Mitochondrial mutations and rearrangements as a diagnostic tool for the detection of sun exposure, prostate cancer and other cancers

This invention is related to a method of detecting skin cancer, UV exposure or UV damage in a biological sample. The invention also relates to a method for determining cumulative UV exposure in a biological sample or monitoring a subject for sun exposure and associated non-melanoma skin cancer. The invention also relates to a method for monitoring long term safety of a clinical UV phototherapy regime. The invention also relates to a use of a quantified level of mutated mitochondrial DNA (mtDNA) in a biological sample, to detect UV damage in a subject, as a biomarker for skin cancer, or as a biomarker for UV damage. The invention also relates to a A kit when used for quantifying a level of mutated mitochondrial DNA (mtDNA), the mtDNA mutation comprising a deletion spanning approximately nucleotides 547 to 4443 in the minor arc of the human mtDNA genome.

Background of the Invention

The current mega-trend in the biological sciences is the human genome project, and commercial exploitation of the data. However, there is an exceptional limitation to the use and implementation of this information as the data is not specific at the level of the individual. Incredibly the data is from only a few individuals, hardly representative of the variation present in human populations, rendering the data useful in general applications only. The staggering complexity of the human genome makes application on an individual basis impractical. To sequence completely one human nuclear genome the U.S. Department of Energy and the National Institute of Health have invested 2.5 billion dollars since 1988 (http://www.ornl.gov/hgmis/project/budget.html).

Mitochondrial Genome

The mitochondrial genome is a compact yet critical sequence of nucleic acid. The mitochondrial genome codes for enzyme subunits necessary for cellular respiration. Mitochondrial DNA, or "mtDNA", is a minuscule genome of nucleic acid at 16,569 base pairs (bp) Anderson et al., 1981; Andrews et al., 1999) in contrast to the immense nuclear genome of 3.3 billion bp. Its genetic complement is astronomically smaller than that of its nuclear cell mate (0.0005%). However, individual cells carry anywhere from 10^3 to 10^4 mitochondria depending on specific cellular function (Singh and Modica-Napolitano 2002). Communication or chemical signalling, routinely occur between the nuclear and mitochondrial genomes (Sherratt et al., 1997). Moreover, specific nuclear components are responsible for maintenance and integrity of mitochondrial sequence (Croteau et al., 1999). When these nuclear areas are rendered non-functional by nuclear rearrangements indicative of potential disease, then mutations begin to appear in mtDNA sequences. In addition, specific mitochondria may be identified for intracellular destruction by deletions prompted by somatic mutations in the mitochondrial genome. This theoretical mechanism may serve as an indication of impending disease as well. About 3,000 genes are required to make a mitochondrion, with only thirty-seven of these coded by the mitochondrial genome, indicating heavy mitochondrial dependence on nuclear loci (Naviaux, 1997).

All mitochondrial DNA (mtDNA) genomes in a given individual are identical given the clonal expansion of mitochondria within the ovum, once fertilization has occurred. The essential role

Non-Melanoma Skin Cancer

Human non-melanoma skin cancer (NMSC) is the commonest cancer in many Caucasian populations (Weinstock, 1998; Rees, 1998). The majority of these tumours are basal cell carcinoma (BCC) and squamous cell carcinoma (SCC). BCCs are locally invasive and can cause significant morbidity but rarely metastasis. SCCs show significant metastatic potential and the occurrence of multiple NMSCs in patients with immunosuppression causes significant management problems (Rees, 1998). While there are no clinically identified pre-malignant lesions for BCC, some SCCs are thought to arise from precursor lesions, namely actinic keratoses (AKs) or areas of Bowen's disease (in situ carcinoma)(Rees, 1998).

SCCs show loss of heterozygosity affecting several chromosomes which suggests the involvement of several tumour suppressor genes in their development. Interestingly, in AKs, an equal or greater degree of genetic loss is observed in these precursor lesions compared to SCCs (Rehman et al. 1994; Rehman et al. 1996). This may be important for the proposed techniques because it suggests that other mechanisms, in addition to inactivation of tumour suppressor genes, are likely to be involved in the development of SCCs.

A role for mitochondria in tumourigenesis was originally hypothesised when tumour cells were found to have an impaired respiratory system and high glycolytic activity (Shay & Werbin, 1987). Recent findings elucidating the role of mitochondria in apoptosis (Green & Reed, 1998) together with the high incidence of homoplasmic mtDNA mutations in colon cancer (Habano et al. 1998; Polyak et al. 1998, reviewed in Penta et al., 2001), primary tumours of the bladder, neck and lung (Fliss et al. 2000), and gastric tumours (Tamura et al. 1999), further support this hypothesis. Furthermore, it has been proposed that these mitochondrial mutations may affect the levels of reactive oxygen species (ROS) which have been shown to be highly mitogenic (Polyak et al. 1998; Li et al. 1997).

Previous studies by the inventors and others have shown that mutations in mtDNA and the associated mitochondrial dysfunction is an important contributor to human degenerative diseases (Birch-Machin et al. 1993; Chinnery et al. 1999; Birch-Machin et al. 2000b). This is because the mitochondrial genome is particularly susceptible to mutations due to the high amounts of ROS produced in this organelle coupled with the lack of protective histones and a low rate of mtDNA repair (Pascucci et al. 1997; Sawyer & van Houten; LeDoux et al.1999) compared to the nucleus. Indeed, the mutation rate for mtDNA is around ten times higher than that of nuclear DNA (Wallace,1994). Most of the mtDNA mutations identified in the recent human tumour studies have indicated possible exposure to ROS derived mutagens. This is important for the investigation of mtDNA mutations in NMSC because there is recent evidence for the direct involvement of UV induced ROS in the generation of mtDNA deletions in human skin cells (Berneburg et al. 1999, Lowes et al., 2002). In addition, the major determinant of NMSC in individuals without protective pigmentation or genetic predisposition is UV (Weinstock, 1998). The putative precursor lesions of

men over the age of forty-nine (49) have some cancerous prostate cells, yet only 20% to 25% of these men have a clinically significant form of prostate cancer (SpringNet – CE Connection, internet, www.springnet.com/ce/j803a.htm). Prostate cancer exhibits a wide variety of histological behaviour involving both erogenous and exogenous factors, i.e. socio-economic situations, diet, geography, hormonal imbalance, family history and genetic constitution (Konishi et al. 1997; Hayward et al. 1998).

From a risk standpoint *familial* and *hereditary* prostate cancers are not considered synonymous terms. Familial cancers refer to the incidences within a family, but are not inherited. This form accounts for up to 25% of prostate cancers (Walsh & Partin, 1997). Hereditary refers to a subtype of prostate cancer with a Mendelian inheritance of a predisposing gene(s) and accounts for approximately 9% of reported cases. A positive family history of prostate cancer for this disease suggests that these predisposing gene(s) play an important role in prostate cancer development and progression. Recently, susceptibility genes on chromosomes 1 and X have been identified as predisposing men to prostate cancer, providing greater insight into the etiology of hereditary cancer (Berthon et a]. 1998; Xu et al. 1998).

Prostate cancer prognosis mainly depends on the tumour stage and grade at diagnosis. Only localized prostate cancer can be cured by radical treatment. Standard detection still relies on digital rectal examination, PSA testing and histopathologic examination of prostatic biopsied tissues. Biopsy of a mass is used to confirm malignancy, it is not an early detection technique. Unfortunately, some early tumours are impossible to identify during rectal exams. PSA tests have a specificity of 60 to 70% and a sensitivity of 70 to 80% (personal communication, Dr. Sunil Gulavita, Northwestern Ontario Cancer Centre). A newer technique which refines diagnosis for tumours of common histologic grade is ploidy-DNA analysis employing flow cytometry (Shankey et al. 1995); however, this technique measures chromosomal changes that are only apparent in later stages of cancer development and is not sufficiently sensitive for the detection of minor alterations in DNA structure or chromosomal inversions, or reciprocal trans-locations in early cancers. The disclosed techniques focus on early detection since prognosis is heavily dependent on the stage of disease at diagnosis.

Our understanding of genetic abnormalities in prostate cancers is scanty. Research into prostate cancer has focussed on the development of knowledge in the following areas: 1) proto-oncogenes (Buttyan et al. 1987); 2) tumour suppressor genes (p53, p73, KAI1 and MMACI/PTEN; Dong et al. 1995; Cairns et al. 1997) and 3) telomere/telomerase activity in metastasis. Up-regulation of telomerase and amplification of telomeric DNA in prostate cells may provide effective markers for diagnosis. Moreover, telomeres may serve as a site for therapy (Ozen et al. 1998). A number of groups have provided evidence for a "prostate cancer gene" in the short arm of chromosome 1 (Berthon et al. 1998). More work is needed to identify the specific locus within this region. It has been suggested that this marker is only one of several possible genes predisposing men to familial prostate cancer. Other studies have shown possible marker loci on the X chromosome (Xu et al.

however, in order to obtain accurate data the age specific deletion and/or mutation rate must be determined concisely. Attributing mutations to the aging process as opposed to a particular disease at the population level is vital. This information is imperative to an understanding of how mtDNA damage accrues over time. Moreover, the consequences of these particular mutations, their frequencies, and associations in the temporal aspects of aging must be known in order to forecast and eventually slow aging at the molecular level. Researchers have not yet determined this rate, which requires evaluation of population data through maternal lines. Accordingly, there is a need for a biomarker which tracks the aging process.

Accordingly, there is a need for a simple, straightforward system of monitoring the mitochondrial genome for mutations which indicate early stage cancer, aging or other human diseases with a DNA component. There is also a need for a simple diagnostic system for sun exposure, non-melanoma skin cancer, prostate cancer, lung cancer and aging linked to defects in the mitochondrial genome. There is a need for a diagnostic system which differentiates between mutations in mtDNA which cause disease, and those which simply represent variation within and between populations.

Summary of the Invention

Aspects of the present invention are listed in the claims.

For example, there is disclosed a method for detecting a predisposition to cancer, early detection of cancer, genesis of cancer, presence of cancer, progression of cancer, absence of cancer, or sun exposure in a biological sample having mtDNA, comprising (a) providing a biological sample comprising mtDNA and (b) detecting a deletion in the mtDNA.

There is also disclosed an array comprising a plurality of nucleic acid members, and a solid substrate, wherein each of the nucleic acid members is associated with at least one deletion associated with predisposition to cancer, early detection of cancer, genesis of cancer, presence of cancer, progression of cancer, absence of cancer, sun exposure or aging, and is chosen from mitochondrial DNA, RNA transcribed from mitochondrial DNA, and cDNA, wherein each nucleic acid member has a unique position on said array and is stably associated with the solid substrate.

There is also disclosed a kit for diagnosing a predisposition to cancer, early detection of cancer, genesis of cancer, presence of cancer, progression of cancer, absence of cancer or sun exposure comprising at least one member selected from the group consisting of: solid support, the array describe above, means for holding the solid support, means for extraction of mitochondrial DNA, means for access to a database of mitochondrial DNA sequences, primers, reagents and instructions

There is also disclosed a database containing mitochondrial DNA sequences chosen from normal control sequences associated with non-disease states, sequences associated with interpopulation variations, sequences associated with intrapopulation variations, or

sequences associated with a predisposition to cancer, early detection of cancer, genesis of cancer, presence of cancer, progression of cancer, absence of cancer, sun exposure or aging.

There is also disclosed a method of detecting a deletion spanning approximately nucleotides 10744 to 14124 of the mtDNA genome, wherein said deletion is associated with prostate cancer, in a subject having mtDNA, comprising (a) providing a biological sample from the subject and (b) detecting the presence of the deletion in the mtDNA. The deletion may be in the range of 3000 to 4000 bp. The deletion may be approximately 3379 bp. The deletion may deletes all or part of base pairs between 10744 and 14124, comprising substantially genes encoding NADH dehydrogenase subunit 4L, NADH dehydrogenase subunit 4, NADH dehydrogenase subunit 5, tRNA histidine, tRNAserine2, and tRNA leucine2.

There is also disclosed a nucleic acid primer 3.4 forward comprising TAG ACT ACG TAC ATA CTA ACC CTA CTC CTA (SEQ ID NO: 139) and a nucleic acid primer 3.4 reverse comprising GAG GTA GGA TTG GTG CTG T (SEQ ID NO: 140).

There is also disclosed an array comprising a plurality of nucleic acid members, and a solid substrate, wherein one of the nucleic acid members is associated with the mtDNA deletion at approximately 10744 to 14124, wherein the nucleic acid member has a unique position on said array and is stably associated with the solid substrate.

There is also disclosed a kit for diagnosing skin cancer comprising at least one member selected from the group consisting of: a disposable chip, the array of claim 31, means for holding the disposable chip, means for extraction of mtDNA, primers, reagents and instructions

There is also disclosed a method of monitoring a person for the progression toward prostate cancer or progression of prostate cancer, in a biological sample from a subject, comprising: providing a biological sample from the subject; extracting DNA from the biological sample; detecting the absence or presence of deletions of the mtDNA; determining whether the deletions are associated with normal interpopulation or intrapopulation variations, or whether the delettions are associated with the absence or presence of a predisposition to prostate cancer, progression toward prostate cancer, prostate cancer or progression of prostate cancer, and; repeating the steps.

There is also disclosed a use of a deletion between approximately 10744 and 14124 of the mtDNA comprising all or part of NADH dehydrogenase subunit 4L, NADH dehydrogenase subunit 4, NADH dehydrogenase subunit 5, tRNA histidine, tRNAserine2, and tRNA leucine2 to detect a predisposition to prostate cancer, early detection of prostate cancer, genesis of prostate cancer, presence of prostate cancer, or progression of prostate cancer in a subject having mtDNA.

There is also disclosed a method for confirming or refuting a prostate cancer biopsy test from a biopsy sample, comprising: obtaining normal tissue from a biopsy sample;

and detecting the absence or presence of a mtDNA deletion of approximately 3379 bp in the normal tissue.

There is also disclosed a method for three-dimensional prostate tumour mapping, comprising: obtaining sextant needle biopsy samples; and detecting the absence or presence of a mtDNA deletion of approximately 3379 bp in each of the sextant samples.

There is also disclosed a method for the collection of a patient sample for use in diagnosing prostate cancer sun exposure by using a deletion spanning approximately nucleotides 10744 to 14124 of the mtDNA genome, comprising: providing a biological sample from the subject.

There is also disclosed a mitochondrial deletion spanning approximately nucleotides 10744 to 14124 of the mtDNA genome for use in diagnosing prostate cancer.

There is also disclosed a method for detecting sun exposure or non-melanoma skin cancer.

There is also disclosed a method of detecting a deletion spanning approximately nucleotides 547 to 4443 in the minor arc of the mtDNA genome, wherein said deletion is associated with skin cancer and/or sun exposure, in a subject having mtDNA, comprising: providing a biological sample from the subject; and detecting the presence of the deletion in the mtDNA. There is also disclosed is to provide a method for determining the cumulative UV exposure, in a subject having mtDNA, comprising: providing a biological sample from the subject; and detecting the presence of the deletion in the mtDNA. A similar method may also be used of monitoring the long term safety of clinical UV phototherapy regimes. The deletion may be in the range of 3500 to 4000 bp. The deletion may be approximately 3895 bp, comprising a span of mtDNA from approximately mtTF1 binding site in the D-loop to tRNA methionine. The deletion may delete all or part of base pairs between 547 to 4443 in the minor arc of the mtDNA genome, comprising substantially a 12s rRNA gene, a 16s rRNA gene, a ND1 gene and promoters for transcription of H and L strands.

There is also disclosed a nucleic acid primer L404 comprising CTT TTG GCG GTA TGC ACT TT (SEQ ID NO: 145) and a nucleic acid primer H4676 comprising GAT TAT GGA TGC GGT TGC TT (SEQ ID NO: 146).

There is also disclosed a nucleic acid probe, 3895-probe, comprising TGC TAA CCC CAT ACC CCG AAA ATG TTG G Tamra (SEQ ID NO: 153).

There is also disclosed an array comprising a plurality of nucleic acid members, and a solid substrate, wherein one of the nucleic acid members is associated with the mtDNA deletion at approximately 547 to 4443, wherein the nucleic acid member has a unique position on said array and is stably associated with the solid substrate.

There is also disclosed a kit for diagnosing skin cancer comprising a disposable chip, the array described above, means for holding the disposable chip, and means for extraction of mtDNA.

There is also disclosed a method of monitoring a person for sun exposure and non-melanoma skin cancer, in a biological sample from a subject, comprising: providing a biological sample from the subject; extracting DNA from the biological sample; detecting the absence or presence of deletions of the mtDNA; determining whether the deletions are associated with normal interpopulation or intrapopulation variations, or whether the deletions are associated with sun exposure, and; repeating the steps.

There is also disclosed a use of a deletion between approximately 547 to 4443 in the minor arc of the mtDNA genome in a subject having mtDNA to detect sun exposure or non-melanoma skin cancer.

There is also disclosed a method for the collection of a patient sample for use in diagnosing skin cancer by using a deletion spanning approximately nucleotides 547 to 4443 in the minor arc of the mtDNA genome, comprising: providing a biological sample from the subject.

There is also disclosed a mitochondrial deletion spanning approximately nucleotides 547 to 4443 of the minor arc of the mtDNA genome for use in diagnosing sun exposure or skin cancer.

There is also disclosed a use of the deletion junction sequence of SEQ ID NO: 147 to confirm the presence of the 3895bp mtDNA deletion associated with sun exposure or NMSC.

There is also disclosed a method for a sensitive detection of a rearrangement in a DNA sample, wherein the rearrangement creates a newly formed junction in the DNA sequence of a sample, comprising: providing a DNA sample comprising or suspected of comprising a rearrangement; providing a primer or a probe which spans the newly formed junction created by the rearrangement; detecting the rearrangement by amplifying or probing the junction. The rearrangement may be in mtDNA. The rearrangement may be a deletion.

Brief Description of the Figures

Figure 1 is a histogram showing the number of mutations at nucleotide position in mitochondrial DNA from patients with prostate cancer.

Figure 2 shows the first half of the non-synonymous clusters produced using Hierarchal Clustering Explorer (HCE).

Figure 3 shows the second half of the non-synonymous clusters produced using Hierarchal Clustering Explorer (HCE).

Figure 4 is a copy of Figure 3 wherein the important cluster is shaded.

Figure 5 is a schematic diagram showing the design and sequence of a primer useful for the detection of the 3.4 kb deletion.

Table 4 is a summary of mitochondrial mutations (including D-loop) from prostate needle biopsies and complete genome mutations from malignant, adjacent and distant benign prostate glands from patients with prostate cancer.

Table 5 is a list of primers used for complete mitochondrial genome amplification for formalin fixed and normal tissues from blood.

Table 6 is a list of amplification primers for use with Example 12.

Table 7 is qPCR Components of Example 12.

Table 8 shows the cycling parameters for Example 12.

Table 9 is a list of probes used in Example 14.

Detailed Description

The method disclosed can be used to diagnose diseases linked to mtDNA. The method disclosed provides for analysis of the mitochondrial genome of an individual from a biological sample, for example by amplification of the mitochondrial genome, sequencing a portion of the mitochondrial genome, preferably the entire mitochondrial genome of the individual using any known means. Denaturing high performance liquid chromatography (DHPLC) may also be used to rapidly screen many samples. DHPLC can focus on hotspots of mutations. DHPLC is more sensitive than automated sequencing in terms of detecting mutations, and can even detect 2% heteroplasmy, compared with 20-25% for ordinary sequencing. Methods for detecting lower levels of heteroplasmy (<2%) may also be developed.

As used herein, "actinic kerotoses" means proposed precursor epidermal lesion of a squamous cell carcinoma.

As used herein, "aging" refers to an accumulation of changes with time, both at the molecular and cellular levels.

As used herein, "alleles" means one of several alternative forms of a given DNA sequence occupying a specific place on a chromosome.

As used herein, "artificial neural network (ANN)" means a virtual device in which several interconnected elements process information simultaneously, adapting and learning from past patterns.

As used herein, "attaching" or "spotting" refers to a process of depositing a nucleic acid onto a solid substrate to form a nucleic acid array such that the nucleic acid is irreversibly bound to the solid substrate via covalent bonds, hydrogen bonds or ionic interactions.

As used herein, "atypical" or "abnormal" means cellular appearance which is not normal, but also does not appear to be malignant.

As used herein, "basal cell carcinoma" means a type of cancer of skin cells.

As used herein, "benign" means of no danger to health; not recurrent or progressive; not malignant.

As used herein, "mutation" encompasses any modification or change in a DNA or RNA sequence from the wild type sequence, including without limitation point mutations, transitions, insertions, transversions, translocations, deletions, inversions, duplications, recombinations or combinations thereof. The modification or change of the sequence can extend from a single base change to the addition or elimination of an entire DNA or RNA fragment.

As used herein, "mutation load" refers to an increase in mutations in mtDNA which may eventually lead to compromised function of the involved gene or the entire genome or may accumulate in non-coding regions.

As used herein, "neoplasia" means a pathological process which may result in transformation to malignant status.

As used herein, "non-involved tissue" means tissue from a part of the body which is not associated with the disease in question.

As used herein, a "non-synonymous" mutation of a polynucleotide is a mutation which results in a different encoded amino acid.

As used herein, "normal tissue" means tissue with no visible manifestations of disease as determined by histology.

As defined herein, a "nucleic acid array" refers to a plurality of unique nucleic acids attached to one surface of a solid support at a density exceeding 20 different nucleic acids/cm^2 wherein each of the nucleic acids is attached to the surface of the solid support in a non-identical preselected region. In one embodiment, the nucleic acid attached to the surface of the solid support is DNA. In a preferred embodiment, the nucleic acid attached to the surface of the solid support is cDNA. In another preferred embodiment, the nucleic acid attached to the surface of the solid support is cDNA synthesized by polymerase chain reaction (PCR). Preferably, a nucleic acid array comprises nucleic acids of at least 150 nucleotides in length. Preferably, a nucleic acid array comprises nucleic acids of less than 6,000 nucleotides in length. More preferably, a nucleic acid array comprises nucleic acids of less than 500 nucleotides in length. In one embodiment, the array comprises at least 500 different nucleic acids attached to one surface of the solid support. In another embodiment, the array comprises at least 10 different nucleic acids attached to one surface of the solid support. In yet another embodiment, the array comprises at least 10,000 different nucleic acids attached to one surface of the solid support. The term "nucleic acid", as used herein, is interchangeable with the term "polynucleotide".

As used herein, a "nucleic acid target" or "a target nucleic acid" is defined as a nucleic acid capable of binding to a nucleic acid member of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a nucleic acid target may include natural (i. e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in nucleic acid probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus,

and upper thigh pain, or painful ejaculation, but have been diagnosed as benign by examination of biopsy tissue by a qualified pathologist.

As used herein, a "synonymous" mutation is a mutation in a polynucleotide which does not have an affect on the encoded amino acid.

As used herein, "transitions" means substitution of like nitrogenous bases, pyrimidine to pyrimidine, purine to purine. A mutation in which one pyrimidine is substituted by the other, or in which one purine is substituted by the other.

As used herein, "transversions" means substitution of unlike nitrogenous bases, purine to pyrimidine, pyrimidine to purine. A mutation in which a purine is substituted or replaced by a pyrimidine or vice versa.

As used herein, "usually sun exposed skin" means skin that is usually or often exposed in an individual. For example, depending on the individual, it may include scalp, face, neck and ears.

MtDNA and diagnosis of specific diseases

Methods are provided for monitoring aging, sun exposure, and diagnosing specific diseases such as prostate cancer and non-melanoma skin cancer through comparisons of mtDNA sequences. Diagnosing diseases such as prostate cancer with mtDNA, rather than nuclear DNA has several advantages. Firstly, mtDNA, a less complex genome, is easily understood at an individual and population level, hence a large mtDNA database with normal and disease associated genomes renders individual diagnosis extremely accurate. Accordingly, variation, in relationship to disease, is understood. Secondly, mtDNA has a 10-fold higher mutation rate than nuclear DNA (Wallace 1992). Nuclear rearrangements, suggestive of preliminary disease, are rapidly communicated to mitochondria, where they appear as somatic mutations. Thirdly, mtDNA has a maternal inheritance pattern, and is essentially clonal in that all mitochondria begin with the same mtDNA sequence, hence variation from this clonal condition is easily detected. Additionally, mtDNA does not show convincing evidence of recombination, thus any alterations in sequence are a somatic event. Any one mitochondrion harboring a mutation(s) is in a sense 'recessive' as a consequence of there being many mitochondrial genomes (2-10 copies) per mitochondrion, and many mitochondria per cell (500-2,000). Moreover, mitochondrial genomes can tolerate very high levels (up to 90%) of mitochondria with damaged genomes. This happens through complementation by the remaining wild type mtDNA (Chomyn et al. 1992). However, mutated genomes have a replicative advantage over wild type genomes because they are usually smaller (Hayashi et al. 1991), hence there is clonal expansion of mutated mtDNA (Brierley et al. 1998), suggesting that unlike nuclear genes, there is little or no selection against cells harboring mtDNA mutations. Because of this elevated mutation rate, mutations and/or deletions that appear in mtDNA are maintained through the life span of the cell and may serve as a record of exposures to various mutagens. The integrity of mtDNA is maintained by nuclear repair mechanisms, and a defect at these loci has been suggested to result in an autosomal dominant disorder associated with multiple mitochondrial deletions (Zeviani et al. 1990).

Consequently, mtDNA may function as an early warning sentinel of early nuclear events related to a variety of cancers or other diseases. Finally, the mitochondrial genome can be sequenced and monitored for mutations on an individual basis.

The methods and products disclosed herein detect both heteroplasmic as well as homoplasmic mutations. In fact, heteroplasmic mutations may be key to the detection of the early genesis of disease, disorder or aging. In addition, although specific mutation sites may indicate a particular disease state, disorder or aging process, the total mutation load is also important in determining the genesis, presence and progression of a disease, a disorder or aging.

The present techniques allow for the ability to examine benign or normal tissue or bodily fluids to determine the effects of sun exposure, the genesis and/or presence of disease, disorder or aging. For example, the present invention allows for the ability to examine benign tissue or bodily fluids for the presence of pre-neoplasia, neoplasia, progression toward malignancy and malignancy.

The mitochondrial mutations detected by the methods disclosed are compared to inter and intrapopulation variations in mitochondrial DNA, and may include comparison with mitochondrial DNA from non-involved tissue from the subject, or with mitochondrial DNA from a maternal relative. It is not necessary to analyze the entire mitochondrial genome. For example, it is not necessary to sequence the entire mitochondrial genome, only a select portion of it. Accordingly, a sample of mitochondrial DNA can provide a diagnosis.

Detection of sun exposure

In one embodiment, a system for early detection of mitochondrial DNA mutations and rearrangements, for example deletions, as a result of sun exposure is provided. The particular changes, such as the common deletion, the 3895 bp deletion identified in examples 13 and 14, associated mutations, and the incidence of as yet uncharacterised deletions in mtDNA serve as reliable bio-markers of sun exposure.

Diagnosis of Skin Cancer

In a preferred embodiment, a system for early diagnosis of mtDNA changes in non-melanoma skin cancer (NMSC) and melanoma skin cancer, and their precursor lesions indicative of solid tumour development is provided. The particular changes, such as the common deletion, the 3895 bp deletion identified in examples 13 and 14, associated mutations, and the incidence of as yet uncharacterised deletions in mtDNA serve as reliable bio-markers of sun exposure, and associated skin cancer. Non-melanoma skin cancer in particular is associated with chronic life long sun exposure. Melanoma skin cancer seems to be more related to acute burning episodes. The mutation fingerprint of the entire mtDNA genome in human NMSC and its precursor lesions is determined. Thus mtDNA changes are established as an early bio-marker of human skin cancer and its precursor lesions. Denaturing HPLC can then be used to assess low levels of heteroplasmy at the sequences of

interest. This approach can also provide an insight into the development of early changes in other human tumours.

Diagnosis of prostate cancer

In another embodiment, a system for diagnosis of prostate cancer is provided.

Age related accumulation of mtDNA defects might predispose an individual to the appearance of certain clinical disorders such as prostate cancer which is prevalent in middle age and older men. In a preferred embodiment, routine prostate cancer screening takes place through mitochondrial genome sequencing from prostate massage fluid. The presence of epithelial cells transformed into cancer cells, can be determined through amplification of mtDNA from prostate massage fluid, eclipsing current diagnostic techniques such as digital rectal examination and PSA. Recently Fliss et al. (2000) identified mutated mtDNA in urine samples of patients with bladder cancer. Similar findings in prostate massage fluid provide a non-invasive early detection method for prostate cancer. Different types of prostate cancer can be diagnosed, as well as differentiating between aggressive, fast growing cells in patients in contrast to prostate cancer as a whole. For example, the 3.4 kb deletion identified in the applicant's teachings, can be used as an indication of prostate cancer.

Early detection and monitoring of prostate cancer progression

The system and method disclosed may be used to detect cancer, and in particular prostate cancer, at an early stage, and before any histological abnormalities. For example, the system and method may be used to detect pre-neoplasia in prostate tissue.

The system can be used to detect the genesis and progression of prostate cancer. Mutations, including both subtle and hyper-mutation (Chen et al. 2002; Chen et al. 2003) in mitochondrial DNA from human prostate tissue, or fluid associated with the prostate (for example prostate massage fluid or urine), can be tested for the presence of neoplasia, and retested at intervals to follow cancer transformation, diagnose malignancy, or confirm continued benign status.

These mutations may be determined by comparison to mitochondria extracted from non-involved tissue such as, but not limited to: blood, urine, hair and buccal swabs. This direct comparison eliminates polymorphisms, maternal background or normal haplotype variation unassociated with disease. The mutations can also be compared to mitochondrial sequences associated with inter and intrapopulation variations. One or more mutations from fluid or tissue of the organ or body system in question, indicates possible disease genesis. The person is then monitored, at successive intervals, for an increase in mutations at other sites, and/or an increase in the number of mutated mitochondrial genomes, indicating disease progression. Benign tissue from the prostate cannot always be considered non-involved. In fact, as can be seen in Example 9, below, what appears to be benign tissue may contain mitochondrial mutations associated with pre-neoplasia, neoplasia, progression toward malignancy or malignancy. In addition, mutation load rather than specific

mutations may be instrumental in determining disease and progression of disease. The system and method detects heteroplasmic as well as homoplasmic mutations.

The prostate gland is monitored for mutations in the mitochondrial genome through prostate massage fluid (PMF) taken during an initial digital rectal examination (DRE) of the prostate. Cells within the PMF are concentrated, smeared on a slide and stained with PSA immunoperoxidase for identification of prostate epithelial cells. These prostate cells are selectively recovered through laser capture micro-dissection. The mitochondrial DNA from these cells is analyzed and compared to mitochondrial DNA from non-involved tissue, and/or to sequences of inter and intrapopulation variations. For example, the DNA analysis can comprise sequencing of the mtDNA. Total DNA is extracted from these cells and mitochondrial specific primers, designed for use with biopsy material treated with formalin (Table 5), are used to amplify the entire mtDNA genome with overlapping amplicons. These PCR products are then sequenced by methods well known to those in the art, including DNA resequencing arrays. Sequencing results are screened for heteroplasmies and mutations and compared to a database of known mtDNA mutations associated with malignant and benign prostate tissues. Based on these comparisons a designation is returned as to the condition of the prostate in regards to, but not limited to: benign (no mutations); pre-neoplasia or neoplasia (low level of mutations); or malignancy (high level of mutations). In the situation of benign, pre-neoplasia and neoplasia, the prostate can be monitored for progression through regular PMF screenings as described.

Alternatively, biopsy material which has been diagnosed as benign, atypical, abnormal can undergo similar testing by either laser capture micro-dissection of the biopsy, or the tissue can be scraped off the slides, or mounted tissue sections can be used, followed by DNA extraction, amplification, sequence analysis and database comparison.

As an alternative to sequencing, and comparison to a database, micro-array technology could be used to identify a specific pattern of mutations, or mutation load based on any number, or combination of the mutations listed in Table 4, through the construction of oligonucleotides, or a specific set of oligonucleotides.

Disease progression can be monitored by comparing mtDNA mutations at successive intervals to a database of mutations in mitochondrial genomes associated with pre-neoplasia, neoplasia and prostate cancer, including calculation of total mutation load. Prostate biopsy tissue can be tested for pre-neoplasia, neoplasia and/or malignant progression in cells described clinically as benign, normal, atypical or abnormal by common histological/pathological, or other clinical methods.

Similarly, the DNA may be analyzed for specific deletions that are known to be associated with disease, for example prostate cancer. This can be done by using PCR based technologies to screen for such deletions.

Assessment of mutations associated with aging

The system and method may be used to assess aging, based on the increasing frequency of mutations such as the "common deletion" of 4977-bp and other mutations of the mitochondrial genome (Liu et al. 1997). This information, in conjunction with health survey data, allows crucial statistical discrimination between separate causes resulting in the same mutation/deletion. Fortunately mtDNA is inherited exclusively through the ovum and is essentially clonal in nature (Van De Graaff & Fox, 1995). This permits carefully controlled studies of mutations/deletions within maternal lines through several generations to determine a reliable age related deletion frequency. This information may be used to develop treatment methods which slow the aging process.

Collection of samples

Biological samples can be collected by any known means, whether for the purpose of constructing a mtDNA sequence database, or performing a diagnostic test on an individual. Samples destined for database generation include, but are not limited to: tumour banks, maternal lineage studies involving affected and unaffected individuals from the same maternal lineage, as well as maternal lineage studies from groups or populations with high frequencies of specific disease such as, but not limited to: skin and prostate cancer, assessment of health status and aging. For example, FTA® Gene Cards® may be used to collect and archive biological samples. Suitable samples include any tissue or body fluid derived from mesothelium, epithelium, or endothelium. Such tissues and fluids include, but are not limited to blood, sputum, buccal cells, saliva, prostate massage fluid, sweat, bone, hair, lymph tissue, cervical smears, breast aspirate, fecal matter, ejaculate, menstrual flow, urine and biopsy tissue. Preferably, approximately 100 µl of blood, 100 µg to 25 mg of solid tissue is sampled. In the case of suspected skin cancer, skin cells or tissue, (from normal, NMSC and precursor lesions) is taken from skin biopsy or a routine suction blistering technique. Where a disease is suspected, primary care physicians, oncologists or other practitioners, may extract both normal and suspected disease tissue from the patient. For the purpose of analyzing sun exposure, tissue may be taken from the dermis or epidermis, or a combination of both.

For samples of tumours such as prostate or skin, replicate cross-sections (5 microns) of micro-dissected paraffin embedded tissues are de-paraffinized prior to one slide being stained with hematoxylin and eosin (HE), with the replicate stained with methyl green (MG), as is standard in the art. HE stains are graded by a pathologist for normal, precursor, and applicable grades of tumour progression. Replicate MG slides are used for laser capture, according to manufacturers recommendations (Arcturus) of graded cells.

Extraction of mtDNA

Extraction of DNA may take place using any method known in the art, followed by amplification of all or a region of the mitochondrial genome, and may include sequencing of the mitochondrial genome, as described in Current Protocols in Molecular Biology.

Analyzing mtDNA

The step of detecting the presence of mutations in the mtDNA can be selected from any technique as is known to those skilled in the art. For example, analyzing mtDNA can comprise sequencing the mtDNA, amplifying mtDNA by PCR, Southern, Northern, Western South-Western blot hybridizations, denaturing HPLC, hybridization to microarrays, biochips or gene chips, molecular marker analysis, biosensors, melting temperature profiling or a combination of any of the above. In addition, statistical techniques such as Inductive Rule Extraction, and Neural Networking can be used.

Sequencing of MtDNA

PCR

Polynucleotide sequences disclosed can be amplified by the polymerase chain reaction (PCR). PCR methods are well-known to those skilled in the art. PCR requires the presence of a nucleic acid to be amplified, two single stranded oligonucleotide primers flanking the sequence to be amplified, a DNA polymerase, deoxyribonucleoside triphosphates, a buffer and salts. The method of PCR is well known in the art. PCR is performed as described in Mullis and Faloona, 1987, Methods Enzymol., 155: 335, herein incorporated by reference.

In general, PCR is performed using template DNA (at least 1fg; more usefully, 1-1000 ng) and at least 25 pmol of oligonucleotide primers. A typical reaction mixture includes: 2μl of DNA, 25 pmol of oligonucleotide primer, 2.5 μl of 10X PCR buffer 1 (Perkin-Elmer, Foster City, CA), 0.4 μl of 1.25 μM dNTP, 0.15 μl (or 2.5 units) of Taq DNA polymerase (Perkin Elmer, Foster City, CA) and deionized water to a total volume of 25 μl. Mineral oil is overlaid and the PCR is performed using a programmable thermal cycler.

The length and temperature of each step of a PCR cycle, as well as the number of cycles, are adjusted according to the stringency requirements in effect. Annealing temperature and timing are determined both by the efficiency with which a primer is expected to anneal to a template and the degree of mismatch that is to be tolerated. The ability to optimize the stringency of primer annealing conditions is well within the knowledge of one of moderate skill in the art. An annealing temperature of between 40°C and 72°C is used. In general, initial denaturation of the template molecules normally occurs at between 92°C and 99°C for 4 minutes, followed by 20-40 cycles consisting of denaturation (94-99°C for 15 seconds to 1 minute/kb), annealing (temperature determined as discussed above; 1-2 minutes), and extension (72°C for 1 minute). The final extension step is generally carried out for 4 minutes at 72°C, and may be followed by an indefinite (0-24 hour) step at 4°C.

al. 1999). These heteroplasmic mtDNA changes result in the formation of heteroduplexes after PCR amplification of the mtDNA. Rapid screening for heteroplasmic mtDNA mutations is determined using the relatively new technique of denaturing high performance liquid chromatography (DHPLC) (Oefner & Underhill, 1998). This technique has recently been used to rapidly screen and identify whole mtDNA genomes for heteroplasmic point mutations down to levels <5% (Van den Bosch et al. 2000).

The DHPLC may be performed on the WAVE™ DNA Fragment Analysis System (Transgenomic, Omaha, USA) which provides a fully automated screening procedure. The same technology can be used to screen for mtDNA heteroplasmic mutations. Preferably, the entire mtDNA genome is amplified by PCR in 13 overlapping fragments using two different PCR conditions as described by van den Bosch et al. (2000). The 1-2 kb PCR products are digested into fragments of 90-600bp and resolved at their optimal melting temperature. Mutations are represented as two peaks and mutations with low percentages, such as <2% heteroplasmy as a 'shoulder' in the peak.

DNA sequencing can also take place using a microarray, as is known in the art (Chee et al. 1996).

Data Analysis

Once sequenced, normal and disease associated mtDNA sequences are archived for comparison in a database. Resequencing devices, micro-array technology, integrated microfluidic amplification and analysis systems, high-speed, high-throughput, mutation detection, and other methods may all be used with the methods disclosed.

Data obtained from the sequencing of the individual mitochondrial genome is compared to population level data. The data is obtained through obtaining samples and sequencing mtDNA as described above. Preferably, the database contains information from maternal line studies. The population level data is maintained in a database. Any suitable database can be used.

Preferably, a multidimensional evaluation research database of clinical and biological data is used, which provides the bio-informatics infrastructure necessary for the collection, processing and dissemination of information amassed by the laboratories involved in this venture. The database is a centralized electronic system which links networks resulting in a dynamic and powerful resource.

The database may be accessed through any known means, and preferably through a secure Internet pathway. Preferably, the database is developed using an e-commerce algorithm, built on a server and deployed using an application server which supports a high volume of concurrent users through optimized performance and scalability features. A separate "web" server can provide the foundation of the web-site architecture since it can serve as the central point through which all content, applications, and transactions must flow before reaching users.

Data mining algorithms known in the art are used to discover patterns, clusters and models from data (SAS 2000). Moreover, intelligent algorithms and methods will be developed for: occurrence of mutation and mutation rates, patterns of mutations for disease detection, information retrieval, and other complex sequence analysis software.

Nucleic Acid Members and Probes

The techniques disclosed provide for nucleic acid members and probes that binds specifically to a target nucleic acid sequence. The target nucleic acid sequence is a nucleic acid or a region of a nucleic acid that is to be detected, as indicative of disease such as prostate cancer, non-melanoma skin cancer and the like. The target nucleic acid sequences to be analyzed using a microarray are preferably derived from human tissue or fluid samples. The techniques provide for target nucleic acid sequences comprising RNA or nucleic acid corresponding to RNA, (i.e., cDNA), or DNA. Nucleic acid members are stably associated with a solid support to comprise an array as disclosed herein.

The nucleic acid members may be single or double stranded, and may be a PCR fragment amplified from cDNA.

Also disclosed are polynucleotide sequences comprising a probe. As used herein, the term "probe" refers to an oligonucleotide which forms a duplex structure with a sequence in the target nucleic acid, due to complementarity of at least one sequence in the probe with a sequence in the target region. The probe may be labeled, according to methods known in the art. A probe may be single or double stranded.

Diagnostic devices

Also disclosed are diagnostic devices such as biochips, gene chips or microarrays used to diagnose specific diseases or identify specific mutations. All sequenced mitochondrial genomes are assessed to create a consenus structure of the base pair arrangement and are assigned a prohibiting index for proportion of base pair deletions and mutations associated with a particular disease or disorder. The diagnostic arrangement is then used to create biochips, gene chips, or microarrays.

Once sequences associated with particular diseases, disease states or disorders are identified, hybridization of mtDNA to an array of oligonucleotides can be used to identify particular mutations. Any known method of hybridization may be used. Preferably, an array is used, which has oligonucleotide probes matching the wild type or mutated region, and a control probe. Commercially available arrays such as microarrays or gene chips are suitable. These arrays contain thousands of matched and control pairs of probes on a slide or microchip, and are capable of sequencing the entire genome very quickly. Review articles describing the use of microarrays in genome and DNA sequence analysis is available at www.gene-chips.com.

Microarray

Polynucleotide arrays provide a high throughput technique that can assay a large number of polynucleotides in a sample comprising one or more target nucleic acid sequences. The arrays are useful for gene expression analysis, diagnosis of disease and prognosis of disease (e.g., monitoring a patient's response to therapy, drug screening, and the like).

Any combination of the polynucleotide sequences of mtDNA indicative of disease, aging, or other health related mutations are used for the construction of a microarray.

The target nucleic acid samples to be analyzed using a microarray are derived from any human tissue or fluid which contains adequate amounts of mtDNA, as previously described, preferably prostate massage fluid, solid tumours, benign tissue, blood; or urine. The target nucleic acid samples are contacted with polynucleotide members under hybridization conditions sufficient to produce a hybridization pattern of complementary nucleic acid members/target complexes.

Construction of a microarray

The microarray comprises a plurality of unique polynucleotides attached to one surface of a solid support, wherein each of the polynucleotides is attached to the surface of the solid support in a non-identical preselected region. Each associated sample on the array comprises a polynucleotide composition, of known identity, usually of known sequence, as described in greater detail below. Any conceivable substrate may be employed.

The array is constructed using any known means. The nucleic acid members may be produced using established techniques such as polymerase chain reaction (PCR) and reverse transcription (RT). These methods are similar to those currently known in the art (see e.g. PCR Strategies, Michael A. Innis (Editor), et al. (1995) and PCR: Introduction to Biotechniques Series, C. R. Newton, A. Graham (1997)). Amplified polynucleotides are purified by methods well known in the art (e.g., column purification). A polynucleotide is considered pure when it has been isolated so as to be substantially free of primers and incomplete products produced during the synthesis of the desired polynucleotide. Preferably, a purified polynucleotide will also be substantially free of contaminants which may hinder or otherwise mask the binding activity of the molecule.

In the arrays of the invention, the polynucleotide compositions are stably associated with the surface of a solid support, wherein the support may be a flexible or rigid solid support.

Any solid support to which a nucleic acid member may be attached may be used. Examples of suitable solid support materials include, but are not limited to, silicates such as glass and silica gel, cellulose and nitrocellulose papers, nylon, polystyrene, polymethacrylate, latex, rubber, and fluorocarbon resins such as TEFLON™.

The solid support material may be used in a wide variety of shapes including, but not limited to slides and beads. Slides provide several functional advantages and thus are a preferred form of solid support. Due to their flat surface, probe and hybridization reagents are minimized using glass slides. Slides also enable the targeted application of reagents, are easy to keep at a constant temperature, are easy to wash and facilitate the direct visualization of RNA and/or DNA immobilized on the solid support. Removal of RNA and/or DNA immobilized on the solid support is also facilitated using slides.

The particular material selected as the solid support is not essential, as long as it provides the described function. Normally, those who make or use the invention will select the best commercially available material based upon the economics of cost and availability, the expected application requirements of the final product, and the demands of the overall manufacturing process.

Numerous methods are used for attachment of the nucleic acid members to the substrate (a process referred as spotting). For example, polynucleotides are attached using the techniques of, for example U.S. Pat. No. 5,807,522, which is incorporated herein by reference for teaching methods of polymer attachment. Alternatively, spotting is carried out using contact printing technology.

The amount of polynucleotide present in each composition will be sufficient to provide for adequate hybridization and detection of target polynucleotide sequences during the assay in which the array is employed. Generally, the amount of each nucleic acid member stably associated with the solid support of the array is at least about 0.1 ng, preferably at least about 0.5 ng and more preferably at least about 1 ng, where the amount may be as high as 1000 ng or higher, but will usually not exceed about 20 ng. Where the nucleic acid member is "spotted" onto the solid support in a spot comprising an overall circular dimension, the diameter of the "spot" will generally range from about 10 to 5,000 μm, usually from about 20 to 2,000 μm and more usually from about 50 to 1000 μm.

Control polynucleotides may be spotted on the array and used as target expression control polynucleotides and mismatch control nucleotides to monitor non-specific binding or cross-hybridization to a polynucleotide in the sample other than the target to which the probe is directed. Mismatch probes thus indicate whether a hybridization is specific or not. For example, if the target is present the perfectly matched probes should be consistently brighter than the mismatched probes. In addition, if all central mismatches are present, the mismatch probes are used to detect a mutation.

Target preparation

The targets for the microarrays, are derived from human fluid or tissue samples. It may be desirable to amplify the target nucleic acid sample prior to hybridization. One of skill in the art will appreciate that whatever amplification method is used, if a quantitative result is desired, care must be taken to use a method that maintains or controls for the relative frequencies of the amplified polynucleotides. Methods of "quantitative" amplification are well known to those of skill in the art. For example, quantitative PCR involves simultaneously co-amplifying a known quantity of a control sequence using the same primers. This provides an internal standard that may be used to calibrate the PCR reaction. The high density array may then include probes specific to the internal standard for quantification of the amplified polynucleotide. Detailed protocols for quantitative PCR are provided in PCR Protocols, A Guide to Methods and Applications, Innis et al., Academic Press, Inc. N.Y., (1990). Other suitable amplification methods include, but are not limited to polymerase chain reaction (PCR) (Innis, et al., PCR Protocols. A guide to Methods and Application. Academic Press, Inc. San Diego, (1990)), ligase chain reaction (LCR) (see Wu and Wallace, Genomics, 4: 560 (1989), Landegren, et al., Science, 241: 1077 (1988) and Barringer, et al., Gene, 89: 117 (1990), transcription amplification (Kwoh, et al., Proc. Natl. Acad. Sci. USA, 86: 1173 (1989)), and self-sustained sequence replication (Guatelli, et al., Proc. Nat. Acad. Sci. USA, 87: 1874 (1990)).

Also disclosed are labeled target or labeled probe. Any analytically detectable marker that is attached to or incorporated into a molecule may be used. An analytically detectable marker refers to any molecule, moiety or atom which is analytically detected and quantified. Detectable labels suitable for use include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels include biotin for staining with labeled streptavidin conjugate, magnetic beads (e.g., DynabeadsTM), fluorescent dyes (e.g., fluorescein, texas red, rhodamine, green fluorescent protein, and the like), radiolabels (e.g., ^{3}H, ^{125}I, 35S, ^{14}C, or ^{32}P), enzymes (e.g., horseradish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic (e.g., polystyrene, polypropylene, latex, etc.) beads. Patents teaching the use of such labels include U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

Means of detecting such labels are well known to those of skill in the art. Thus, for example, radiolabels may be detected using photographic film or scintillation counters, fluorescent markers may be detected using a photodetector to detect emitted light. Enzymatic labels are typically detected by providing the enzyme with a substrate and detecting the reaction product produced by the action of the enzyme on the substrate, and colorimetric labels are detected by simply visualizing the colored label.

The labels may be incorporated by any of a number of means well known to those of skill in the art. However, in a preferred embodiment, the label is simultaneously incorporated during the amplification step in the preparation of the sample polynucleotides. Thus, for example, polymerase chain reaction (PCR) with labeled primers or labeled nucleotides will provide a labeled amplification product. In a preferred embodiment, transcription amplification, as described above, using a labeled nucleotide (e.g. fluorescein-labeled UTP and/or CTP) incorporates a label into the transcribed polynucleotides. Alternatively, a label may be added directly to the original polynucleotide sample (e.g., mRNA, polyA mRNA, cDNA, etc.) or to the amplification product after the amplification is completed. Means of attaching labels to polynucleotides are well known to those of skill in the art and include, for example nick translation or end-labeling (e.g. with a labeled RNA) by kinasing of the polynucleotide and subsequent attachment (ligation) of a polynucleotide linker joining the sample polynucleotide to a label (e.g., a fluorophore).

In a preferred embodiment, the target will include one or more control molecules which hybridize to control probes on the microarray to normalize signals generated from the microarray. Labeled normalization targets are polynucleotide sequences that are perfectly complementary to control oligonucleotides that are spotted onto the microarray as described above. The signals obtained from the normalization controls after hybridization provide a control for variations in hybridization conditions, label intensity, "reading" efficiency and other factors that may cause the signal of a perfect hybridization to vary between arrays.

binding affinity between associated oligonucleotides and/or polynucleotides and the test polynucleotides.

Following detection or visualization, the hybridization pattern is used to determine quantitative information about the genetic profile of the labeled target polynucleotide sample that was contacted with the array to generate the hybridization pattern, as well as the physiological source from which the labeled target polynucleotide sample was derived. By genetic profile is meant information regarding the types of polynucleotides present in the sample, e.g. in terms of the types of genes to which they are complementary, as well as the copy number of each particular polynucleotide in the sample.

Diagnostic or Prognostic Tests

Also disclosed are diagnostic tests for detecting diseases. Also disclosed are prognostic tests for monitoring a patient's response to therapy. According to the method the presence of disease or the patient's response to therapy is detected by obtaining at fluid or tissue sample from a patient. A sample comprising nucleic acid is prepared from the fluid or tissue sample. The nucleic acid extracted from the sample is hybridized to an array comprising a solid substrate and a plurality of nucleic acid members, wherein each member is indicative of the presence of disease or a predisposition to a disease or disorder. According to this diagnostic test, hybridization of the sample comprising nucleic acid to one or more nucleic acid members on the array is indicative of disease, a predisposition to a disease or disorder, or in the case of a prognostic test, indicative of a patient's response to therapy.

Kits

Kits containing reagents and instructions to carry out the methods are provided. For example, the kit may comprise reagents and instructions for detecting mitochondrial deletions, mutations, heteroplasmies, homoplasmies in tissue specific samples and tissue associated fluids. The kits may also comprise one or more primers which hybridize to the mitochondrial genome for making a primer extension product. Kits may also include a solid support such as a disposable chip, means for holding the solid support, means for extraction of mtDNA and means for access to a database of mtDNA sequences.

For example, a kit to detect a mtDNA deletion associated with prostate cancer may include the forward and reverse 3.4 primers, reagents, and instructions.

Similarly, a kit to detect a mtDNA deletion associated with sun exposure or NMSC may include L404 primer, H4676 primer, the 3895 probe, reagents and instructions.

Other utilities for the present techniques, such as that described above and in the following examples, will be readily apparent to those skilled in the art.

The present invention is more particularly described in the following examples which are intended as illustrative only since numerous modifications and variations will be apparent to those skilled in the art.

benign tissue, removed from any surrounding tissue pathology. Tissues were laser capture microdisssected (LCM) by a qualified, clinical pathologist. Sequences were compared to mitochondrial DNA (mtDNA) extracted and sequenced from the patient's blood. Sequencing results indicate that mitochondrial mutations, often heteroplasmic, appear in all three tissue types. Contrary to the analysis of Alonso et al. (2005), mutations in mitochondrial DNA from malignant tissue, adjacent benign and distant benign were found in comparison with mtDNA sequences from the patient's blood. Heteroplasmy, mutant and normal mitochondrial genomes existing within the same individual, is considered evidence of recent mutation (Huoponen, 2001). Most mutations are not held in common between the three tissue types from the same individual. In fact, the comparative mutation loads for each tissue type is roughly the same, indicating that mtDNA mutations are active within the prostate in the presence of tumor, irrespective of prostate tissue type. Moreover, tissue which appears histologically benign, is often found with mutations in the mitochondrial genome. These mutations are not associated with mitochondrial sequences embedded in the nucleus, or NUMTS (nuclear/mitochondrial sequences) (Lopez, 1994). A lengthy cloning and sequencing study of human NUMTS utilizing a Rho^0 cell line, as well as comparison to known NUMTS archived in the NCBI database was done and no mutations were detected to known mitochondrial pseudogenes archived in public databases to ensure data clean of pseudogene data points.

In the 12 males clinically symptomatic for prostate cancer, wherein the pathology results indicated that the prostate tissue was not malignant, benign glandular tissue was recovered by LCM and the mitochondrial genomes amplified and sequenced from prostate needle biopsy tissues. Results indicate a complete absence of mtDNA mutations in four patients, mutations only in the non-coding control region in four patients, or mutations in both the control region and gene coding regions in four patients. However, the mutation locations were significantly different than the malignant cohort ($P > 0.01$). Additionally, only one coding region mutation was observed in these individuals and of these mutations one fell into the regions mentioned above (i.e. NDI, ND2, COXI and CytB) (patient 375 CytB at position 15081).

When both synonymous and non-synonymous mutations, inclusive of the four protein coding regions, were used as malignant disease markers, 30/31 (97%) of the prostate cancer group was identified. Table 1c lists the synonymous and non-synonymous mutations found in ND1, ND2, COX1, and CytB in tissue samples from the 31 individuals with prostate cancer. Mutations in ND1, ND2, COX1, and CytB are clearly associated with prostate cancer. Although the present techniques have identified the mutations listed in Table 1c and Figures 2 to 4, they also include any mutations in ND1, ND2, COX1 and CytB. In the same analyses, 1 of the 12 benign, symptomatic patients was included with the malignant group; however, this patient had a mutation in CytB, which may indicate early cancer progression. The rest of this group clustered with the blood, or normal control.

Example 13: The use of a 3895 bp mitochondrial DNA deletion as a marker for sunlight exposure in human skin

Since the initial discovery of the possible association of the 3895bp deletion with sun exposure as reported by Harbottle et al., 2004 and Durham et al., 2003, further work has confirmed that the association is present. As well, the additional work has enabled the use of this deletion for a diagnostic test of sun exposure or skin cancer detection, through the development of a novel method for detection of deletions and sequence rearrangements in a fully quantitative manner. This method provides for the annealing of a primer or probe to the newly formed sequence created by the deletion or insertion involved in the rearrangement and thereby allows the use of real-time quantitative PCR (qPCR) as a means of detection. The use of the qPCR platform enables the quantitative detection of the deletion rather than a simple presence and absence scenario. This quantitation is the basis of the test as it is the relative amounts of deletion that measures the level of sun exposure or the character of the malignancy rather than the simple presence or absence semi-quantitation previously reported. As well the qPCR platform enables the use of apparently unaffected or non-exposed tissue as the detection sensitivity is substantially greater than conventional PCR and ethidium bromide detection.

Background:

The incidence of non-melanoma skin cancer (NMSC) is increasing in populations of European origin (Severi and English, 2004). For example, one million new cases are diagnosed each year in the USA (Wesson and Silverberg, 2003) and 65,000 in the UK (figures provided by Cancer Research. UK). NMSC accounts for around 90% of skin cancers and consists of basal cell and squamous cell carcinomas (BCC and SCC, respectively). BCC are the commonest form of NMSC and arise predominantly from the basal keratinocytes of the epidermis but also from cells in hair follicles and sebaceous glands. They are locally invasive but rarely metastasise. SCCs are also derived from basal keratinocytes; however, in contrast to BCC, SCC may metastasize. Compared with BCC, SCC shows the greatest increase with age and is concentrated in the elderly (Severi and English, 2004). The relative density of NMSC is highest on body sites "usually" exposed to the sun when outdoors such as scalp, face, neck, and ears as defined by Armstrong (2004). SCC, however, differs appreciably from BCC in having a much lower density on body sites which are "occasionally" exposed to the sun such as shoulders, back, and chest as defined by Armstrong (2004).

Therefore, the major determinant of NMSC is the ultraviolet radiation (UVR) component of sunlight that induces DNA damage. Importantly it is both the pattern (more continuous versus intermittent) and the cumulative amount of sun exposure that influences the development of NMSC (Armstrong and Kricker, 2001). To determine a reliable marker of cumulative UVR exposure in human skin, the inventors and others have examined the novel idea of using mitochondrial DNA (mtDNA), rather than nuclear DNA, as a biomarker of UV-induced DNA damage (Pang et al, 1994; Berneburg et al, 1997; Birch-Machin et al, 1998; Birch-Machin, 2000;). Compared with mutation screening of nuclear DNA

Claims:

1. A method of detecting skin cancer, UV exposure or UV damage in a biological sample, the method comprising:

(a) quantifying, in the sample, a level of mutated mitochondrial DNA (mtDNA), wherein the mutation comprises a deletion spanning approximately nucleotides 547 to 4443 in the minor arc of the human mtDNA genome; and

(b) comparing the level of the mutated mtDNA in the sample to a reference value.

2. The method of claim 1, wherein: the skin cancer is basal cell carcinoma (BCC); the reference value is the level of mutated mtDNA clinically normal perilesional skin; and wherein a detection of BCC is found if the level of the mutated mtDNA in the sample is higher than the level of the mutated mtDNA in the clinically normal perilesional skin.

3. The method of claim 1, wherein: UV exposure of UV damage is detected; the reference value is the level of the mutated mtDNA in occasionally sun-exposed, rarely sun-exposed or sun-protected skin tissue; and wherein UV exposure of UV damage is detected if the level of the mutated mtDNA in the sample is higher than the level of the mutated mtDNA in the occasionally sun-exposed, rarely sun-exposed or sun-protected skin tissue.

4. The method of any one of claims 1 to 3, wherein the biological sample is: (i) a skin or tissue sample from a subject; or (ii) a tissue culture sample.

5. The method of claim 4, wherein the biological sample is a tissue culture sample, and wherein the method further comprises exposing the tissue culture sample to at least one sub-lethal dose of ultraviolet radiation (UVR) prior to step (a).

6. The method of claim 5, wherein the biological sample is exposed to a series of repetitive sub-lethal doses of UVR.

7. The method of claim 6, wherein the series comprises exposing the biological sample to a daily dose of UVR.

8. The method of any one of claims 5 to 7, wherein the UVR is from a solar-simulated UVR source.

9. The method of any one of claims 5 to 8, wherein the UVR comprises UVA, UVB, or UVA/UVB.

10. A method for determining cumulative UV exposure in a biological sample or monitoring a subject for sun exposure and associated non-melanoma skin cancer (NMSC), the method comprising:

(a) providing the biological sample or a tissue sample obtained from a subject;

(b) quantifying a level of mutated mitochondrial DNA (mtDNA) in the biological sample or tissue sample, the mutation comprising a deletion spanning approximately nucleotides 547 to 4443 in the minor arc of the human mtDNA genome;

(c) repeating steps (a) and (b) after an interval of time; and

(d) comparing the quantified levels of the mutated mtDNA.

11. A method for monitoring long term safety of a clinical UV phototherapy regime, the method comprising:

(a) providing a biological sample or a tissue sample obtained from a subject;

(b) quantifying a level of mutated mitochondrial DNA (mtDNA) in the biological sample or tissue sample, the mutation comprising a deletion spanning approximately nucleotides 547 to 4443 in the minor arc of the human mtDNA genome;

(c) repeating steps (a) and (b) after conducting the UV phototherapy regime on the biological sample for an interval of time; and

(d) comparing the quantified levels of the mutated mtDNA.

12. The method of claim 10, wherein the biological sample is a skin sample and wherein the NMSC is basal cell carcinoma (BCC).

13. The method of any one of claims 1 to 12, wherein the mtDNA mutation comprises: a deletion spanning approximately 3895 bp; a deletion comprising a span of mtDNA from approximately a mtTF1 binding site in the D-loop to tRNA methionine; or a deletion comprising substantially a 12s rRNA gene, a 16s rRNA gene, a ND1 gene and promoters for transcription of H and L strands.

14. The method of any one of claims 1 to 13, wherein the step of quantifying the level of the mutated mtDNA comprises amplification of the mtDNA.

15. The method of claim 14, wherein the step of quantifying the level of the mutated mtDNA comprises PCR analysis selected from: deletion specific PCR analysis, radioactive PCR analysis, quantitative PCR, real-time PCR (RT-PCR) analysis, or RT-PCR using recombinant, thermostable, Taq DNA polymerase.

16. The method of claim 14, wherein the step of quantifying the level of the mutated mtDNA comprises a quantitative PCR method.

17. The method of claim 15 or claim 16, wherein the level of the mutated mtDNA is quantified using a PCR primer chosen from: a PCR primer that bridges the mtDNA junction formed by the deletion spanning approximately nucleotides 547 to 4443; or a primer having a sequence corresponding to SEQ ID NO:145, SEQ ID NO:146, or SEQ ID NO:153.

18. The method of any one of claims 1, 2, 10, 11 and 13 to 17, wherein the biological sample: is from non-lesional skin; is from a dermis or an epidermis layer of a subject; or comprises occasionally sun exposed tissue, usually sun exposed tissue, or rarely sun exposed tissue.

19. The method of claim 18, wherein the biological sample is occasionally sun exposed tissue, usually sun exposed tissue, or rarely sun exposed tissue and wherein the reference value comprises a level of the mutated mtDNA in a reference sample, and wherein the reference sample comprises: rarely sun exposed skin tissue; occasionally sun exposed skin tissue; usually sun exposed skin tissue; or a blood sample.

20. Use of a quantified level of mutated mitochondrial DNA (mtDNA) in a biological sample, to detect UV damage in a subject, as a biomarker for skin cancer, or as a biomarker for UV damage, wherein the mtDNA mutation comprises a deletion between approximately nucleotides 547 to 4443 in the minor arc of the human mtDNA genome.

21. A kit when used for quantifying a level of mutated mitochondrial DNA (mtDNA), the mtDNA mutation comprising a deletion spanning approximately nucleotides 547 to 4443 in the minor arc of the human mtDNA genome, wherein the deletion is associated with skin cancer and/or UV damage, the kit comprising at least one primer for amplifying the mutated mtDNA.