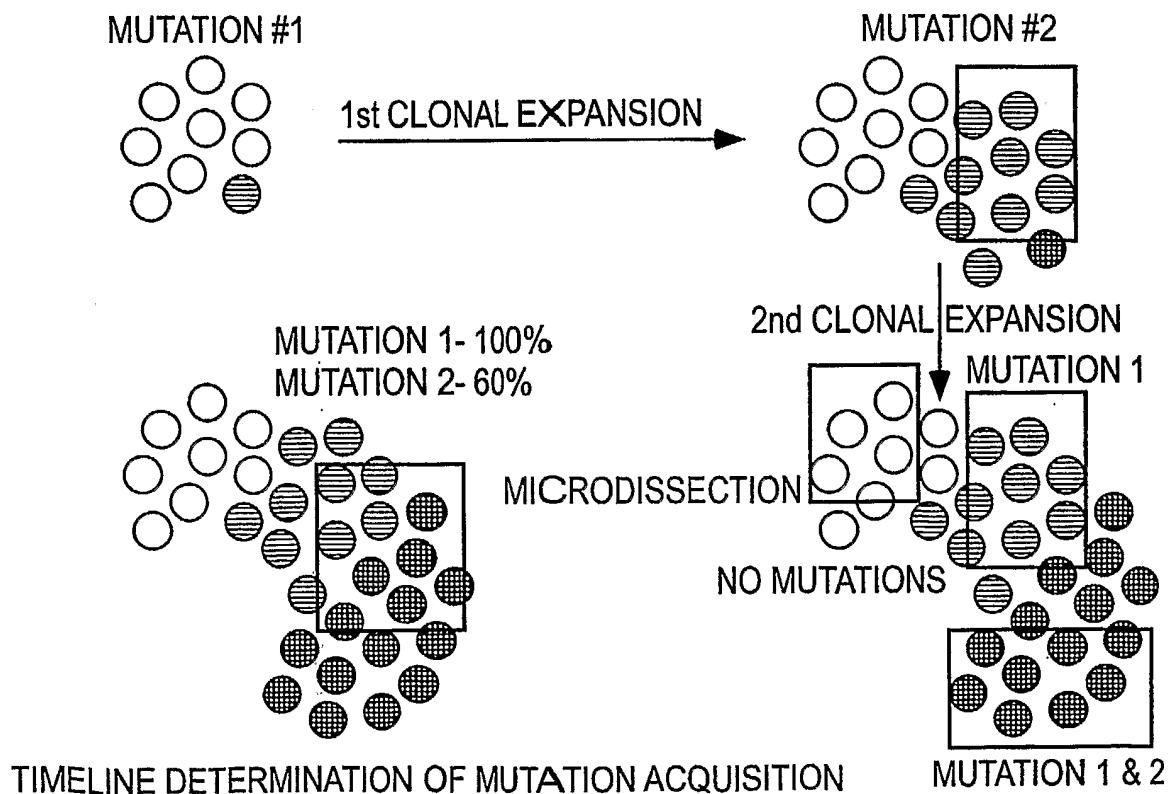


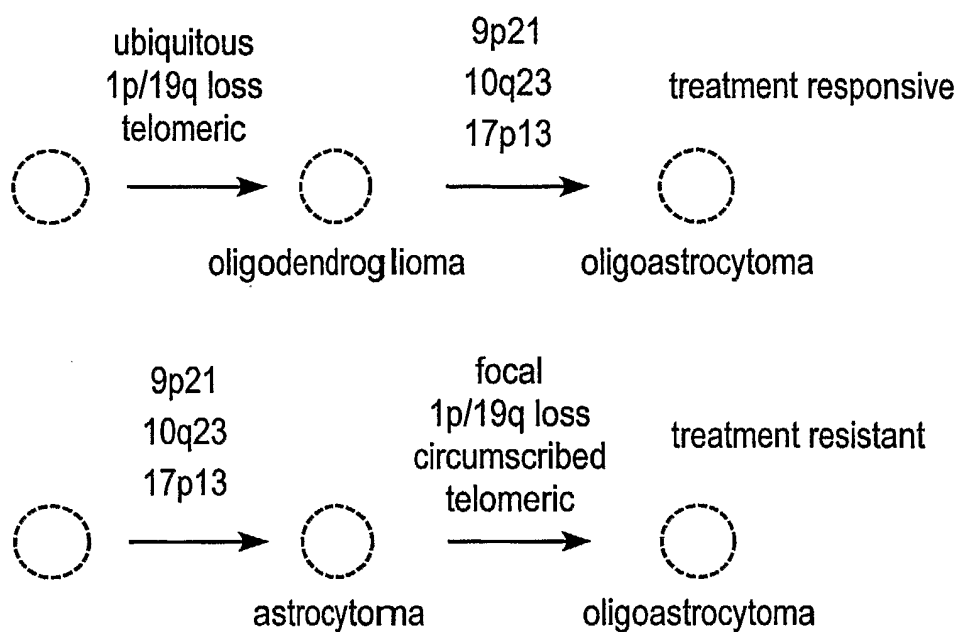
FIG. 3B



4/8

FIG. 4

INTEGRATED MOLECULAR PATHOLOGY CLASSIFICATION OF GLIOMAS



QUALITY CONTROL MEASURES TO MONITOR ANALYTIC PRECISION

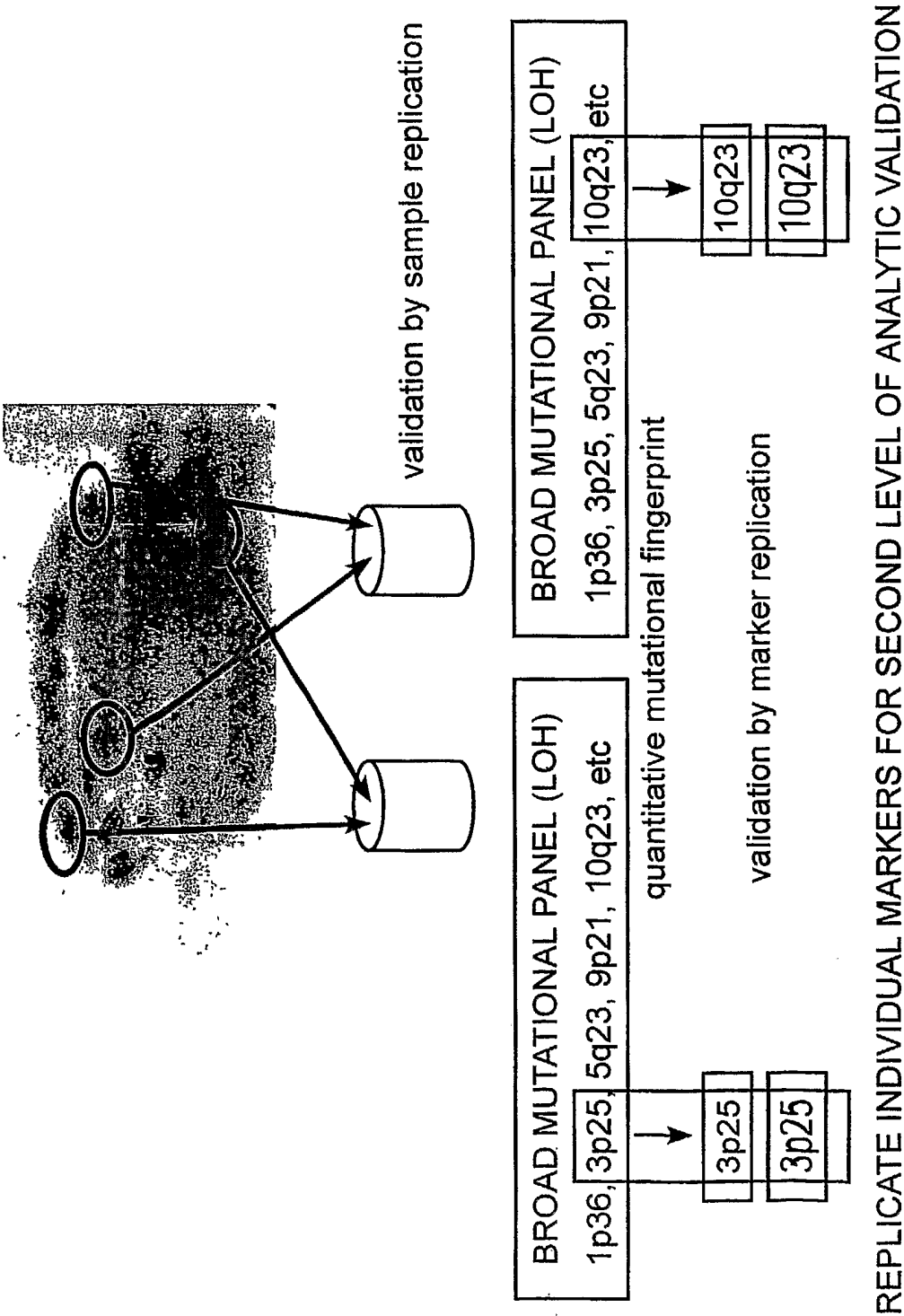


FIG. 5

FIG. 6 ELECTROPHEROGRAM SHOWING MARKER REPLICATION

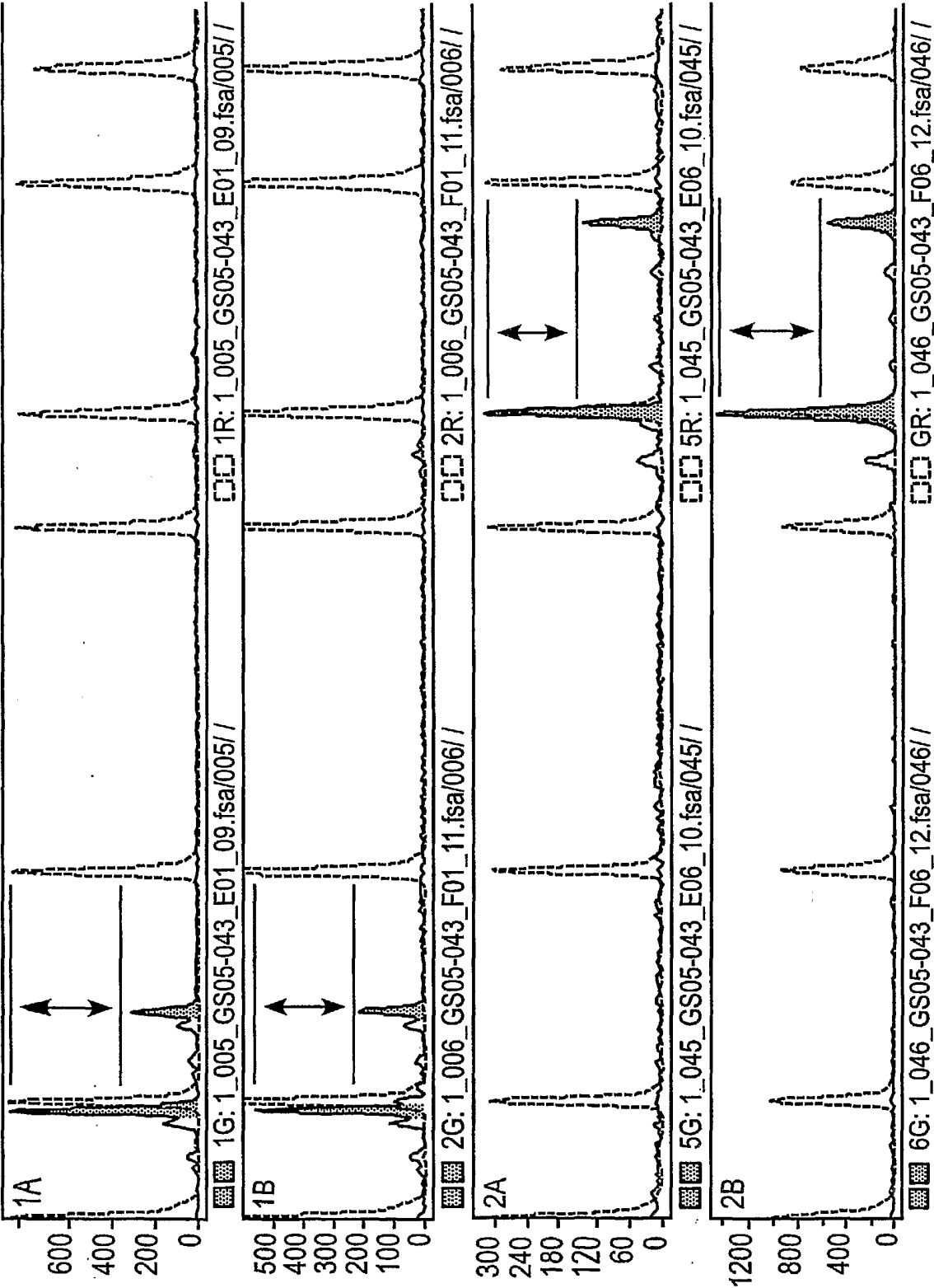
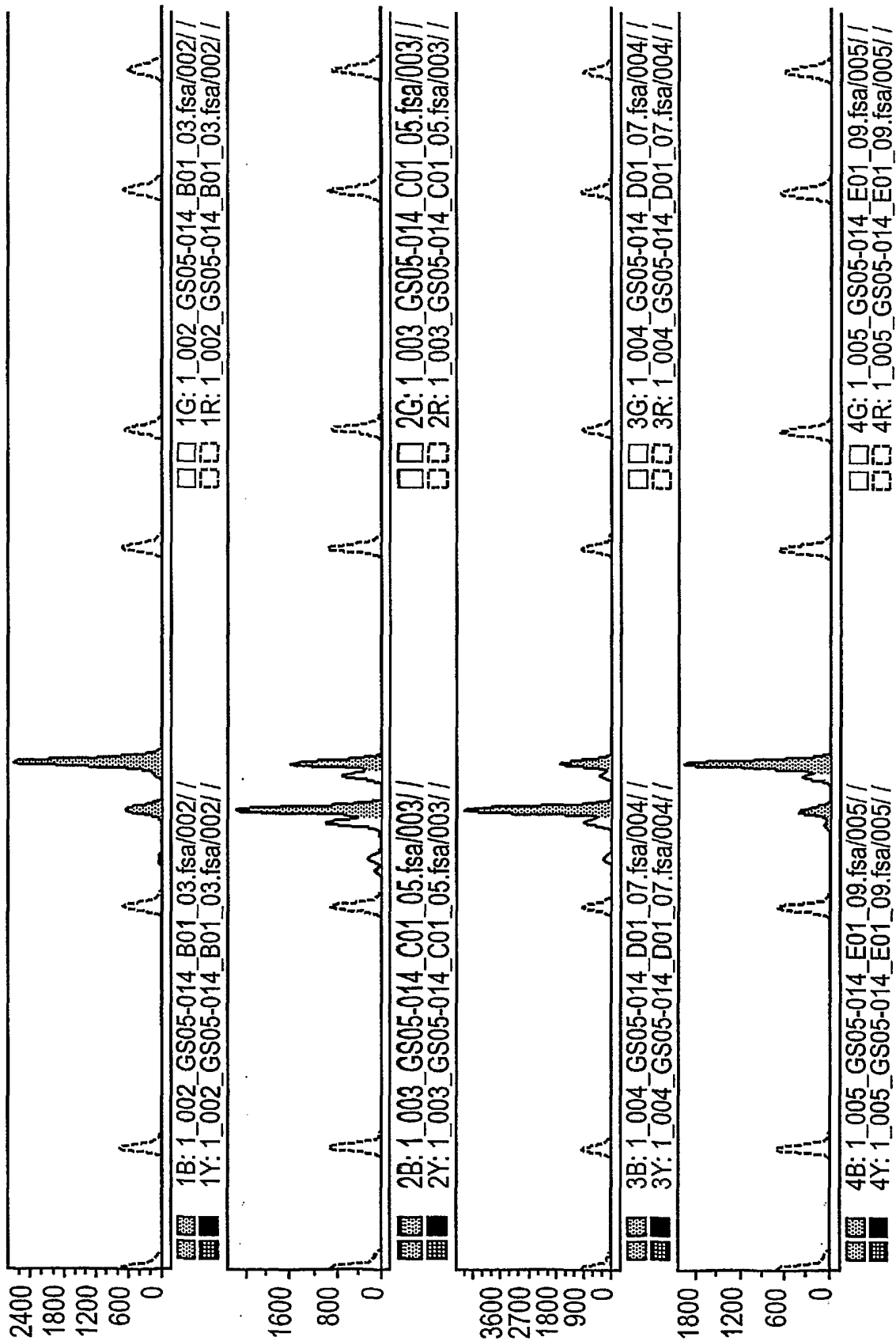


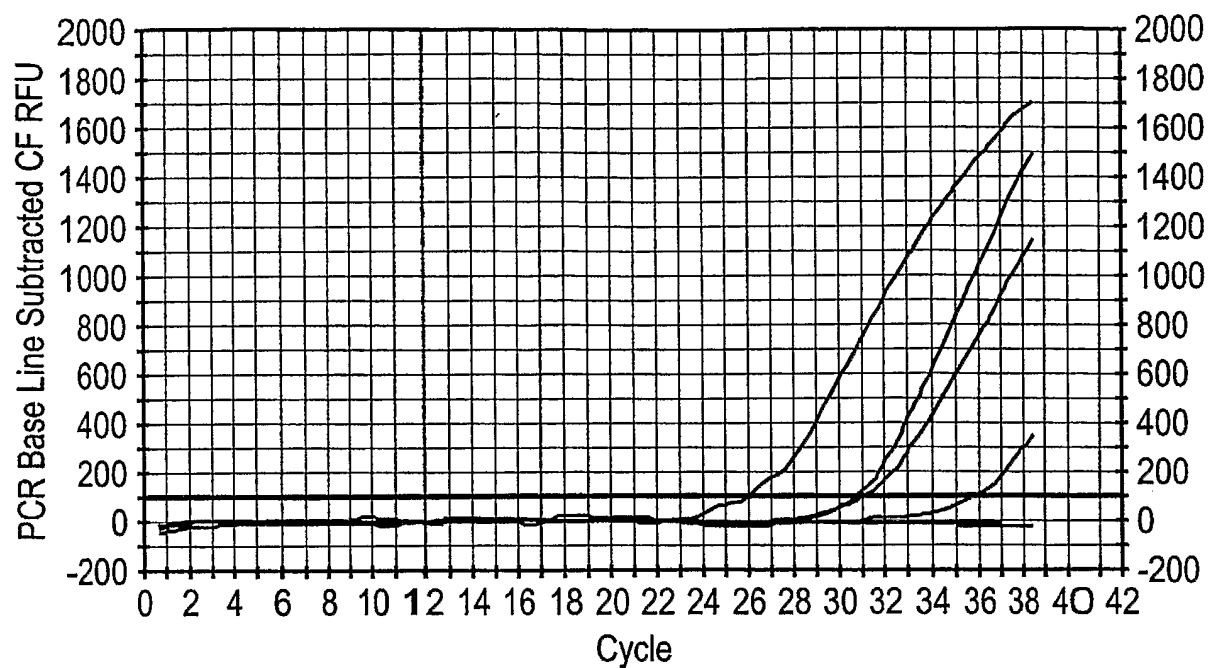
FIG. 7
ALLELIC IMBALANCE DUE TO INADEQUATE AMOUNTS OF AMPLIFIABLE
DNA PRODUCING FALSE POSITIVE IMBALANCE



Inconsistent imbalance due to insufficient DNA

8/8

FIG. 8



PCR Quantification Spreadsheet Data for SYBR-490

Well	Identifier	Ct	Setpoint
A01	04515 KRAS	N/A	
B01	04515 KRAS	35.9	
C01	045J2 KRAS	31.2	
D01	045J2 KRAS	30.8	
E01	513 CTRL KRAS	26	
F01	NTC KRAS	N/A	