NUCLEOTIDE DETECTION BY FLUORESCENCE RESONANCE ENERGY TRANSFER AND METHODS THEREOF

FIELD OF THE INVENTION

THIS INVENTION relates to methods for detecting the identity of a nucleotide in a nucleic acid using the Fluorescence Resonance Energy Transfer (FRET) technology.

BACKGROUND OF THE INVENTION

Fluorescence resonance energy transfer (FRET) is a spectroscopic method based on the energy transfer occurring between the electronic excited states of two dyes, called donor and acceptor, via a dipole-dipole interaction which can involve one or more molecules (Morrison LE, 1999; Ha T, 2001). The energy transfer can be induced only if the emission spectrum of the donor overlaps with the excitation spectrum of the acceptor, and if donor and acceptor are located in a sufficient proximity to each other (between 1 and 10 nanometers).

The relationship between energy transfer and donor/acceptor distance is defined mathematically as proportional to the inverse sixth power of the distance, with a maximal effect that is obtained in a range of distance specific for each donor-acceptor couple.

When the donor is irradiated at specific wavelengths, its excited state energy is transferred to the acceptor and dissipated as fluorescence, quenching the fluorescence at the donor site and simultaneously increasing the fluorescence intensity at the acceptor site. The fluorescence emission of the donor can be restored by separating the donor and the acceptor by a distance sufficient to considerably reduce the energy transfer.

In an alternative configuration, the acceptor group can be a quencher, so that the energy from the excited donor fluorophore is transferred to the quencher and dissipated as heat rather than fluorescence energy, and only a reduction of the fluorescence associated to the donor site can be monitored.

Various molecules have been tested as donor, acceptor, or quencher, as well as different excitation wavelengths and linker groups, so that specific combination of those elements have been optimized to achieve the better energy transfer efficiency and larger change in signal with respect to each specific FRET application.

As a general concept, FRET analysis provides means to indirectly measure the distance between chemical groups modified with the fluorophores and located on either the same molecule, or distinct, but interacting, molecules. The observed variations of fluorescence intensity are a consequence how close such groups are. However, this effect is not exploited to calculate absolute distances since FRET efficiency is also influenced by the angle between the dipole moments of the fluorophores, parameters often difficult to quantify with the necessary precision, and the most common applications are therefore qualitative rather than quantitative. Indeed, this technology allows the dynamic monitoring the molecular reactions or interactions by the means of the macroscopic changes in the observed fluorescence. In fact, even relative small distance changes between donor and acceptor sites can be magnified and detected.

FRET-based technologies have been extensively used in biological systems for detecting, *in vitro* and *in vivo,* the formation (or the destruction) of intra- or intermolecular interactions of proteins and nucleic acids. The progress in sensitivity of detection in the instrumentation make possible the extensive use of FRET for applications in biochemistry and biology, providing different kinds of information as a consequence of different experimental designs.

When applied to the study of protein-protein or protein-nucleic acid interactions, FRET permits to demonstrate that the two molecules bind to each other by virtue of their features such as three-dimensional structure and charge distribution.

In the case of nucleic acid-nucleic acid interactions, the variability and the complexity of the system is much less important, since these molecules interact essentially by hybridization on the basis of the classical A-T / A-U / C-G base pairing rules, making such interactions easily predictable. Therefore, FRET technologies were developed in the field of nucleic acid studies mostly with the scope of deducing information not on the structure but on the sequence of nucleic acids.

One or more nucleic acid molecules can be labeled at the level of two nucleotides and, following process such as de-/hybridization and/or enzymatic modification (polymerization, ligation, cleavage or amplification reactions), the proximity of two labeled nucleotides can be measured by FRET. Various strategies have been developed with respect to the configuration of the nucleic acid complex necessary to bring the FRET labels at an appropriate distance, and improve the sensibility and the capabilities of the methods. These assays can be performed using nucleic acids in solution or bound on a solid support. The first approach allows the hybridization

reaction to be faster, without unspecific adsorption, but the assay results generally less sensitive than the second approach, which provides means for removing unhybridized nucleic acids, enzymes or any other undesirable molecule before further processing.

Various reviews recently published discuss the approaches so far developed on the basis of FRET for identifying nucleotides in a specific sequence, especially with aim of detect known or unknown polymorphic sites (Morrison LE, 1999; De Angelis DA, 1999; Kwok PY, 2001; Didenko VV, 2001). The prior art discloses how two (or even three) FRET-generating groups can be brought into proximity by virtue of the intra- and/or inter-strand hybridization and reactions between complementary nucleic acid sequences, providing the means of genotyping nucleic acids of different origin, in particular with the aim of identifying single nucleotide polymorphisms (SNPs). The actual location of the donor/acceptor groups can be onto different strands which are paired to each other or both paired to a third strand or on the same strand

As shown in the literature (Ju J et al., 1995), DNA molecules containing one fluorophore linked to the 5' terminal nucleotide and the other to an internal nucleotide have been synthetized and used simply as primers to be paired to a target nucleic acid before starting a conventional sequencing reaction starting from the available 3' terminal nucleotide. In this case, the FRET effect is not even actually used to directly sequence a nucleic acid but only as a way to provide primers having an improved fluorescent signal for traditional electrophoresis-based sequencing technologies.

A FRET-based, single base extension monitoring system involves the template-directed incorporation of the fluorophore-labeled nucleotide by making use of a primer labeled with the other fluorophore at the 5' end (Chen X et al., 1997; Chen X and Kwok PY, 1997; Kwok PY and Chen X, 1999). The FRET effect is however measurable only when the double-labeled primer is released from the template, since the distance between the two fluorophores makes the energy transfer irrelevant when the primer is hybridized to the template. Then, the extended primer needs to be dehybridized from the template, since it is only in the form of a single stranded, free bending molecule that the donor and acceptor groups can be at a distance sufficient for generating a measurable intra-strand FRET signal (WO 97/22719).

Prior art discloses other technologies involving PCR or denaturation evaluation steps, as well as various enzymes (nucleases, ligases, retrotranscriptases) and complex arrangements of fluorophores and nucleic acids, in order to generate and detect intra- and/or intra-strand FRET interactions (Bernard PS et al., 1998; Nauck M et

al., 1999; Hirshhorn JN et al., 2000; Tyagi S et al., 2000; Myakishev M et al., 2001; Solinas A et al., 2001).

However, the prior art fails to demonstrate how a FRET system for nucleic acid analysis can be provided directly by extending the 3' end of a primer, internally labeled in an appropriate position and paired to a template molecule, in order to identify a nucleotide in the template molecule.

SUMMARY OF THE INVENTION

It has been found that FRET can allow the direct and specific detection of a nucleotide incorporated in nucleic acid following single base extension at the 3' end, without thermocycling or other manipulations and reactions.

The present invention provides novel methods for detecting the incorporation of a nucleotide after a single base extension reaction. The methods make use of the Fluorescence Resonance Energy Transfer (FRET) effect generated by two fluorophores linked to the same nucleic acid strand. The methods of the invention provide a FRET-detectable signal following the single base extension of an internally FRET fluorophore-labeled primer hybridized to a complementary nucleic acid sequence, by the means of a nucleic acid polymerase and of an appropriate FRET fluorophore-labeled nucleotide precursor. The distance between the labeled internal nucleotide and the nucleotide at 3' end in the primer has been optimized for the scope of the invention, avoiding any further processing following the base extension.

The methods can be used, either in solution or a solid support, for applications requiring the rapid and precise identification of a nucleotide in a specific position of a nucleic acid, such as in genotyping, as well as to characterize nucleic acid polymerases and molecules interacting with these enzymes.

DESCRIPTION OF THE FIGURES

Figure 1 illustrates one of the embodiments of the invention, wherein a single pair of template and labeled primer molecules are exposed to two distinct single base extension reactions, each including a different labeled nucleotide.

Figure 2 illustrates one of the embodiments of the invention, wherein different template molecules are exposed to the same primer molecule and single base extension reaction including two, alternatively labeled nucleotides.

Figure 3 illustrates the possible molecular events when the methods of the invention are applied on nucleic acids immobilized on a solid surface in the form of DNA colonies *via* its 5' end. After exposing such surface to the nucleic acid polymerase together with the dNTP modified with an acceptor fluorophore and primers internally labeled with a donor fluorophore, allowing the extension reaction, and finally washing away the reactants, the modified dNTP may still be found on the solid surface in a undesired manner following an unspecific extension due to the hybridization of two immobilized nucleic acids leading to a 3' extension (a). Other undesirable donor and acceptor groups may also be immobilized on the surface due to the non-specific binding of the labeled dNTP or primers to the immobilized nucleic acid (b' and b") or to the surface (c' and c"). Nonetheless, only specific signals resulting from the specific hybridizations and extensions events are actually detected when applying the excitation wavelength for the donor fluorophore (donor excitation) suitable for obtaining a measurable fluorescent emission from the acceptor fluorophore (acceptor emission) due to FRET effect (F).

Figure 4 shows the sequence of the codons associated to the ApoE alleles used as a model for genotyping in the examples (the polymorphic position is boxed and bold).

Figure 5 shows the sequence and the complementarity of the primers and templates used in the examples. The nucleotides associated to the SNP1 or SNP2 variants of ApoE alleles, as defined in Figure 4, are the boxed and bold nucleotides in the templates sequences. The nucleotides in the tested primer sequences modified with a FRET-generating fluorophore are underlined and bold.

Figure 6 provides a comparison of the fluorescence intensity monitored in real-time after the addition of a DNA polymerase and either a complementary (dATP-Cy5) or not complementary (dGTP-Cy5) labeled nucleotide to the hybridized molecules (SNP1fT / P1r-t4). The curves were generated by interpolation of the measured data.

Figure 7 shows the increase of fluorescence monitored in real-time after the addition of an exonuclease to a sample containing the double labeled complex generated in the experiment represented in Figure 6. The curve was generated by interpolation of the measured data.

Figure 8A and 8B provide different representation of the data monitored in real-time on the DNA polymerase preference for the unlabeled (dATP) versus labeled

(dATP-Cy5) forms of the same nucleotide. The curves were generated by interpolation of the measured data.

Figure 9A and 9B shows the spectra before and after single base extension (60 minutes reaction) for SNP1fT / P1r-t6 (A) and SNP1fT / P1r-t9 (B) couples of template / primer molecules.

Figure 10A and 10B show different ways to arrange and detect a template containing two polymorphic sites, like the ApoE SNP1 and SNP2 presented in the examples, on a solid support.

Figure 11 shows the increase of the signal ratio FRET filter set / donor filter set following primer extension reaction using different dNTP-Cy5 and synthetic templates immobilized on a solid support mimicking the three relevant ApoE haplotypes ε2 (A), ε3 (B), and ε4 (C) associated to SNP1 and SNP2. The data are plotted against each added dNTP-Cy5.

Figure 12 shows the ratio of FRET signal before and after polymerase-mediated incorporation of dNTP-Cy5 using templates immobilized on a solid support as DNA colonies The data are plotted against each added dNTP-Cy5.

Figure 13A and 13B show the variation of FRET signal following the addition of different combinations of un -/labeled nucleotide and/or fluorophores in a primer extension reaction involving alternative templates for SNP2 and a labeled primer (see Figure 5), as measured using the appropriate FRET filter. TMR means Tetramethylrhodamine, while (+) and (-) stand for, respectively, positive and negative nucleotide, according to the definition given in figures 4, 5, and 9, or Table I.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to novel FRET-based methods allowing the direct and specific detection of a nucleotide incorporated, after a single base extension, in a nucleic acid complex, comprising a primer and a template molecule, by making use of two fluorophores, the first one linked to an internal nucleotide of the primer molecule, and the second one is linked at the 3' terminal nucleotide of the primer molecule. This FRET-compatible arrangement of the two fluorophores results from the single base extension of the internally labeled primer molecule hybridized to the complementary nucleic acid sequence in the template, followed by the exposure to a nucleic acid polymerase and to an appropriate fluorophore-labeled nucleotide precursor. The FRET

effect can be directly measured and associated to the identity of the nucleotide complementary to the labeled one in the template molecule.

The main object of the present invention is to provide a method for identifying a nucleotide in a nucleic acid by means of Fluorescence Resonance Energy Transfer (FRET), comprising the following steps:

a) providing a first nucleic acid containing the nucleotide to be identified in a single-stranded conformation at least in the portion including the nucleotide to be identified and the sequence immediately in 3' to said nucleotide;

b) providing a second nucleic acid comprising a single stranded 3' end complementary to the sequence of the first nucleic acid immediately in 3' to the nucleotide to be identified, and a first FRET-generating fluorophore linked to a nucleotide located from 1 to 12 nucleotides far apart from the 3' terminal nucleotide of said second nucleic acid;

c) hybridizing the first and the second nucleic acid, resulting in a nucleic acid wherein the sequence immediately in 3' to said nucleotide is in the double stranded conformation;

d) adding a nucleic acid polymerase and at least a nucleotide chemically linked to a second FRET-generating fluorophore;

e) allowing the nucleic acid polymerase to perform a single base extension at the 3' end of the first nucleic acid; and

f) monitoring the incorporation, or the absence of incorporation, of the nucleotide chemically linked to the second FRET-generating fluorophore using a FRET-based analytical technology.

In the sense of the present application, the terms "first nucleic acid" and "second nucleic acid" are equivalent to "template" and "primer", respectively. Template and primer molecules are usually distinct molecules but, exceptionally, can belong to the same molecule, for example in the so-called "hairpin" or "hammerhead" structures, where a single stranded molecule has a 3' end sequence is complementary to an internal sequence and can therefore work as a traditional primer.

Naturally occurring or non-naturally occurring nucleotides may be present in the template and in the primer molecules, but they have to be compatible with nucleic acid polymerase and the labeled nucleotides used in the following steps of the invention. Usually, the template is a genomic DNA, complementary DNA, recombinant DNA, RNA, or mRNA, which can be directly purified from cells, amplified using Polymerase

Chain Reaction (PCR), or replicated using a plasmid in an host cell prior to be submitted to the methods of the invention.

The primer molecule is usually obtained by chemical synthesis. A large variety of synthetically modified nucleic acids have been developed for chemical and biological methods in order to improve the detection and/or the functional diversity of nucleic acids. These modified molecules can be fully compatible with natural polymerizing enzymes, maintaining the base pairing and replication properties of the natural counterparts, as recently reviewed (Thum O et al., 2001).

The template sequence may be completely or only partially in a single-stranded conformation. The complementary sequence comprised into the primer should have be chosen to provide an hybridizing activity sufficient to provide both a stable complex to be used by the nucleic acid polymerase and very low non-specific hybridizing activity to any other sequence. The primer can be 5 to 100 bases in length, but preferably 15 to 25 bases in length. The segment of the primer hybridizing with the template should be at least 5 bases in length, and preferably it should be hybridized with the at least 5 to 20 bases located at the 3' end.

As shown in the examples, the single base extension reaction and the base assignment can be performed according to the methods of the invention even if the labeled nucleotide internal to the primer is not complementary to the corresponding nucleotide in the template, generating a mismatch. If the bases adjacent to this unpaired position are enough to guarantee the specificity and the stability of the complex, the FRET is still detectable.

Many chemical groups suitable as donor/acceptor probes are known in the literature, such as Texas Red (TR), Tetramethylrhodamine (TMR) , Bodipy-TMR, Cyanine dyes (Cy3, Cy5), or Alexa dyes (Alexa 594, Alexa 568, Alexa 546). These groups can provide a FRET signal when irradiated at the wavelengths comprised between 428 and 650 nanometers using lamps and filters working in this range. The suitable fluorescent groups can be covalently bound to the precursor nucleotide or a primer by making use of the reactions and of the optimization protocols described in the examples or known in the literature (Glazer AN and Mathies RA, 1997; Hung et al., 1997).

The examples demonstrate that particularly efficient donor/acceptor combinations are the ones having Texas Red as donor, and a Cyanine Dye molecule, such as Cy5, as acceptor. Other possible donor-acceptor pairs covering intermolecular distances

compatibles with FRET based applications and applicable here are Cy3-Cy5 and Tetramethylrhodamine-Texas Red, and other combinations described in the literature (Hung et al., 1998; Panchuk-Voloshina N et al., 1999). The label can be linked to the nucleotides by the technologies known in the art, for example using the C6-amino linker as shown in the examples.

Cy5 acts as acceptor of many fluorophores emitting in the red region and can be selected for its good spectral properties, very high extinction coefficient, moderate to good quantum yield. When the measurement by fluorescence spectroscopy is made using filters and Arc-Mercury lamp make this fluorophore the most sensitive one. Another interesting feature concerning Texas Red-Cy5 is that Cy5 signal is very weak following direct irradiation, meanwhile the transfer from Texas Red by FRET is clearly visible. This lamp performs poorly at the wavelength of optimal irradiation of Cy5 but is excellent to irradiate Texas Red. Finally, the integral overlap between emission of Texas Red and absorption of Cy5 is such that allows measurements of FRET in the context of a double stranded nucleic acid when the fluorophores are separated by a number of nucleotides compatible with FRET parameters. Such number, as calculated in the example, is comprised between 1 and 12 bases, and in particular between 4 and 10 bases.

The fluorescent moieties can be combined either with a donor site on the internal labeled nucleotide and an acceptor site on the 3' terminal nucleotide, or in the opposite conformation. Even if the number of combinations tested in the examples is limited when compared to the variety of fluorophores described in the literature, any combination of donor / acceptor fluorophores which shows a measurable FRET effect can be used. However, a specific approach can be preferred in certain applications.

For example, if the complex is generated in solution, the donor is preferably located in the internal position, and the acceptor is preferably linked to the nucleotide incorporated as 3' terminal nucleotide. In this way, being the precursor is in molar excess to the internally labeled primer, the fluorescence change before and after the single base extension due to FRET effect is larger and more easily measurable. In a typical experimental set-up, a single stranded DNA primer internally labeled with Texas Red is hybridized with a template, then a DNA polymerase is added together with a Cy5 labeled nucleotide precursor. The analysis of the FRET signal after the single base extension can confirm the presence of a nucleotide complementary to the incorporated

precursor in the unpaired position available to the extension reaction and sufficiently near to the Texas Red labeled internal nucleotide.

The optimal distance for obtaining the FRET effect is typical for each combination but it can be calculated and tested as shown in the examples, so that a combination can be chosen according to the expected range of distance to be analyzed and/or to the experimental set-up. Moreover, as in embodiments shown below, if two or more second FRET-generating fluorophore are used (each associated to a specific nucleotide), the formulas allow to choose the fluorophores and the experimental set-up in way that each second FRET-generating fluorophore generate a FRET effect distinct from the other one(s). Some further basic modifications to the methods can be obtained by looking to some applications in the literature. For example, in order to improve the FRET efficiency, a third fluorophore can be also linked to another internal nucleotide primer or in the template (Kawahara S et al, 1999).

The methods of the present invention can be also comprised in methods for the evaluation of the extension properties of nucleic acid polymerases, for the evaluation of the preference of nucleic acid polymerases for the same nucleotide when labeled or unlabeled, or for screening modulators of nucleic acids polymerases.

The methods of the present invention can be applied also not only in solution but also to solid phase systems, that is, when the primer and / or the template molecules are covalently or non-covalently linked to a support. The immobilization of nucleic acid on a solid surface is of particular interest in any situation wherein large number of samples have to be analyzed and high throughput technologies need to be applied. An important application is the generation and the analysis of nucleic acids immobilized onto a discrete area of a solid support, wherein each discrete area corresponds to a template or a primer having a specific sequence. For example, nucleic acids immobilized on a solid surface can be provided in form of microarrays (Cheung VG et al., 1999) or DNA colonies (WO 96/04404; WO 98/44151; WO 00/18957; WO 02/46456).

The methods of the invention are compatible to commercially available FRET-based analytical technology like multiwell plates, fluorometers, fluorescence microscopes, CCD cameras, lamps with beam regulation, image analysis software. Interferences or other problems are possible but the necessary adjustments and materials needed to maximize the efficiency of the analysis are well known in the art.

The elements of the methods of the invention, as well the basic hybridization and polymerization conditions, are well known in the literature as well their variants, which can used in the methods of the invention if they do not have properties or features clearly interfering with FRET and/or enzymatic polymerization. For example, fluorophore labeled nucleotide triphosphates (NTPs), as well as deoxyribonucleotide triphosphates (dNTPs) and dideoxyribonucleotide triphosphates (ddNTPs) can be chosen as precursors to be labeled and used in the polymerization reactions. Examples of nucleic acid polymerases which can be used in the present invention are Eukaryotic and Prokaryotic DNA polymerases, which can be purified from cells or cloned and expressed in a recombinant system (Perler FB et al., 1996).

The single base extension reaction should contain the correct combination of polymerase, nucleic acids, and labeled nucleotide as well as other compounds, such as dimethyl sulfoxide (DMSO), Bovine Serum Albumin (BSA), or $MgCl_2$, known to be effective in supporting the activity of the chosen polymerase. at specific concentrations well known in the prior art and easily adjustable. For example, the nucleic acid polymerase is a DNA polymerase and the nucleotide linked to the second FRET-generating fluorophore is a deoxyribonucleotide triphosphate, or a dideoxyribonucleotide triphosphate (the latter one is used when it is desired to block any further elongation).

The methods of the present invention provide reliable information on the variants of a sequence into a sample when it results form the incorporation of the 3' terminal labeled nucleotide by a single base extension reaction using a nucleic acid polymerase, able to elongate the paired, internally labeled primer, and a labeled nucleotide precursor. The generation of a FRET-detectable complex depends from the presence of a nucleotide complementary to the labeled precursor. Therefore the comparison of fluorescence signal before and after distinct reactions in which different labeled precursors are present allows the identification of the nucleotide present in the unpaired position at the 5' of the nucleotide paired to the unlabeled paired 3' terminal nucleotide of the internally labeled primer. However, the primer should have a 3' terminal sequence large and specific enough to allow the correct pairing of the primer immediately in 3' to the nucleotide to be identified. Some basic situations can be summarized as follows.

If the single labeled intermediate complex obtained by the hybridization of primer and template is exposed to two distinct single base extension reactions, each

including different labeled nucleotides, the FRET signal will be measurable only when the precursor is complementary to such unpaired nucleotide. In this way, the parallel analysis of different alternative nucleotides for a single position can be provided (Figure 1).

Alternatively, a single base extension reaction may contain different precursors of two or more distinct nucleotides, each labeled with a specific group that is distinguishable from the other when it is incorporated in the internally labeled strand and provokes a energy transfer with the group already present in the primer. Thus, different template molecules can be exposed to the same primer molecule and to the same single base extension reaction. In this way, the parallel analysis of different templates for alternative nucleotides in a single position can be provided (Figure 2).

For example, the extension reaction can be performed adding at the same time two distinct nucleotide precursor types, each labeled with a fluorophore molecule having, in one case, donor properties and, in the other case, acceptor properties when put in proximity of the fluorophore internal to the primer. This solution allows to reduce the number of reactions needed to identify a polymorphic site nucleotide from 4 to 2 for two SNPs (or from 2 to 1 for a single SNP), considerably improving and speeding the entire process. This way of multiplexing the methods can be obtained, for example, using Texas Red in the primer, and Cy5 or Cy3 on each distinct nucleotide. The first combination allows measuring the FRET effect of Texas Red on Cy5, the second one of Cy3 on Texas Red. Other combinations that can be used with Texas Red as donor fluorophore with Cy5 as acceptor fluorophore, are Tetramethylrhodamine as donor fluorophore with Texas Red as acceptor fluorophore, and Alexa 546 as donor fluorophore with Texas Red as acceptor fluorophore.

The methods of the invention can be adopted in any application needing a fast and reliable identification of a nucleotide in a nucleic acid, and in particular comparing the possible alternative bases in a specific nucleotide position in different samples or across a population, and associating allelic variation typical of a genotype to a phenotype, such in the case of single nucleotide polymorphisms (SNPs). The methods of the invention allow the identification of a single nucleotide polymorphism or mutation in a template potentially containing the polymorphism or the mutation, the position of said polymorphism or mutation corresponding to the first not hybridized nucleotide in 5' of the sequence of the template hybridized to the primer.

The present invention results particularly useful in methods involving the detection of nucleotide polymorphisms by making use of nucleic acids immobilized on a solid surface, especially when in the form of DNA colonies (WO 98/44151; WO 98/44152; WO 00/18957; WO 02/46456). When the DNA colonies have to be analyzed by primer extension, the methods of the invention for FRET-based detection for primer extension do not need the in troduction of a blocking step prior to the primer hybridization to prevent non-specific extensions. The blocking step is usually carried out by performing one or several extension reaction in presence of a polymerase or a terminal transferase and nucleotides with blocked 3'-end such as, but not limited to, ddNTPs. The blocking extension reactions can be repeated several time if necessary, with or without denaturation and re-hybridization step in between. The use of the appropriately internally labeled primers descried by the present invention on the DNA colonies avoids any FRET signal associated to non-specific binding and/or extensions (Figure 3). The process results therefore more precise and efficient by obtaining better signal to noise ratio and/or longer read lengths. At the same time, the method is less complex and considerable efforts and time can be saved.

Once established by any other technology (PCR, cytogenetics, DNA sequencing, etc.) that a segment of DNA represents a good candidate for containi ng such variants, and/or that some variants have a direct effect on a cellular, physiogical, or pathological process, the iterative application of the methods of the invention on the same segment of DNA isolated from a statistically relevant number of individuals can lead to the identifications of positions wherein specific mutations are significantly associated to altered phenotypes, or help the identification of samples containing the variants known to be relevant (Carlson CS et al., 2001; Kwok PY, 2001). The present invention provides an alternative solution to the problem of identifying this information expressed or genomic DNA sequences in order to identify haplotypes, or specific SNPs, associated a disease or to other relevant phenotypes.

The methods of the present invention allowed the characterization of primers and template models for genotyping two polymorphic sites of the human ApoE gene, in particular in the case of the variants associated to codons 112 and 158. The methods show to be readily amenable to both solution- and solid phase-implemented high throughput format, in isothermal condition and with no other enzyme or manipulation needed. Furthermore, the real-time monitoring of specific nucleotide incorporation can

allow the acquisition of kinetic data useful for testing different polymerases and / or different fluorescently-labeled nucleotides.

The methods of the present invention can be also used for determining the kinetic of nucleotide incorporation by a DNA polymerase, and therefore they can be comprised in methods for evaluating the efficiency of nucleic acid polymerases, either identified by screening genomic DNA or generated by mutagenizing known ones. These enzymes can be screened to determine the conditions for their optimal use (buffer composition and pH, additives, consumption of probes and reagents, sensitivity of detection, selectivity), by time-monitoring the fluorescence variation in the context of identical primer/template complexes and labeled precursors. Another possibility is to verify the interfering effect of other compounds (small molecule or other proteins) on these enzymes, with the aim of screening modulators of nucleic acids polymerases. The previous embodiments also permit measuring the preference of the polymerase when exposed to a mixture of labeled and non-labeled nucleotides, which is important, for example, when applying a method of *in situ* sequencing method on DNA colonies (WO 98/44152).

The methods of the invention can be eventually adapted to generate and identify further bases by extending the labeled hybridized primer not only using a single base, but also by adding more than one base. It is however essential that the labeled nucleotide allow further elongation, or that the eventual blocking group to be removed. Another condition is that the labeled nucleotide(s) incorporated after the first one are separated by a number of nucleotides making not exceeding the distance from the internal fluorophore needed to achieve a FRET signal for the chosen donor / acceptor combination. Moreover, if the totality of precursor molecules is labeled, the nucleotides incorporated after the first one alter the fluorescence profile of the complex in a non-linear, limited way, due to collision-related and/or self-quenching problems. A way to avoid such limitation is to provide, at each extension reaction, a mixture of unlabeled and labeled precursors to a large number of internally labeled primer / template intermediate complexes (WO 98/44152). Even if all such intermediates are potentially extended, only a limited proportion will be FRET-detectable, while the rest of complexes are still available to be modified into FRET-detectable complexes by a new elongation reaction. The fluorescence variations measured at each round of elongation result more linear and more easily measurable.

Given that nucleic acid polymerases usually shows to incorporate the unlabeled precursor rather than the labeled one and that this preference varies from nucleotide to nucleotide for a given enzyme, preliminary experiments have to be performed in order to identify the proportion of modified and unmodified molecules to be added for each nucleotide. If, for example, the preliminary measurements show that the modified nucleotide is incorporated even 10 times less efficiently than the unmodified, an equal proportion of FRET-detectable and FRET-undetectable complexes can be theoretically obtained by mixing the labeled and unlabeled precursors in a 10:1 proportion. Another possibility to improve sequencing capabilities of the methods of the invention is to allow the removal the FRET-generating fluorophore on the nucleotide already incorporated by any of the methods known in the art (WO 93/21340), and repeating the steps d) to f) described above.

The present invention provide also kits, comprising at least a nucleic acid having an internal nucleotide located from 1 to 12 nucleotides far apart from the 3' terminal nucleotide linked to a first FRET-generating fluorophore and at least a nucleotide chemically linked to a second FRET-generating fluorophore, which are suitable for the identification of single nucleotide polymorphisms or mutations, for the evaluation of the extension properties of nucleic acid polymerases, for the evaluation of the preference of nucleic acid polymerases for the same nucleotide when labeled or unlabeled, or for screening modulators of nucleic acids polymerases.

EXAMPLES

EXAMPLE 1: Materials & Methods

a) Mathematical models for FRET analysis

The distance (R) between a donor and acceptor fluorescent group to generate a FRET signal is calculated by Förster's equation (Förster T, 1948)

$$R = R_0 (1/E - 1)^{1/6}$$

wherein R_0 is the distance for 50% transfer efficiency E. R_0 is calculated in Angstrom [Å]) using the following formula

$$R_o = 9.786 \times 10^3 (\kappa^2 n^{-4} \Phi_D J)^{1/6} \text{ [Å]}.$$

wherein:

κ^2 is a geometric factor that accounts for the relative orientation in space of the donor emission and acceptor absorption transition dipoles. If both donor and acceptor

can rotate isotropically within times much shorter than the time scale of fluorescence emission, as in the present case, then $\kappa^2 = 2/3$.

n is the refractive index of the medium between the donor and the acceptor, and has been taken to be 1.3 for DNA samples.

Φ_D is the donor quantum yield.

J is a measure of the spectral overlap between emission spectrum of the donor and the absorption spectrum of the acceptor and corresponds to:

$$J = \int_0^\infty \varepsilon_A(\lambda) F_D(\lambda) \lambda^4 d\lambda / \int_0^\infty F_D(\lambda) d\lambda \text{ [cm}^3\text{M}^{-1}\text{]}$$

wherein:

λ is the wavelength in centimeters.

$F_D(\lambda)$ is the corrected fluorescence of the donor at wavelength λ normalized by the integral in the nominator.

$\varepsilon_A(\lambda)$ is the absorption of the acceptor at wavelength λ.

The previous formulas, when applied to the fluorescent labels used in the examples, provided the following results. Using an extinction coefficient for Cy5 of 250,000 $M^{-1}cm^{-1}$ at its maximum of absorption (Schobel U et al., 1999), the overlap integral *J* for the couple Texas Red-Cy5 was calculated to be 8.4×10^{-13} cm^3M^{-1}. The quantum yield (Φ_D) for dNTP-Texas Red was 0.29, in agreement with the reported value of 0.27 (Panchuk-Voloshina N et al., 1999). The calculated value of R_O was 57.6 Å. The efficiency of fluorescence energy transfer (E) was quantified as the fractional decrease of Texas Red donor fluorescence due to the binding of Cy5-labelled acceptor and was expressed by $E = 1 - F_{DA}/F_D$ where F_{DA} and F_D are the relative yield of fluorescence of the donor in the presence and absence of the acceptor, respectively.

The extent of quenching was calculated using an Excel macro command by fitting the resulting spectrum with donor and acceptor alone spectra using the equation:

$$R = D(Q) + A(I)$$

where D, A, and R are the donor, acceptor and resultant spectra respectively and Q is the quenching factor of the donor fluorescence and I the increase factor of acceptor fluorescence after incorporation of the Cy5 labeled dNTP. The efficiency of quenching of donor fluorescence is defined by E=1-Q.

b) DNA sequences

Probes and templates used in the following experiments have been designed on the sequence of the human Apolipoprotein E (ApoE) mRNA (EMBL Acc. No. K00396; SEQ ID NO: 1) and protein (SWISSPROT Acc. No. P02649; SEQ ID NO: 2).

ApoE was discovered as a plasma protein involved in lipoprotein metabolism, synthesized by the liver but also made locally in the brain. There are three major isoforms of ApoE (ε2, ε3, ε4), which are differentially represented amongst populations (Smith JD, 2000). These ApoE gene variants result from single base substitutions in the first nucleotide within the two codons (Figure 4), leading to the consequent substitutions of Arginine for Cysteine at amino acid 130 (or 112, excluding the leader sequence in the numbering) and Cysteine for Arginine at amino acid position 176 (or 158, excluding the leader sequence in the numbering). The ε3 allele is the most prevalent, "wild-type" form. The three allelic forms have different charges, determining detectable differences in the interaction with other proteins and in the electrophoretic mobility, and are associated to specific physiological effects such as alterations of cholesterol levels in plasma or cardiovascular diseases. The ApoE model is therefore particularly indicated for validating genotyping methods.

Oligonucleotide templates and primers were designed on the sequence surrounding SNP1 and SNP2 of the human ApoE gene and used to demonstrate the utility of the methods of the present invention for identifying single base mutations. The template SNP1fT (SEQ ID NO: 3) is 62-mer having the sequence corresponding to the nucleotides 407-468 of the forward strand in the ε2 and ε3 alleles. The template SNP1fC (SEQ ID NO: 4) differs from SNP1fT for the nucleotide substitution T to C in the position corresponding to the nucleotide 448, specific for the ε4 allele. The template SNP1rA (SEQ ID NO: 5) is a 59-mer having the sequence corresponding to the reverse strand of nucleotides 428-486 in the ε2 allele. The template SNP2rA (SEQ ID NO: 6) is a 61-mer having the sequence corresponding to the reverse strand of the nucleotides 566-626 in the ε2 allele. The template SNP2rG (SEQ ID NO: 7) differs from SNP2rA for the nucleotide substitution A to G in the position corresponding to the nucleotide 586, specific for the ε3 and ε4 alleles. The primers used in the examples are un-/labeled 20-mers which are designed to be complementary to the sequence 3' to the target nucleotide (Figure 5).

Unlabeled oligonucleotide templates and probes were synthesized by Microsynth (Switzerland). Fluorescently labeled oligonucleotide probes were synthesized by Eurogentec (Belgium) using standard phosphoramidite chemistry. Due to availability of commercial precursors, Thymidine was chosen as the nucleotide to be labeled whenever the labeling is at the level of an internal nucleotide in the primer. A deoxy-Thymidine, modified at C6 using an amino modifier linker, was introduced during the synthesis of the primer. The fluorescent label was added to the linker only post-synthetically. Synthetic oligonucleotides were purified using reverse-phase high pressure liquid chromatography (RP-HPLC). The oligonucleotides were dissolved in 10 milliMolar sodium borate (pH 7.3), at a final concentration of 100 microMolar and analyzed using an HPLC system (HP 1090; Hewlett Packard) equipped with a column Aquapore RP300, C18 (220x2.1 mm; Perkin Elmer). The binary gradient used was based on 100mM TEAA (triethyl ammonium acetate, pH 7.0) as A and acetonitrile as B, and consisted of 1% acetonitrile / minute over 30 minutes followed by 5% acetonitrile / minute over 10 minutes. The traces were recorded at the wavelength of 260 nanometers for unlabeled oligonucleotides and also at the wavelength of 570 nanometers for fluorescently labeled samples.

c) Synthesis, purification and analysis of labeled nucleotides

The labeled Cytidine and Uridine nucleotides used in the examples are commercially available (Amersham). The labeled Adenosine and Guanidine nucleotides used in the examples have been prepared starting from dATP and dGTP aminopropargyl precursors (AP3) commercially available (NEN).

Fifty nanomoles of dATP(AP3) or dGTP(AP3) were dissolved in 475 microliters of $NaHCO_3$ 0.1 Molar (pH 8.4), then added into a vial of Cy5 monofunctional reagent (Amersham), and mixed thoroughly. The labeling reaction was allowed to proceed at room temperature for 14 hours. The reaction mixture was then diluted to 2 milliliters with TMACl (tetramethyl ammonium chloride, 0.1 Molar, pH 7.5) and subjected to ion exchange chromatography purification at a flow rate of 2.6 milliliters/minute on a 2 milliliters column which was filled with DEAE-Sephacel (Sigma) and equilibrated with 10 milliliters of TMACl 1 Molar (pH 7.5), followed by 10 milliliters of 0.1M TMACl (pH 7.5). The reaction mixture was loaded onto the column and washed with 40 milliliters of TMACl 0.1 Molar (pH 7.5) to remove excess of hydrolyzed reagent. The product was eluted with TMACl 0.2 Molar (pH 7.5) and fractions (each 1.5 milliliters) were collected. Fractions containing the labeled product were identified by UV-visible spectroscopy,

and then pooled before being vacuum evaporated in a speed vacuum concentrator. The final product was dissolved in 2 milliliters of TEAA 0.1 Molar (triethyl ammonium acetate, pH 7.0) and subjected to RP-HPLC purification column (Macherey Nagel ET250/8/4 Nucleosil 7C18; 120 Å) at a flow rate of 1 milliliter/minute using a Waters system (Millipore Corp.) equipped with two Model 510 pumps and an automated gradient controller. The elution was monitored at 279 nanometers using a Model 481 UV-visible detector. The binary gradient applied is 0% B over 1 minute, 0 to 18% B over 9 minutes, 18 to 30% B over 44 minutes, 30 to 80% B over 4 minutes, where the phase mobile A is 100 milliMolar TEAA (pH 7.0) and B is acetonitrile. Typical yields of purified Cy5 derivatives were about 60%, calculated by spectroscopy using the molar extinction coefficient of 250,000 M^{-1} cm^{-1} at 650 nanometers (Schobel U et al., 1999).

The purified fluorophore-labeledes dGTP-Cy5 and dATP-Cy5 were subjected to ElectroSpray Ionization Mass Spectrometry (ESI MS) to verify their molecular mass. The experimental masses were of 1197.94 for dGTP-Cy5 and of 1182.04 for dATP-Cy5 for theoretical average masses of 1198.00 and 1182.00 respectively.

d) Evaluation of the donor-acceptor couples for FRET applications

In addition to the usual criteria for a donor-acceptor pair related to the specific spectroscopic properties, other requirements were taken into account in the choice of fluorescent labels. In particular, the necessity to allow FRET detection by fluorescence microscopy of DNA attached to solid surfaces with the use of standard microscopy optical filter sets mostly affected the selection of dyes. In order to get a reliable signal from each dye without the interference from the other one, it is preferable to achieve a difference as large as possible in wavelength for their maximum emission. Furthermore lamp performance was also considered.

For example, the advantages for the couple Texas Red/Cy5 selected and its use in fluorescence microscopy arise from three observations. Firstly, the lamp profile of the Arc-Mercury lamp is well suited for the donor irradiation and little direct irradiation of Cy5 is observed using the FRET filter set. Direct irradiation of the acceptor at the excitation wavelength of the donor is usually a source of noise implicating further corrections of signal. Secondly, little fluorescence is detected with the FRET filter due to the donor alone and autofluorescence of the system is minimal for the FRET filter set up due to the separation between excitation and emission wavelengths. Thirdly, corresponding fluorescently labeled precursors are commercially available.

Texas Red oligonucleotides are custom synthesized with the possibility to introduce the label at Thymidine (or Uridine) at internal positions in the sequence. Cy5 nucleotides for dUTP, and dCTP are commercially available, while the two other bases dATP and dGTP were labeled as described before using standard protocols.

The integral overlap and Ro values have been calculated for the pair Texas Red - Cy5, obtaining an integral overlap, *J* factor of 8.4 x 10^{-13} $cm^{-3}M^{-1}$ and a Ro value of 57.6 angstrom (5.76 nanometers), therefore in a range of values well suited for sensitive FRET measurements (Clegg RM, 1992). Other fluorophores were tested as adequate, for instance, Alexa 594 with better spectral properties than Texas Red (Panchuk-Voloshina N et al., 1999). These fluorophores offer an alternative to the Texas Red/Cy5 pair. Another alternative combination is represented by the Texas Red, as acceptor fluorophore covalently linked to the oligonucleotide, and Cy3, Tetramethylrhodamine, or Alexa 546 as donor fluorophore.

As described below, these different possibilities permits the simultaneous use of one label for each possible base at one polymorphic site or the simultaneous use of two primers for identifying two polymorphic sites (Figure 1 and 2), reducing the number of experiments. A single reaction incorporation reaction and two visualizations with two FRET filter sets allow the correct assignment of the correct base at the variation site, in solution or on a solid surface (Figure 3).

e) Formation and measurement of FRET-detectable nucleic acid complexes in solution

Intermediate, non-FRET detectable nucleic acid complexes are formed by mixing equimolar amounts of template and labeled probes in TMN buffer (40 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, 50 mM NaCl). The solution is heated 3 minutes at 95°C and cooled to 20° at 1°C/minute followed by dilution with the polymerase buffer (40 mM Tris HCl pH 7.5, 20 mM $MgCl_2$, 50 mM NaCl, 0.1 mg/ml BSA, 6 mM DTT, 20% DMSO (v/v))

Two samples are typically prepared and compared for demonstrating the correct enzymatic incorporation of labeled nucleotide by single base extension (Figure 1). In the first one, the labeled nucleotide is the one expected to be incorporated at the 3' end since it is complementary to the target site (positive sample), and another in which the labeled nucleotide is not complementary to the target site (negative or misincoporation control). A 5-fold molar excess of the two possible dNTPs for the target site to be determined are added, followed by the addition of 25 Units of T7 Sequenase (US Biochemical; 13 Units / microliter) and the recording continued for 1 hour. Alternatively,

a mixture of labeled and unlabeled nucleotides (dATP-Cy5 and dATP, with dATP representing from 0% to 20% of the mixture) was used.

For real-time measurements, using Texas red as a donor, the sample is excited at 560 nanometers and the emission recorded at 610 nanometers. Fluorescence emission spectra and time-course emission recordings of fluorescently labeled species were performed in a Jasco FP 750 spectrofluorometer equipped with a temperature controller and stirring unit Jasco ETC-272T. Unless stated otherwise, measurements were performed at 20°C using excitation and emission bandwidths of 5 and 10 nanometers respectively.

Experimental denaturation curves were generated in a spectrophotometer (Jasco V550) equipped with a temperature controller (Jasco ETC-505T). Each couple of synthetic complementary oligonucleotides was diluted at a final concentration of 0.5 microMolar in 1 milliliter of a solution containing 100 milliMolar NaCl, 10 mM sodium cacodylate (pH 7.2)

f) Formation and measurement of FRET-detectable nucleic acid complexes on a solid support

The embodiments related to the methods involving a solid support were demonstrated by the means of the covalent attachment of 5' phosphate templates according to the literature (Adessi C et al., 2000).

Briefly, 5' phosphorylated templates were prepared as 1 microMolar solution in 10 mM 1-methyl-imidazole (pH 7.0; Sigma Chemicals) containing 40 mM 1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC; Pierce). 25 microliters of oligonucleotide solution was added per NucleoLink well (NUNC), and incubated at 50°C for 2 hours. The wells were then washed three times (1 minute each) with 5xSSC buffer (0.75 M NaCl, 75mM sodium citrate pH 7.0). Texas Red labeled primers were diluted in TMN buffer at a concentration of 0.5 microMolar and incubated with template-functionalized NucleoLink tubes using a Peltier thermal cycler (PTC-200, MJ research). The temperature program applied was; 2 minutes at 65°C, 1 min at 63.5°C, followed by a 1°C decrease per minute up to a final temperature of 20°C. The tubes were finally washed 4 times in 0.5xSSC (1 minute per wash) and twice with 5xSSC before measuring fluorescence emission.

The intermediate complexes covalently linked to the support were incubated with 25 microliters of a polymerase buffer (40 mM Tris HCl pH 7.5, 20 mM $MgCl_2$, 50 mM NaCl, 0.1 mg/ml BSA, 6 mM DTT, 20% DMSO (v/v)) containing 5 microMolar dNTP

labeled with Cy5, Cy3, Tetramethylrhodamine, or Alexa 546. Finally 1.3 units of T7 Sequenase 2.0 DNA polymerase (Amersham) was added and the reaction mixture incubated for 5 min at 37°C. The fluorescence emission was measured after rinsing the wells as described (Adessi C et al., 2000).

EXAMPLE 2: FRET-detectable nucleic acids complexes in solution

Nucleic acid complex intermediates were obtained in solution by mixing the combination of template and primers shown in Figure 5 at the concentration of 0.25 nanoMolar in a volume of TMN buffer of 100 microliters, with a final volume of 1 milliliter after the dilution with the polymerase buffer.

A FRET-generating fluorophore (Texas Red) was inserted in various positions along primers of equal length to evaluate the range of distances leading to measurable FRET signals after the incorporation of a labeled nucleotide at the 3' end, as well as the effect of the label on nucleic acid complex stability. The nucleotides separating into the primer the fluorophore labeled nucleotide from the 3' end to be extended were comprised between 1 and 19, so that it was possible to measure the FRET interaction in a double stranded configuration for a discrete range of distances. Thymidine was chosen as internal nucleotide to be labeled which can be naturally present in the sequence (as for P1f-t and P2f-t series), or obtained by replacing other nucleotides in the original sequence and generating a mismatch (as for P1r-t series). Thus, the effect of internal fluorophore was studied not only in optimal conditions (perfect matching between primer and template), but also in sub-optimal conditions (labeled nucleotide not complementary to the corresponding nucleotide on the template). Unlabeled and 5' labeled primers were used as controls. In the latter case, however, a distance corresponding to 19 bases is unfavorable for obtaining a FRET signal in the double stranded configuration.

The effect on the stability of the nucleic acid complexes resulting from the single base extension was studied by generating denaturation curves of the templates when hybridized with the probes of equal length but labeled in different positions. The transition from the double stranded complex to single stranded molecules is measurable in solution as a shift of the absorbance at 260 nanometers when the melting temperature (Tm), which is proportional to the strength of the complex, is reached.

In a typical experiment, a solution containing the double stranded complexes was gradually heated starting from 20°C up to 80°C at the rate of 0.5°C per minute, measuring the absorbance at 260 nanometers during the all process. The Tm value measured for each couple of oligonucleotides, together with the FRET efficiency measured using positive and negative 3' extensions, are reported in Table I. These values represent the averaged result of at least three experimental determinations, unless otherwise indicated.

The comparison of the Tm values obtained for the different template / primer couples shows that the nucleic acid complexes resulting from the hybridization of labeled, mismatch containing probes (P1r-t4, -t6, and t9) displayed a reduction in Tm of about 10°C when compared to the unlabeled or to the 5' labeled primers, which showed the same Tm (within experimental error). For the other couples of templates and labeled probes, where no mutation was introduced, there was no relevant variation of the Tm value when compared to the unlabeled or the 5' labeled primers. This suggests that the label itself linked to the nucleotide via the C6-amino linker does not interfere with the duplex stability, while an internal non-complementary labeled nucleotide interferes with the duplex stability only in a limited way.

Fluorescence spectroscopy, homogeneous solution assays were performed using the same couples of probes and templates for which the denaturation curve was determined, with the aim to quantify the extent of fluorescence signal variation for labeled probes upon duplex formation and to real-time measure the quenching of donor emission after incorporation of acceptor (Cy5)-labeled dNTP complementary to the target site SNP1 or SNP2. The extent of donor quenching by FRET for each internally labeled probe was calculated from their spectra. All the experiments were performed with the two possible labeled dNTPs, the positive (the one that should be incorporated) and the negative (the one that should not be incorporated) ones, clearly demonstrating the method allows to discriminate correctly the incorporated base by making use of the Texas Red quenching values. The fluorescence emission was recorded before and after the single base extension .

Satisfactory results in terms of quenching efficiency have been obtained when the Texas Red labeled thymidine was from 1 to 12 nucleotides far apart from the 3' terminal nucleotide to be extended with the dNTP-Cy5. This result appears to be not affected from the decrease of the duplex DNA stability due to the mismatch-associated, internal fluorophore in certain primers. The more intense FRET effect and the better

ratio between the signal obtained using the positive or the negative base were measured when the internal label was from 4 to 10 bases far apart from the Cy5 labeled nucleotide incorporated at the 3' end (Table I).

Other experiments were performed in solution to better characterize the primer extension reactions provided by the methods of the invention.

Real-time monitoring shows that, after adding the DNA polymerase and a nucleotide complementary to the polymorphic site in SNP1, the emission signal of the donor decreases over time, while it remains stable if the nucleotide is not complementary to such site (Figure 6). In order to further confirm that the measured quenching was due to energy transfer from the labeling of the same molecule, an exonucleolytic reaction was performed. Shrimp alkaline phosphatase (5 units) was added to a sample where the single base extension positive reaction for the SNP1fT / P1r-t4 hybridized couple of molecule proceeded for 2 hours. After a further incubation of 30 minutes, the recording is then restarted followed by the addition of 5 units of T4 DNA polymerase, which start degrading the dephosphorylated 3' of the strands. The fluorescence was almost recovered to its initial value before FRET as a consequence of the loss of spatial proximity of the donor-acceptor pair arising from the internucleotide bonds breakage (Figure 7).

Another aspect studied by the means of real-time monitoring was the preference of DNA polymerase for labeled versus unlabeled nucleotides , in order to check how the methods of the invention work when different ratios of labeled and unlabeled nucleotides are applied. The analysis of the signal (quenching of Texas Red by FRET) observed after such extensions with these mixtures provides a simple, precise and rapid method for measuring the nucleic acid polymerase preference, since the extent of quenching of Texas Red is proportional to the degree of incorporation of dNTPs-Cy5 following primer extension.

The kinetic of incorporation was measured by real-time monitoring for different mixtures of dATP-Cy5 and dATP, containing from 0% to 20% dATP. At saturation values, it can be seen that the increase of the content of unlabeled dATP in the mixture result in a decrease of fluorescence (Figure 8A). The same data can be represented in linearized form if transformed as a function of the ratio [dATP]/[dATP-Cy5 + dATP] using the equation [100/Q]-1, where Q is the value of Texas Red fluorescence quenching at saturation for a given mix of labeled and unlabeled dATP, expressed in % (Figure 8B). The slope of the linear function reflects the enzymatic preference of the

polymerase for dATP compared to dATP-Cy5. If there is no preference between the two molecules, a value of 1 should be obtained. However, at least for the DNA polymerase used in the example (T7 Sequenase), a preference value of 5.49 was obtained for dATP-Cy5 (Figure 9B). Considering this preference value, in order to achieve a 10% fluorescent label incorporation into the extended DNA molecule, a mix of 61% dATP-Cy5 and 39% dATP should be used (5.49 x 100 % / (10 –1) = 61%).

The preference value for any other combination of acceptor-donor fluorophores, nucleotides, and nucleic acid can be easily calculated with the above described approach. This evaluation is helpful for defining the mixing proportions in order to incorporate the desired proportion of labeled fluorophores by primer extension using internally labeled primers.

Finally, an alternative way to illustrate the data provided in Table I is to compare the emission spectra recordings (Figure 9). The spectra were obtained with SNP1fT / P1r-t6 (A) and SNP1fT / P1r-t9 (B) template / primer couples, but similar results for other primer / template described before with the two possible dNTPs-Cy5, as well as at different times of the process of hybridization and extension. It is evident the ratio between the fluorescence emission measured when the positive base (dATP-Cy5) is present in the extension reaction, with a maximal difference at a wavelength of 610 nanometers, and the almost overlapping spectra when the negative nucleotide (dGTP-Cy5) is present.

The previous experiments demonstrate how the methods of the invention can allow, in a direct and reliable way, the discrimination of single base extension reactions associated to a polymorhypic site in solution, also when the primer contains a labeled internal nucleotide not complementary to the template. This information is obtained applying simple conditions, i.e. the same required for the single base extension. For example, the FRET effect can be directly measured while the reaction proceeds, at the same temperature, or at end of the reaction, at room temperature. This represents a clear advantage over prior art (Morrison LE, 1999; Chen X et al., 1997; WO 97/22719) which teaches more complex requirements, such as conditions leading to the denaturation of the double stranded structure resulting from the hybridization of template and primer molecules after the single base extension, two or three primer molecules, or additional reactants and further manipulation of the samples.

EXAMPLE 3: FRET-detectable nucleic acids complexes in solid phase

The detection of the single base extension by the means of the methods of the invention on a solid support was validated using templates and internally labeled primers by analyzing the model system used in solution and a DNA array generated by a self-patterning process (DNA colonies; WO 00/18957). The genotyping of ApoE using a primer-guided nucleotide incorporation assay on a support has been already performed (Syvanen AC et al., 1990), but no specific protocol has been published for the identification of such mutations on a surface using a FRET-based nucleotide incorporation assay.

The templates can be generated and prepared for the single base extension in different ways, allowing each immobilized strand to contain either one or two polymorphic sites, as shown in Figure 10A and 10B. The situations here represented are simply examples on how templates and primers can be arranged on a solid support before performing the single base extension reaction.

The templates, once immobilized on the surface via the 5' end and then enzymatically digested and/or denaturated in order to obtain single stranded molecules, are exposed to the reaction mixture containing the polymerase and the labeled nucleotide. In the present genotyping model system, the possible bases complementary to the site of variation can be A or G for SNP1 and C or T/U for SNP2 (see Figures 4, 5, and 9).

Equimolar amounts of templates pairs, each representing a specific genotype (SNP1fT and SNP2rA for ε2; SNP1fT and SNP2rG for ε3; SNP1fC and SNP2rA for ε4), were mixed and immobilized on a solid support before being hybridized with fluorescent primers P1r-t9 and P2f-t7, each selected for detecting a specific variable site (on the forward strand for SNP1 and on the reverse strand for SNP2).

Each dNTP-Cy5 was assayed for T7 Sequenase incorporation in an individual well in triplicate copies. Measurements have been taken before and after extension with each possible dNTP using the fluorescence microscope with the donor filter and the FRET-specific filter, and calculating the ratio of the signal obtained using the filter FRET in respect to the donor filter. The increase of this ratio is then plotted as a function of the added base giving a 'fingerprint' for each genotype with a highly specific signal (Figure 11). The signal obtained for negative bases is at the level of expected values for misincorporation of dNTP-Cy5, usually between 5 to 15%. In the ε4 example

(Figure 11C), it can be seen a higher misincorporation value for A but the specific signal is still sufficiently higher to assign the correct genotype unambiguously.

In another experiment, DNA colonies resulting from solid phase amplification of a DNA fragment isolated from genomic DNA and including ApoE gene were obtained as described in WO 00/18957. The sum of intensities of individual DNA colonies has been calculated before and after incorporation of each particular dNTP-Cy5. A clear assignment of the genotype can be deduced, where C and A are the positive bases as expected for an ApoE ε3 homozygous individual (Figure 12).

Finally, the variation of FRET signal following the addition of nucleotides labeled with different fluorophores was evaluated in primer extension reactions including the same labeled primer (P2f-t7) and different templates for SNP2 (SNP2rG or SNP2rA), together with differently labeled nucleotides (dCTP-Cy5 and dUTP-Cy3; Figure 13A), or with different fluorophores for the same nucleotide (Cy3, Alexa 546, Tetramethylrhodamine; Figure 13B). All conditions and materials provided the expected indication on the template, demonstrating that different combinations of fluorophores, templates, and primers are equally effective and precise in identifying the incorporated nucleotide.

The previous experiments demonstrate how the methods of the invention can also be performed using template molecules immobilized on a solid support, applying essentially the same conditions and criteria shown to be working in solution, at the same facilitating the process of genotyping using this approach (see Figure 3).

Table I

Template (target nucleotide)	Primer	Nucleotides separating donor and acceptor fluorophores	Tm (°C) (n ≥ 3)	Texas Red quenching efficiency [2]	
				(Positive base)	(Negative base)
SNP1fT (T)	P1r	N.A.	77.5±0.5	N.A.	N.A.
	P1r-t19	19	78.6±0.5	N.D.	N.D.
	P1r-t4	4	68.4±0.5	0.64 (A)	0.13 (G)
	P1r-t6	6	66.6±0.5	0.74 (A)	0.05 (G)
	P1r-t9	9	67.4±0.5	0.68 (A)	0.05 (G)
SNP1rA (A)	P1f	N.A.	71.9±0.5	N.A.	N.A.
	P1f-t1	1	71.7±1.1	0.45 (T)	0.10 (C)
	P1f-t10	10	71.5 [1]	0.67 (T)	0.01 (C)
SNP2rA (A)	P2f	N.A.	66.7±1.2	N.A.	N.A.
	P1f-t19	19	66.6±0.7	N.D.	N.D.
	P1f-t7	7	66.2±1.2	0.72 (T)	0.04 (C)
	P1f-t12	12	65.8 [1]	0.13 (T)	N.D.

[1] n=1

[2] Calculated using intensity at maximum emission wavelength with an excitation wavelength of 560 nanometers.

N.A.: not applicable

N.D.: not determined

REFERENCES

Adessi C et al., Nucleic Acids Res 28, E87(2000).

Bernard PS et al., Am J Pathol 153, 1055-1061(1998).

Carlson CS et al., Curr Opin Chem Biol 5, 78-85 (2001).

Chen X et al., Proc Natl Acad Sci USA 94, 10756-61 (1997).

Chen X and Kwok PY, Nucleic Acids Res 25, 347-53 (1997).

Chen X and Kwok PY, Genet Anal 14, 157-63 (1999).

Cheung VG et al., Nat Genet 21, 15-19 (1999).

Clegg RM. Meth Enzymol 211, 353-388 (1992).

De Angelis DA, Physiol Genomics 1, 93-99 (1999).

Didenko VV, Biotechniques 31, 1106-1 (2001).

Förster T, Ann Phys 2, 55-75 (1948).

Glazer AN and Mathies RA, Curr Opin Biotechnol 8, 94-102 (1997).

Ha T, Methods 25, 78-86 (2001).

Hirschhorn JN et al., Proc Natl Acad Sci U S A 97, 12164-12169 (2000).

Hung SC et al., Anal Biochem 252, 78-88 (1997).

Hung SC et al., Anal Biochem 255, 32-38 (1998).

Ju J et al., Proc Natl Acad Sci USA 92, 4347-4351 (1995).

Kawahara S et al., Nucleic Acids Symp Ser 42, 241-242 (1999).

Kwok PY, Annu Rev Genomics Hum Genet 2, 235-258 (2001).

Morrison LE, J Fluorescence 9, 187-196 (1999).

Myakishev MV et al., Genome Res 11, 163-169 (2001).

Nauck M et al., Clin Chem 45, 1141-1147 (1999).

Panchuk-Voloshina N et al., J Histochem Cytochem 47, 1179-1188 (1999).

Perler FB et al., Adv.Protein Chem 48, 377-435 (1996)

Schobel U et al., Bioconjugate Chemistry 10, 1107-1114 (1999).

Smith JD, Ann Med 32, 118-127 (2000).

Solinas A et al., Nucleic Acids Res 29, E96 (2001).

Syvanen AC et al., Genomics 8, 684-692 (1990).

Thum O et al., Angew. Chem Int Ed 40, 3990-3993 (2001).

Tyagi S et al., Nat Biotechnol 18, 1191-1196 (2000).

CLAIMS

1. A method for identifying a nucleotide in a nucleic acid by means of Fluorescence Resonance Energy Transfer (FRET), comprising the following steps:
   a) providing a first nucleic acid containing the nucleotide to be identified in a single-stranded conformation at least in the portion including the nucleotide to be identified and the sequence immediately in 3' to said nucleotide;
   b) providing a second nucleic acid comprising a single stranded 3' end complementary to the sequence of the first nucleic acid immediately in 3' to the nucleotide to be identified, and a first FRET-generating fluorophore linked to a nucleotide located from 1 to 12 nucleotides far apart from the 3' terminal nucleotide of said second nucleic acid;
   c) hybridizing the first and the second nucleic acid, resulting in a nucleic acid wherein the sequence immediately in 3' to said nucleotide is in the double stranded conformation;
   d) adding a nucleic acid polymerase and at least a nucleotide chemically linked to a second FRET-generating fluorophore;
   e) allowing the nucleic acid polymerase to perform a single base extension at the 3' end of the first nucleic acid; and
   f) monitoring the incorporation, or the absence of incorporation, of the nucleotide chemically linked to the second FRET-generating fluorophore using a FRET-based analytical technology.

2. The method of claim 1 wherein the first FRET-generating fluorophore is linked to a nucleotide located from 4 to 10 nucleotides far apart from the 3' terminal nucleotide of said second nucleic acid.

3. A method as claimed in claim 1 or 2, wherein the first FRET-generating fluorophore is a fluorescence donor site and the second FRET-generating fluorophore is a fluorescence acceptor site.

4. A method as claimed in claim 1 to 2 wherein the first FRET-generating fluorophore is a fluorescence acceptor site and the second FRET-generating fluorophore is a fluorescence donor site.

5. The methods of claim 3 or 4 wherein the fluorescence donor site is a Texas Red or Alexa 594 molecule and the fluorescence acceptor site is a Cyanine Dye molecule.

6. The methods of claim 3 or 4 wherein the fluorescence acceptor site is Texas Red, and the fluorescence donor site is Cy3, Tetramethylrhodamine, or Alexa 546.

7. A method as claimed in any of the previous claims, wherein two or more second FRET-generating fluorophore are used, each associated to a specific nucleotide, and wherein each second FRET-generating fluorophore generate a FRET effect distinct from the other one(s).

8. A method as claimed in any of the previous claims, wherein the FRET-generating fluorophores are chemically linked to the nucleotides using a C6 amino linker.

9. A method as claimed in any of the previous claims, wherein the 5' end of the first and / or the second nucleic acid is linked to a solid support.

10. The method of claim 9 wherein the link is a non-covalent link.

11. The method of claim 10 wherein the link is a covalent link.

12. The method of claims 10 or 11 wherein multiple copies of the first or of the second nucleic acid are immobilized onto a discrete area of a solid support, and wherein each discrete area corresponds to a first or a second nucleic acid having a specific sequence.

13. The method of claim 12 wherein the nucleic acids immobilized on a solid surface are provided in the form of microarrays or DNA colonies.

14. A method as claimed in any of the previous claims, wherein the nucleic acid polymerase is a DNA polymerase and the nucleotide linked to the second FRET -

generating fluorophore is a deoxyribonucleotide triphosphate, or a dideoxyribonucleotide triphosphate.

15. A method for the identification of a single nucleotide polymorphism or mutation using the method of any of the previous claims, wherein the second nucleic acid is a nucleic acid molecule potentially containing the polymorphism or the mutation, the position of said polymorphism or mutation corresponding to the first not hybridized nucleotide in 5' of the sequence of the first nucleic acid hybridized to the second nucleic acid.

16. The method of claim 15 wherein the second nucleic acid comprises the sequence corresponding to the codons 112 and/or 158 of human ApoE gene.

17. A method for evaluating the extension properties of nucleic acid polymerases comprising a method of claims 1 to 14.

18. A method for the evaluation of the preference of nucleic acid polymerases fo r the same nucleotide when labeled or non-labeled comprising a method of claims 1 to 14.

19. A method for screening modulators of nucleic acids polymerases comprising the method of claims 1 to 14.

20. A method for sequencing nucleic acids according to the method s of claims 1 to 14, further comprising the removal the second FRET-generating fluorophore and repeating the steps d) to f).

21. A kit for the identification of single nucleotide polymorphisms or mutations comprising at least a nucleic acid having an internal nucleotide located from 1 to 12 nucleotides far apart from the 3' terminal nucleotide linked to a first FRET-generating fluorophore polymerase and at least a nucleotide chemically linked to a second FRET-generating fluorophore.

22. A kit for the evaluation of the extension properties of nucleic acid polymerases, comprising at least a nucleic acid having an internal nucleotide located from 1 to 12 nucleotides far apart from the 3' terminal nucleotide linked to a first FRET-generating fluorophore and at least a nucleotide chemically linked to a second FRET-generating fluorophore.

23. A kit for the evaluation of the preference of nucleic acid polymerases for the same nucleotide when labeled or unlabeled, comprising at least a nucleic acid having an internal nucleotide located from 1 to 12 nucleotides far apart from the 3' terminal nucleotide linked to a first FRET-generating fluorophore and at least a nucleotide chemically linked to a second FRET-generating fluorophore.

24. A kit for screening modulators of nucleic acids polymerases comprising at least a nucleic acid having an internal nucleotide located from 1 to 12 nucleotides far apart from the 3' terminal nucleotide linked to a first FRET-generating fluorophore and at least a nucleotide chemically linked to a second FRET-generating fluorophore.