(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 April 2003 (10.04.2003)

PCT

(10) International Publication Number
WO 03/029485 A2

(51) International Patent Classification[7]: C12Q 1/68, C12N 15/10, G06F 19/00

(21) International Application Number: PCT/DK02/00653

(22) International Filing Date: 2 October 2002 (02.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA 2001 01438 2 October 2001 (02.10.2001) DK

(71) Applicant *(for all designated States except US)*: **AZIGN BIOSCIENCE A/S** [DK/DK]; Vestre Teglgade 10, DK-2450 København SV (DK).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **JENSEN, Jens, Bitsch** [DK/DK]; Trøjborggade 7, 5 tv., DK-1757 København V (DK). **THIRSTRUP, Kenneth** [DK/DK]; Frederiksborgvej 141 st.tv., DK-2400 København NV (DK). **HUMMEL, Rene** [DK/DK]; Ellemosevej 73, DK-2900 Hellerup (DK).

(74) Agents: **HALBERG, Kristian** et al.; NeuroSearch A/S, Patent Department, 93 Pederstrupvej, DK-2750 Ballerup (DK).

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

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

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

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

WO 03/029485 A2

(54) **Title:** SPECIFIC DIFFERENTIAL DISPLAY ARRAYS

(57) **Abstract:** The invention relates to a method for identifying a gene fragment that is differentially expressed in a non-stressed and stressed state of a biological source, said method comprising i) identifying a gene that is differentially expressed in a non-stressed and stressed state, and ii) identifying a fragment of the identified gene, said fragment comprising a non-conserved region.

SPECIFIC DIFFERENTIAL DISPLAY ARRAYS

FIELD OF INVENTION

The invention relates to methods for identifying gene fragments and proteins that is differentially expressed, specific differential display arrays, methods for preparation of specific differential display arrays, and kits containing a specific differential display array, and further to the use of specific differential display arrays in methods for the determination of expression profiles in biological materials in which there is an interest in the expression of polynucleotides.

BACKGROUND OF INVENTION

Different kinds of arrays have become increasingly important tools in the biotechnology industry and related fields. These arrays have a plurality of polynucleotide spots deposited on a solid surface in form of an array. Arrays of both polypeptides and polynucleotides have been developed and find use in a variety of applications. One of the applications is differential gene expression, where expression of genes in different cells or tissues (normally a control sample and a sample of the cell or tissue of interest) is compared, and any difference in the mRNA expression profile is determined.

In gene expression analysis using arrays, an array of “probe” nucleotides is contacted with a nucleic acid sample of interest such as mRNA or mRNA concerted into cDNA or cRNA from a particular tissue or cell. Contact is carried out under hybridisation conditions favourable for hybridisation of nucleic acids complementary to the “probe“ nucleotides on the array. Unbound nucleic acid is then removed by washing. The resulting pattern of hybridised nucleic acid provides information regarding the gene expression profile of the sample tested on the array. Gene expression analysis is used in a variety of applications including identification of novel expression of genes, correlation of gene expression to a particular tissue or a particular disease, identifying effects of agents on the cellular expression such as in toxicity testing and in identifying drugs.

A variety of different array techniques have been developed during the years in order to meet the growing demands from the biotechnology industry see, e.g., Lockhart et al., Nature Biotechnology (1996) 14: 1675-1680, Shena et al., Science (1995) 270: 467- 470, WO 98/51789 (Display Systems Biotech ApS) and WO 01/51667 (Integriderm). However, there is still a need for new improved arrays having specific applications.

The present invention provides methods for identifying gene fragments for a completely new array being capable of a more specific determation of the polynucleotide expression profile of polynucleotides in a biological material. Furthermore, the invention provides an array with an improved signal to noise ratio of specific polynucleotide expression profiles.

BRIEF DISCLOSURE OF THE INVENTION

Accordingly, in a first aspect the invention relates to a method for identifying a gene fragment that is differentially expressed in a non-stressed and stressed state of a biological source, said method comprising:

- identifying a gene that is differentially expressed in a non-stressed and stressed state;
- identifying a fragment of the identified gene, said fragment comprising a non-conserved region.

In a further aspect, the invention relates to a specific differential display array comprising a plurality of polynucleotide spots associated with a surface of a solid support, wherein the individual polynucleotide spot comprises a gene fragment as identified by the method according to the present invention.

In still further aspects, the invention relates to a method for preparing a specific differential display array, a method for the determination of an expression profile in a biological material, a method for the determination of a difference in expression profiles from at least a first and a second biological material, a method for identifying a therapeutic, prophylactic and/or toxic agent involved in a direct or indirect action on the expression profile in a biological material, a diagnostic method for the determination of the differences of expression profiles between two biological materials, a database comprising a plurality of gene fragments and gene expression profiles of stressed and non-stressed states, and a kit for use in a hybridisation assay.

Further objects of the invention will be apparent to the person skilled in the art from the following detailed description and examples.

DETAILED DESCRIPTION OF THE INVENTION

Genomics – method of identifying a differentially expressed gene fragment

In a first aspect, the present invention provides a method for identifying a gene fragment that is differentially expressed in a non-stressed and stressed state of a biological source, said method comprising:

- identifying a gene that is differentially expressed in a non-stressed and stressed state;

- identifying a fragment of the identified gene, said fragment comprising a non-conserved region.

In one embodiment of the method for identifying a gene fragment, the identification of a gene that is differentially expressed in a non-stressed and stressed state of a biological source comprises:

- providing a first nucleotide sample from a first biological source being in a non-stressed state;
- providing a second nucleotide sample from a second biological source being in a stressed state;
- identifying a gene that is differentially expressed in the first and second biological source.

In a second embodiment of the method for identifying a gene fragment, the stressed state of the second biological source is directly or indirectly related to a disease state, a chemical treatment, a biological treatment, environmental influence or any other physiological or pathophysiological change, or a combination thereof. In a special embodiment, the disease state is selected from CNS related diseases and treatments, depression, diabetes type II, or obesity. In a further special embodiment, the chemical treatment is treatment with one or more chemicals selected from the group consisting of naturally occurring chemical entities or synthetically derived chemical entities. In a still further special embodiment the chemical treatment is a treatment with one or more antidepressants or CNS drugs.

In a further embodiment of the method for identifying a gene fragment, the biological source is selected from the group consisting of tissues, organs, biological fluids or parts or combinations thereof. In a special embodiment, the first and second biological sources are of the same kind, i.e. of the same kind of origin, such as coming from the same type of tissue, organ, biological fluid or part or combinations thereof, or of the same organism or the same type of organism or the same cell type etc.

In a still further embodiment of the method for identifying a gene fragment, the non-conserved region has less than 90% intraspecies identity. In a special embodiment, the non-conserved region has less than 85% intraspecies identity. In a further special embodiment, the non-conserved region has less than 80% intraspecies identity. In a still further special embodiment, the non-conserved region has less than 75% intraspecies identity.

In a further embodiment of the method for identifying a gene fragment, the non-conserved region has at least 50% interspecies identity. In a special embodiment, the non-conserved region has at least 60% interspecies identity, such as at least 65% interspecies identity. In a further special embodiment, the non-conserved region has at least 70% interspecies identity, such as at least 75% interspecies identity. In a further special embodiment, the non-conserved region has at least 80% interspecies identity, such as at least 85% interspecies identity. In a further special embodiment, the non-conserved region has at least 90% interspecies identity. By the use of such a strategy in which the non-conserved regions are selected on the basis of their interspecies identity, it is possible to use the same type of specific differential display array for detection of expression profiles in several different species as long as the selected non-conserved regions shares such a high degree of homology to enable hybridisation.

In a further embodiment of the method for identifying a gene fragment, the identified fragment has a length of from 25 to 1000 nucleotides. In a special embodiment, the identified fragment has a length of from 25 to 750 nucleotides. In a further special embodiment, the identified fragment has a length of from 25 to 500 nucleotides.

In a still further embodiment of the method for identifying a gene fragment, the step of identifying a gene that is differentially expressed in the first and second biological source comprises using a differential display technique. In a special embodiment, the differential display technique used is restriction fragment differential display. In a further special embodiment, the differential display technique used is the restriction fragment differential display (RFDD) PCR analysis as described in US Patent No. 6,261,770 (Azign Bioscience). In another preferred embodiment, the differential display technique is a microarray analysis using any commercial available microarray based e.g. on oligonucleotides, polynucleotides, cDNA, LNA, PNA or INA. Preferably, the microarray used has 100-100.000 gene fragments (25-1000 nt) attached to the surface. Examples of microarrays suitable for use in the present invention are those disclosed in US 5,445,934 and US 5,744,305.

In a further embodiment of the method for identifying a gene fragment, the step of identifying a fragment of the identified gene comprising a non-conserved region comprises comparing by alignment the identified gene with other intraspecies genes.

Having identified a gene that is differentially expressed in the first and second biological source it is a possibility to identify two or more fragments of the identified gene, all said fragments comprising a non-conserved region.

<u>Proteomics – method of identifying a differentially expressed gene fragment</u>

In a second aspect, the identification of a gene that is differentially expressed in a non-stressed and stressed state of a biological source comprises:

- identifying a protein that is differentially expressed in a non-stressed and stressed state;
- identifying the gene corresponding to the differentially expressed protein; and

- identifying a fragment of the identified gene, said fragment comprising a non-conserved region.

Preferably, the identification of a protein that is differentially expressed in a non-stressed and stressed state comprises

- providing a first protein sample from a first biological source being in a non-stressed state;
- providing a second protein sample from a second biological source being in a stressed state;
- identifying a protein that is differentially expressed in the first and second biological source.

In a second embodiment of the method for identifying a protein, the stressed state of the second biological source is directly or indirectly related to a disease state, a chemical treatment, a biological treatment, environmental influence or any other physiological or pathophysiological change, or a combination thereof. In a special embodiment, the disease state is selected from CNS related diseases and treatments, depression, diabetes type II, or obesity. In a further special embodiment, the chemical treatment is treatment with one or more chemicals selected from the group consisting of naturally occurring chemical entities or synthetically derived chemical entities. In a still further special embodiment the chemical treatment is a treatment with one or more antidepressants or CNS drugs.

In a further embodiment of the method for identifying a protein, the biological source is selected from the group consisting of tissues, organs, biological fluids or parts or combinations thereof. In a special embodiment, the first and second biological sources are of the same kind, i.e. of the same kind of origin, such as coming from the same type of tissue, organ, biological fluid or part or combinations thereof, or of the same organism or the same type of organism or the same cell type etc.

In a still further embodiment of the method for identifying the corresponding gene fragment of the differentially regulated protein, the non-conserved region of the gene fragment has less than 90% intraspecies identity. In a special embodiment, the non-conserved region has less than 85% intraspecies identity. In a further special embodiment, the non-conserved region has less than 80% intraspecies identity. In a still further special embodiment, the non-conserved region has less than 75% intraspecies identity.

In a further embodiment of the method for identifying the corresponding gene fragment, the non-conserved region has at least 50% interspecies identity. In a special embodiment, the non-conserved region has at least 60% interspecies identity, such as at least 65% interspecies identity. In a further special embodiment, the non-conserved region has at least 70% interspecies identity, such as at least 75% interspecies identity. In a further special embodiment, the non-conserved region has at least 80% interspecies identity, such as at least 85% interspecies identity. In a further special embodiment, the non-conserved region has at least 90% interspecies identity. By the use of such a strategy in which the non-conserved regions are selected on the basis of their interspecies identity, it is possible to use the same type of specific differential display array for detection of expression profiles in several different species as long as the selected non-conserved regions shares such a high degree of homology to enable hybridisation.

In a further embodiment of the method for identifying the corresponding gene fragment, the identified fragment has a length of from 25 to 1000 nucleotides. In a special embodiment, the identified fragment has a length of from 25 to 750 nucleotides. In a further special embodiment, the identified fragment has a length of from 25 to 500 nucleotides.

In a still further embodiment of the method for identifying a gene fragment, the step of identifying the protein that is differentially expressed in the first and second biological source comprises using a differential display technique. In a special embodiment, the differential display technique used is MDLC-MSMS (Multi-Dimensional-Liquid-Chromatograhy-Mass-Spectrometry MS), cf. Muticultural Chromatography and the Signature Peptide Approach to Proteomics", Regnier F. et al. LCGC, Vol. 19 number 2, February 2001. In a further special embodiment, the differential display technique used is technologies as the ICAT technology (available from Applied Biosystems), cf. "Quantitative ananlysis of complex protein mixtures using isotope-coded affinity tags", Gygi, S.P. et al. 1999, Nature Biotechnology. Vol.17, p 994-999), Protein 2D-gel electrophoresis and DALPC (Direct Analysis of Large Protein Complexes), cf. "Direct analysis of protein complexes using mass spectrometry"; Link, A.L. et al. 1999. Nature Biotechnology. Vol.17, pp. 676-682).

In a further embodiment of the method for identifying the corresponding gene fragment, the step of identifying a gene fragment of the identified protein comprising a non-conserved region comprises: comparing by alignment the identified gene with other intraspecies genes.

Having identified a protein that is differentially expressed in the first and second biological source it is a possibility to identify two or more gene fragments of the identified protein, all said fragments comprising a non-conserved region.

In a second aspect, the present invention provides a database comprising a plurality of gene fragments as identified by the method as described above. In a special embodiment, the database further comprises additional expressed genes that are substantially relevant to the biological state.

Specific differential display array

In a further aspect, the present invention provides a specific differential display array comprising a plurality of polynucleotide spots associated with a surface of a solid support, wherein the individual polynucleotide spot comprises a gene fragment as identified by the method as described above in the sections “Genomics” and “Proteomics”. In a preferred embodiment, the specific differential display array comprises polynucleotide spots identified partly by the method of described in the section “Genomics” and partly by the method described in the section “Proteomics”.

In one embodiment of the specific differential display array, the individual polynucleotide spot is in single stranded or double stranded form.

In a further embodiment of the specific differential display array, the individual polynucleotide spot is of DNA, RNA, cDNA, natural, synthetic, semisynthetic origin or is a chemical analogous such as LNA, PNA, INA and other MNA’s (Modified Nucleic Acid) with DNA, RNA or protein backbone structure.

In a still further embodiment of the specific differential display array, the solid support is made of a flexible or rigid material.

In a further embodiment of the specific differential display array, the array comprises from 2 to 50,000 polynucleotide spots.

In a still further embodiment, the array comprises two or more spots comprising a gene fragment from the same differentially expressed gene or protein. In a special embodiment, the spots comprising a gene fragment from the same differentially expressed gene comprise the same gene fragment. In a further special embodiment, the spots comprising a gene fragment from the same differentially expressed gene or protein comprise different gene fragments from the same gene. In a special embodiment thereof, all said fragments comprise a non-conserved region. In a further special embodiment thereof, all said fragments are chosen in such a way that they are non-overlapping regions. The number of the gene fragments to be chosen is dependent inter alia on the length of the expressed gene of the differentially expressed gene or protein and the degree of non-conserved regions. The use of two or more gene fragments from the same differentially expressed gene ensures an improved specificity of the array. Furthermore, the use of two or more gene fragments from the same differentially expressed gene will allow a statistical evaluation of the signal from the individual gene fragment.

A specific differential display array according to the invention has a multiplicity of individual polynucleotide spots, stably associated with a surface of a solid support. The specific differential display array comprises spots comprising a polynucleotide composition, wherein the polynucleotide regions within the composition are of known identity, usually of known sequence, as described later on in detail. The polynucleotide spots may be of convenient shape but most often circular, oval or any other suitable shape. The polynucleotide spots may be arranged in any convenient pattern across the surface of the solid support, such as in row or columns to form a grid, in a circular pattern and the like. Preferably the pattern of polynucleotide spots are arranged as a grid to facilitate the evaluation of the results obtained from the analyses in which the specific differential display array is used.

The specific differential display array according to the invention may be of a flexible or rigid solid support and the polynucleotide spots are stably associated thereto. By stably associated is meant that the polynucleotide spots will be associated in their position on the solid support during the analysis in which the specific differential display array is used, such as during different hybridisation, washing and detection conditions. The polynucleotide regions contained in the spots may be covalently or non-covalently associated to the surface of the solid support. Methods how to covalently or non-covalently bind the polynucleotide regions to the surface of the solid support are well known for a person skilled in the art and may be found in Ausubel et al., Current protocols in Molecular Biology, Greene Publishing Co. NY, (1995).

The solid support to which the individual polynucleotide spots are stably associated to is made of a flexible or rigid material. By flexible is meant that the support is capable of being bent or folded without breakage. By rigid is meant that the support is solid and does not readily bend, i.e. the support is not flexible. The support may be fabricated from a variety of materials, including plastics, ceramics, metals, gels, nitrocellulose, nylon, glass and the like.

The array may be produced according to any convenient methodology, such as preparing or obtaining the polynucleotides and then stably associate them with the surface of the support or growing them directly on the support. A number of array configurations and methods for their production are known to those skilled in the art and disclosed in US Patents: 5,445,934, 5,532,128, 5,556,752, 5,242,974, 5,384,261, 5,405,783, 5,412,087, 5,424,186, 5,429,807, 5,436,327, 5,472,672, 5,527,681, 5,529,756, 5,545,531, 5,554,501, 5,561,071, 5,571,639, 5,593,839, 5,599,695, 5,624,711, 5,658,734 and 5,700,637.

The solid support of the invention may have several configurations ranging from a simple to a more complex configuration depending on the intended use of the specific differential display array. The size and thickness of the specific differential display array is not critical as long as the array will function in the expected way and as long as the results obtained after use of the array are not changed. The number and amount of the polynucleotide spots is dependent on the intended use of the arrays as well as the detection system use to determine the expression profile of the biological material being evaluated by the aid of the specific differential display array. The number of the polynucleotide spots may vary from about 2 to about 100,000 such as, e.g., from about 2 to about 50,000, from about 10 to about 25,000, from about 100 to about 10,000, from about 100 to about 5,000, from about 100 to about 1,000, from about 400 to about 600 or about 500 polynucleotide spots, or at least 2 such as, e.g. at least 10, at least 25, at least 50, at least 100, at least 300, at least 400, at least 500 or at least 600 spots, or even more than 100,000 spots. The limitations of the number of the polynucleotide spots are dependent on the way in which the evaluation of the expression profile of the biological material is performed. The amount of the polynucleotide regions present in the polynucleotide spot may vary and the amount will be sufficient to provide adequate hybridisation and detection of the target nucleic acid. Generally the polynucleotides will be present in each spot at a concentration corresponding to an amount of 1 pg – 100 µg or less than 100 µg of the polynucleotide. Normally, only 1 polynucleotide region is present in each spot.

The copy number of the polynucleotide present in each polynucleotide spot will be sufficient to provide enough hybridisation for a target nucleic acid to yield a detectable signal, and generally range from about 50 fmol or less.

Other polynucleotide spots (control spots), which may be present on the specific differential display array, include spots comprising genomic DNA, housekeeping genes, negative and positive control polynucleotides and the like. These polynucleotide spots comprise polynucleotides, which are not unique, i.e they are not polynucleotide regions corresponding to differentially expressed genes. They are used for calibration or as control polynucleotides, and the function of these polynucleotide spots are not to give information of the expression of these polynucleotides, but rather to provide useful information, such as background or basal level of expression to verify that the analysis and the expression profiles obtained are relevant or not. Furthermore these control spots may serve as orientation spots.

The polynucleotide composition also comprises an excipient. Suitable excipients are solvents like e.g. water or any other aqueous medium, pH adjusting agents like buffering agents, stabilising agents, hybridising agents, coloring agents, labelling agents and the like. In general, the excipients used are inert, i.e. they do not have any polynucleotide related effect.

<u>Method of preparing a specific differential display array</u>

In a still further aspect, the present invention provides a method for preparing a specific differential display array as described above comprising

- generating a plurality of compositions each comprising a gene fragment as identified by the method for identifying a differentially expressed gene fragment as described in the sections 'Genomics' and 'Proteomics' above;

- stably associating each composition in one or more individual spots on a surface of a solid support.

In a preferred embodiment of the method, the plurality of compositions comprises compositions identified partly by the method for identifying a differentially expressed gene fragment as described in the section "Genomics" above and partly by the method for identifying a differentially expressed gene fragment as described in the section "Proteomics" above

In a special embodiment of the method for preparing a specific differential display array, the individual gene fragment is produced using one or more primers specific for said gene fragment. In this case the gene fragment preferably has a length of 25-500 nucleotides, more preferably 200-300 nucleotides. A gene fragment of this length has the advantage that it has a high specificity. In another preferred embodiment, the individual gene fragment consists of a synthetically synthesised oligonucletide of length 25-80 nucleotides. A gene fragment of this length has the advantage that it may be prepared automatically by a suitable apparatus.

The specific differential display array may be prepared (produced) using any convenient method and several methods are well known for a person skilled in the art, such as standard procedures according to Sambrook et al., (Molecular cloning: A laboratory manual 2nd edition. Cold Spring Harbour Laboratory Press, New York.).

One means of preparing the array is: i) synthesising or otherwise obtaining the above mentioned gene fragments, ii) preparing the polynucleotide compositions to be used in each spot and then iii) depositing in the form of spots the polynucleotide compositions comprising the gene fragment onto the surface of the solid support. The gene fragments may be of DNA, RNA, cDNA, natural, synthetic, semisynthetic origin or chemical analogous such as LNA, INA or PNA or any other molecule with a DNA, RNA or protein backbone. The gene fragments may be obtained from any biological material such as, e.g., tissues or cells and/or produced by a cell culture. The biological material may be an organism, such as a microorganism, plant, fungus (e.g. yeast or mushrooms) or animal.

The gene fragments may be prepared using any conventional methodology such as automated solid phase synthesis protocols, PCR using one or more primers specific for the gene fragments and the like. In general, PCR is advantageous in view of the large numbers of gene fragments that must be generated for specific differential display array. The amplified non-conserved gene fragments may further be cloned in any suitable plasmid vector to enable multiplication and storage of the amplified non-conserved gene fragments.

The prepared non-conserved gene fragments may be spotted onto the solid support using any convenient methodology, including manual and automated techniques, e.g. by micro-pipette, ink jet pins etc. and any other suitable automated systems. An example of an automated system is the automated spotting device, Beckman Biomek 2000 (Beckman Instruments). The ready arrays may then be stored at suitable conditions until use.

Fragment that is differential expresses and comprises a non-conserved region

The fragments for use in the specific differential display array according to the invention are unique by being fragments from genes differentially expressed in a stressed and non-stressed state and comprising a non-conserved region.

Information on genes that are differentially expressed in a stressed and non-stressed state may be obtained in a number of different ways.

Scientific literature may be consulted in order to identify differentially expressed genes. Furthermore, electronic databases – such as Entrez, GenBank, SwissProt - may be consulted.

Moreover, the information may be obtained experimentally. In a special embodiment, the identification of a gene that is differentially expressed in a non-stressed and stressed state of a biological source comprises:

- providing a first nucleotide sample from a first biological source being in a non-stressed state;
- providing a second nucleotide sample from a second biological source being in a stressed state;
- identifying a gene that is differentially expressed in the first and second biological source.

The identification of a gene that is differentially expressed in the first and second biological source may be performed by a number of ways known by a person skilled in the art. Examples are the use of a differential display technique, such as restriction fragment differential display (RFDD-PCR), northern blot analysis, microarray analysis and Protein Differential Display analysis, such as MDLC-MSMS and 2D-gel electrophoresis.

Method for the determination of an expression profile

In a further aspect, the present invention provides a method for the determination of an expression profile in a biological material, said method comprising:

- obtaining a polynucleotide sample from the biological material,
- labelling said sample to obtain a labelled target polynucleotide sample
- contacting at least one labelled target polynucleotide sample with a specific differential display array as described above under conditions which are sufficient to produce a hybridisation pattern, and

- detecting said hybridisation pattern to obtain the expression profile of the biological material.

In a still further aspect, the present invention provides a method for the determination of a difference in expression profiles from at least a first and a second biological material, said method comprising:

- obtaining a first expression profile of the first biological material according to the method for the determination of an expression profile in a biological material as described above,
- obtaining a second expression profile of the second biological material according to the method for the determination of an expression profile in a biological material as described above,
- comparing the first and the second expression profiles to identify any difference in the expression profiles from the first and second biological material.

In one embodiment of the method for the determination of a difference in expression profiles, the first and the second biological material are of the same kind of biological material.

In a second embodiment of the method for the determination of a difference in expression profiles, the first biological material is in a non-stressed state and the second biological material is in a stressed state. In a special embodiment, the stressed state of the second biological material is directly or indirectly related to a disease state, a chemical treatment, a biological treatment, environmental influence or any other physiological or pathophysiological change, or a combination thereof. In a further special embodiment, the chemical treatment is treatment with one or more chemicals selected from the group consisting of naturally occurring chemical entities or synthetically derived chemical entities. In a still further special embodiment, the disease state is selected from depression, CNS related diseases, diabetes type II, or obesity.

In a third embodiment of the method for the determination of a difference in expression profiles, the stressed state of the second biological material is directly or indirectly related to the stressed state of the second biological source used in the method for identifying a gene fragment as described above.

In a fourth embodiment of the method for the determination of a difference in expression profiles, each material is labelled with a unique label (e.g. Cy3 and Cy5 for each sample, respectively).

In a further aspect, the present invention provides a method for identifying a therapeutic, prophylactic and/or toxic agent involved in a direct or indirect action on the expression profile in a biological material, said method comprises:

- obtaining a first expression profile of a first biological material according to the method for the determination of an expression profile in a biological material as described above,
- obtaining a second expression profile of a second biological material according to the method for the determination of an expression profile in a biological material as described above,
- applying a test compound to the second biological material and obtaining a third expression profile thereof according to the method for the determination of an expression profile in a biological material as described above,
- comparing the first, second and third expression profiles, and
- identifying any differences in the expression profiles so as to identify any biological response of the test compound on the expression profile.

In one embodiment of the method for identifying a therapeutic, prophylactic and/or toxic agent, the method further comprises:

- applying a test compound to the first biological material and obtaining a fourth expression profile thereof according to the method for the determination of an expression profile in a biological material as described above,
- comparing the first, second, third and fourth expression profiles, and
- identifying any differences in the expression profiles so as to identify any biological response of the test compound on the expression profile.

In a second embodiment of the method for identifying a therapeutic, prophylactic and/or toxic agent, the first and second biological material are of the same kind of biological material.

In a further embodiment of the method for identifying a therapeutic, prophylactic and/or toxic agent, the first biological material is in a non-stressed state and the second biological material is in a stressed state.

In a still further embodiment of the method for identifying a therapeutic, prophylactic and/or toxic agent, the expression profile of the second biological material is direct or indirect measure of a disease state, a chemical treatment, a biological treatment, environmental influence or any other physiological or pathophysiological change, or a combination thereof.

In a further embodiment of the method for identifying a therapeutic, prophylactic and/or toxic agent, the test compound is a chemical or biological derived compound such as compounds selected from the group consisting of therapeutic, prophylactic and/or toxic chemical entities, physiologically chemical entities, hormones, vitamins, nutrients, pesticides, fungicides, bacteriocides and any other organic chemical entity.

The method according to the above embodiment of the invention is used to identify potential therapeutic, prophylactic and/or toxic agents useful for the treatment of diseases caused by an alteration in the expression profile of the polypeptides. One example is the use of a biological model, such as a rat model in which a first, second, third and/or fourth group are used. The first and third group is non-stressed and the second and fourth group stressed in such a way that the expression of one or more polynucleotides are influenced in such as way that an increase or a decrease of the gene expression is obtained. The third and fourth groups are treated with a test compound.

Determination of expression profiles typically means determination of the expression level of multiple mRNAs, all of them corresponding to the spotted gene fragments on the array. The detection limit of the expression level of a mRNA may be approximately 0.2 ng or less of total RNA of the biological material used to hybridise each polynucleotide spot.

The expression profiles can be produced by any means known in the art, including but not limited to the methods disclosed by: Liang et al., (1992) Science 257: 967-971; Ivanova et al., (1995) Nucleic Acids Res 23: 2954-2958; Guilfoyl et al., (1997) Nucleic Acids Res 25(9): 1854-1858; Chee et al., (1996) Science 274: 610-614; Velculescu et al., (1995) Science 270: 484-487; Fiscker et al., (1995) Proc Natl Acad Sci USA 92(12): 5331.5335; and Kato (1995) Nucleic Acids Res 23(18): 3685-3690.

According to one embodiment of the invention the specific differential display array will be used for the evaluation of the expression profile of one or more biological materials or a mixture of biological materials. The method for the determination of an expression profile in a biological material or in a mixture of biological materials comprises obtaining a polynucleotide from the biological material(s), labelling said polynucleotide to obtain a labelled target polynucleotide sample, contacting at least one labelled target polynucleotide sample with an array as defined above under conditions which are sufficient to produce a hybridisation pattern and detecting said hybridisation pattern to obtain the polynucleotide expression profile of the biological material or the mixture of biological materials. The expression profile in the biological material may be determined to correspond to the expression of specific genes. The biological material or the mixture of biological materials may be in a non-stressed or a stressed stage. The stress may directly or indirectly influence the expression profile and thereby the polynucleotides identified which react upon that type of stress.

The stress may be caused by any disease or condition. The examples of the present application show the non-limiting diseases depression, obesity and diabetes (diabetes type II).

The analysis of the expression profile includes several steps of procedures in which well known techniques are used, such as those mentioned in Sambrook et al., Molecular Cloning: A Laboratory approach, Cold Spring Harbour Press, NY (1987), and in Ausubel et al., Current protocols in Molecular Biology, Greene Publishing Co. NY, (1995).

The biological material to be evaluated needs to be identified and isolated such as e.g. described in Example 9. For the ability to perform the analysis cDNA are generally produced from isolated total RNA or polyA RNA (mRNA). The total RNA/mRNA can be isolated using a variety of techniques. Numerous techniques are well known (see Sambrook et al., Molecular Cloning: A Laboratory approach, Cold Spring Harbour Press, NY (1987), and Ausubel et al., Current protocols in Molecular Biology, Greene Publishing Co. NY, (1995)). In general, these techniques include a first step of lysing the cells and then a second step of enriching for or purifying RNA.

The isolated total RNA/mRNA are reversed transcribed using a RNA-directed DNA polymerase, such as "reverse transcriptase" isolated from such retroviruses as AMV, MoMuLV or recombinantly produced. Many commercial sources are available (e.g., Perkin Elmer, New England Biolabs, Stratagene Cloning Systems).

Preferably the mRNA is reversed transcribed into cDNA and at the same time a label is incorporated for later detection of the hybridised amplified products on the array. The amplification by PCR may be performed according to Example 2. The label may vary dependent on the system to be used for the detection and several labels are well known in the area of molecular biology (e.g. radioactive labels, fluorescent labels, coloring labels, chemical labels etc.)

The labelled cDNA is then denaturated and used for hybridisation on the array. The hybridisation conditions vary and are dependent on the aim with the expression profile obtained after the hybridisation. One example is found in Example 6. After hybridisation of the labelled cDNA, the specific differential display array is washed to remove the cDNA which have not hybridised to the fragments and the hybridised labelled cDNA are detected by a suitable means and an expression profile obtained.

As known by a person skilled in the art, most diseases or conditions might influence the expression profile of the second biological material. As an example the use of *in vivo* models such as, e.g., a rat model in which at least a first and a second experimental group are used. The first group is non-stressed and the second group stressed in such a way that the expression of one or more of the polynucleotides on the specific differential display array are influenced in such as way that an increase or a decrease of the expression is obtained, when the expression profiles are analysed using array and the method according to the invention. The second group may be either permanently stressed or stressed during a certain period of time and after the period of stress one or more biological materials obtained from the second group and the expression profile determined.

Diagnostic methods and kits

In a still further aspect, the present invention provides a diagnostic method for the determination of the differences of expression profiles between two biological materials, said method comprises:

- obtaining a first expression profile of a first biological material according to the method for the determination of an expression profile in a biological material as described above,
- obtaining a second expression profile of a second biological material according to the method for the determination of an expression profile in a biological material as described above,
- comparing the first and second expression profile, and identifying any difference in the expression profiles.

In one embodiment of the diagnostic method, the expression profile from more than two different biological materials are compared, such as biological materials, which are in different stages of a disease.

Preferably the difference between the first expression profile and the second expression profile is directly or indirectly influenced by a disease. Furthermore, the expression profile from more than two different biological materials are compared, such as biological materials, which are in different stages of a disease. The diagnostic method may be useful in the determination of diseases directly or indirectly caused by different expression profiles and by the use of such a method there will be an enhanced possibility to start the treatment of the disease at an early stage of the disease.

In a further aspect, the present invention provides a kit for use in a hybridisation assay, said kit comprising a specific differential display array as described above.

In one embodiment of the kit, the kit further comprises reagents for generating a labelled target polynucleotide sample.

In a second embodiment of the kit, the kit further comprises a hybridisation buffer.

The kit may be used according to the above-mentioned methods for the determination of expression profiles in a biological material as defined above.

Database

In a further aspect, the present invention provides a database comprising a plurality of gene fragments as identified by the method as described above. In a special embodiment, the database further comprises additional expressed genes that are substantially relevant to the stressed state of the biological source.

In a preferred embodiment the database of the invention contains one or more of the following items of information obtained in gene and protein expression analysis:

1) Genes and proteins identified as differentially expressed (their accession number and sequences)
2) Method used to identify these genes and proteins e.g. RFDD, Microarray analysis, MDLC-MSMS and Protein 2D-gel electrophoresis.
3) The sequence of the gene fragments identified to go onto the Specific Differential Display Array, cf. Example 4.
4) Information about the biological material used for the analysis, e.g. mouse brain.
5) The gene and protein expression profiles performed on the biological materials.
6) Information about the treatment of the biological source, e.g. mouse treated with the 10mg/kg antidepressant fluoxetine one time/day for 21 days.
7) Drug information: Drug name, Dose usage, chemical structure, mechanism of action, Drug Interaction, Clinical Trial information, Indications and usage, Contradictions, Drug/Laboratory test interactions, receptor binding studies results, ADME results, PK results, toxicological profile

In a further preferred embodiment, the database of the invention further contains the following items of information obtained in analysis on the Specific Differential Display Arrays:

1) Genes identified (their accession number and sequence) and spotted onto the Specific Differential Display Array
2) Method used to identify these genes e.g. RFDD, Microarray analysis, MDLC-MSMS and Protein 2D-gel electrophoresis.
3) The sequence of the gene fragments identified to go onto the Specific Differential Display Array, cf. example 4.
4) Information about the biological material used for the analysis, e.g. mouse brain.
5) The gene expression profiles performed on the biological material using the Specific Differential Display Array.
6) Information about the treatment of the biological source, e.g. mouse treated with the 10 mg/kg antidepressant fluoxetine one time/day for 21 days.
7) Drug information: Drug name, Dose usage, chemical structure, mechanism of action, Drug Interaction, Clinical Trial information, Indications and usage, Contradictions, Drug/Laboratory test interactions, receptor binding studies results, ADME results, PK results, toxicological profile

The database of the invention may be used as follows: A gene expression profile is performed on a biological source, e.g. brain tissue, from a treatment of a test animal, e.g. a mouse, with a Novel Chemical Entity (NCE) on a Specific Differential Display Array. This gene expression profile will then be stored in the database and use for a complex datamining against all the gene profiles stored in the database resulting in the possibility to elucidate e.g. potential therapeutic and toxicological effect of a new tested drug and suggesting new modification of this NCE to make it a more efficient NCE towards this therapeutic field.

Disease states

In connection with all aspects of the present invention, the disease state may be any disease state giving rise to a differential expression of one or more genes. Examples of such disease states are:

Adenoma
Amnesia
Arthritis, Rheumatoid (Musculoskeletal Diseases)
Arthritis
Anxiety
Asthma
Autoimmune Diseases
Autonomic Nervous System Diseases
Breast Neoplasms
Bright's Disease (Glomerulonephritis)
Bronchiectasis
Bronchiolitis
Bronchiolitis Obliterans Organizing Pneumonia (BOOP)
Bronchitis
Bronchopulmonary Dysplasia
Buruli Ulcer (Mycobacterium Infections)
Cancer (Neoplasms)
Carcinoid Tumor
Carcinoma
Carcinoma, Merkel Cell
Cardiomyopathy, Hypertrophic
Carnitine Palmitoyl Transferase Deficinecy
Cartilage Diseases
Central Cord Syndrome
Central Core Disease (Myopathies, Structural, Congenital)
Central Pain Syndrome (Thalamic Diseases)
Cerebellar Ataxia

Cerebral Hemorrhage
Cerebral Palsy
Cerebrovascular Disorders
Cervix Dysplasia
Creutzfeldt-Jakob Syndrome
Crohn Disease
Dementia
Demyelinating Diseases
Depressive Disorder
Depression
Dermatitis
Dermatitis, Exfoliative
Dermatitis Herpetiformis
Dermatofibroma
Dermatomycoses
Dermatomyositis
Dermoid Cyst
Desmoplastic Small Round Cell Tumor (Carcinoma, Small Cell)
Diabetes, Bronze (Hemochromatosis)
Diabetes, Gestational
Diabetes Insipidus
Diabetes Insipidus, Nephrogenic
Diabetes Mellitus
Diabetic Ketoacidosis
Diabetic Neuropathies
Diabetic Retinopathy
Diffuse Cerebral Sclerosis of Schilder
Digestive System Neoplasms
Drug Toxicity
Epilepsy
Fatty Acid Oxidation Disorders
Gastroenteritis
Glaucoma
Glioblastoma
Hemangioma
Hematoma, Epidural
Hematoma, Subdural
Hepatitis

Hepatitis, Chronic
Inflammatory Bowel Diseases
Kidney Tubular Necrosis, Acute
Kidney Diseases
Lennox-Gastaut Syndrome (Epilepsy)
Lennox-Gastaut Syndrome (Epilepsy)
Leukemia
Liver Diseases
Lung Cancer (Thoracic Neoplasms)
Lung Diseases, Interstitial
Migraine
Motor Neuron Disease
Multiple Sclerosis
Myocardial Infarction
Myocardial Ischemia
Myocarditis
Nervous System Neoplasms
Obesity
Osteoarthritis
Osteochondritis
Osteoporosis
Pancreatic Neoplasms
Pancreatitis
Parathyroid Diseases
Parkinson Disease
Psoriasis
Retinoblastoma
Sarcoidosis
Sarcoma
Schizophrenia
Seizures
Short Bowel Syndrome
Sjogren's Syndrome
Skin Diseases, Infectious
Skin Neoplasms
Skin Ulcer
Sleep Disorders
Spasm

Spasms, Infantile (West Syndrome)
Spinal Diseases
Stress
Stress Disorders, Post-Traumatic
Stroke (Cerebrovascular Disorders)
Tremor
Urinary Retention
Migraine
Cerebrovascular Disorders *(Stroke)*
Hypoxia, Brain
Brain Ischemia
Ischemic Attack, Transient
Epilepsy
Spasms, Infantile
Dementia
Alzheimer Disease
Parkinsonian Disorders
Parkinson Disease
Multiple System Atrophy

In a special embodiment of the invention, the disease state is a CNS related disease selected from the group consisting of Asthma, cystic fibrosis, chronic obstructive pulmonary disease and rhinorrhea, convulsions, vascular spasms, coronary artery spasms, renal disorders, polycystic kidney disease, bladder spasms, urinary incontinence, bladder outflow obstruction, irritable bowel syndrome, gastrointestinal dysfunction, secretory diarrhoea, ischaemia, cerebral ischaemia, ischaemic hearth disease, angina pectoris, coronary hearth disease, traumatic brain injury, psychosis, anxiety, depression, dementia, memory and attention deficits, drug addiction and/or abuse, including cocaine or tobacco abuse, Parkinson's disease, Alzheimer's disease, dysmenorrhea, narcolepsy, Reynaud's disease, intermittent claudication, Sjorgren's syndrome, migraine, arrhythmia, hypertension, absence seizures, myotonic muscle dystrophia, xerostomi, diabetes type II, hyperinsulinemia, premature labour, baldness, cancer, schizofrenia or psychosis.

Further, CNS related diseases according to the present invention include a variety of disorders associated with the neural system, for example eating disorders, obsessive compulsive disorders, panic disorders, alcoholism, pain, memory deficits and anxiety. Included among these disorders are disorders such as pseudodementia or Ganser's syndrome, migraine pain, bulimia, obesity, pre-menstrual syndrome or late luteal phase syndrome, post-traumatic syndrome, memory loss, memory dysfunction, social phobia, attention deficit hyperactivity disorder, chronic fatigue syndrome, premature ejaculation, erectile difficulty, anorexia nervosa, disorders of sleep, autism, mutism, trichotillomania or mood syndrone.

Also, according to the present invention CNS related diseases include auto-immune diseases, e.g. Addison's disease, alopecia areata, Ankylosing spondylitis, haemolytic anemia (anemia haemolytica), pernicious anemia (anemia perniciosa), aphthae, aphthous stomatitis, arthritis, arteriosclerotic disorders, osteoarthritis, rheumatoid arthritis, aspermiogenese, asthma bronchiale, auto-immune asthma, auto-immune hemolysis, Bechet's disease, Boeck's disease, inflammatory bowel disease, Burkitt's lymphoma, Chron's disease, chorioiditis, colitis ulcerosa, Coeliac disease, cryoglobulinemia, dermatitis herpetiformis, dermatomyositis, insulin-dependent type I diabetes, juvenile diabetes, idiopathic diabetes insipidus, insulin-dependent diabetes mellisis, auto-immune demyelinating diseases, Dupuytren's contracture, encephalomyelitis, encephalomyelitis allergica, endophthalmia phacoanaphylactica, enteritis allergica, auto-immune enteropathy syndrome, erythema nodosum leprosum, idiopathic facial paralysis, chronic fatigue syndrome, febris rheumatica, glomerulo nephritis, Goodpasture's syndrome, Graves' disease, Hamman-Rich's disease, Hashimoto's disease, Hashimoto's thyroiditis, sudden hearing loss, sensoneural hearing loss, hepatitis chronica, Hodgkin's disease, haemoglobinuria paroxysmatica, hypogonadism, ileitis regionalis, iritis, leucopenia, leucemia, lupus erythematosus disseminatus, systemic lupus erythematosus, cutaneous lupus erythematosus, lymphogranuloma malignum, mononucleosis infectiosa, myasthenia gravis, traverse myelitis, primary idiopathic myxedema, nephrosis, ophthalmia symphatica, orchitis granulomatosa, pancreatitis, pemphigus, pemphigus vulgaris, polyarteritis nodosa, polyarthritis chronica primaria, polymyositis, polyradiculitis acuta, psoreasis, purpura, pyoderma gangrenosum, Quervain's thyreoiditis, Reiter's syndrome, sarcoidosis, ataxic sclerosis, progressive systemic sclerosis, scleritis, sclerodermia, multiple sclerosis, sclerosis disseminata, acquired spenic atrophy, infertility due to antispermatozoan antibodies, thrombocytopenia, idiopathic thrombocytopenia purpura, thymoma, acute anterior uveitis, vitiligo, AIDS, HIV, SCID and Epstein Barr virus associated diseases such as Sjorgren's syndrome, virus (AIDS or EBV) associated B cell lymphoma, parasitic diseases such as Lesihmania, and immunosuppressed disease states such as viral infections following allograft transplantations, graft vs. Host syndrome, transplant rejection, or AIDS, cancers, chronic active hepatitis diabetes, toxic chock syndrome, food poisoning, and transplant rejection.

Definitions

In the context of the present application and invention the following definitions apply:

The term "polynucleotide" is intended to mean a single or double stranded polymer composed of nucleotides, e.g. deoxyribonucleotides and/or ribonucleotides from about 10 to about 9,000 nucleotides in length, such as from about 10 to about 6,000, from about 10 to about 3,000, from about 10 to about 1,500, from about 10 to about 1,000, from about 25 to about 1,000, or from about 25 to about 750.

The term “complementary” or “complementarity” is used in relation to the base-pairing rules of nucleotides well known for a person skilled in the art. Polynucleotides may be complete or partial complementary. Partial complementarity means that at least one nucleic acid base is not matched according to the base pairing rules. Complete complementarity means that all nucleotides in a polynucleotide match according to the base pairing rules. The degree of complementary between polynucleotides affects the strength of hybridisation between two polynucleotide strands. The inhibition by hybridisation of the complementary polynucleotide to the target polynucleotide may be analysed by techniques well known for a person skilled in the art, such as Southern blot, Northern blot, and the like under conditions of high stringency. A partially (substantially) homologous polynucleotide will compete for and inhibit the binding of a completely homologous sequence to the target sequence under low stringency.

The term “homology” is intended to mean the degree of identity of one polynucleotide to another polynucleotide. According to the invention the term homology is used in connection with complementarity between polynucleotides within a family or between species. There may be complete homology (i.e. 100% identity) between two or more polynucleotides. The degree of homology may be determined by any method well known for a person skilled in the art.

The term "polynucleotide composition" is intended to mean a composition comprising a polynucleotide together with an excipient. The polynucleotide compositions are applied as spots on the array. The polynucleotide composition comprises a non-conserved region of a polynucleotide family member. The term “polynucleotide composition” includes also control or calibrating compositions such as, e.g. compositions comprising polynucleotides corresponding to housekeeping genes.

The term "non-conserved region" is intended to mean a segment of nucleotides in a polynucleotide that - compared to a segment of nucleotides in another polynucleotide originating from the same species. i.e. intraspecies - has at the most 90% identity.

The term “intraspecies gene” is intented to mean a gene or polynucleotide arising from the same species, such as such humans, mice or rats. The publicly known genes and polynucleotides may generally be found and downloaded from Genbank or EMBL (www.ncbi.nih.org).

The term "intraspecies identity" is intended to mean identity within a group of members belonging to the same species such as such humans, mice or rats (paralogy).

The term "interspecies identity" is intended to mean identity between a group of different species, such as a group comprising humans, mice and rats (orthology).

The terms "expression profile", “differential expression profile” or “gene expression profile” are intended to mean the expression of the mRNAs in a biological material. While an expression profile encompasses a representation of the expression level of at least one mRNA, in practice the typical expression profile represents the expression of several mRNAs. For example, an expression profile used according to the present invention represents the expression levels of at least from 1 to 100,000 or more different mRNAs in a biological material. The expression level of the different mRNAs is the same or different. The expression of mRNAs may be up- or down regulated resulting in different expression profiles.

The term "biological material" include within its meaning organisms, organs, tissues, cells or biological material produced by a cell culture. The biological material may be living or dead. The material may correspond to one or more cells from the organisms, in case the organism is a multicellular organism, the biological material may correspond to one or more cells from one or more tissues creating the multicellular organism. The biological material to be used according to the invention may be derived from particular organs or tissues of the multicellular organism, or from isolated cells obtained from a single or multicellular organism.

The term "biological source" include within its meaning organisms, organs, tissues, cells or biological material produced by a cell culture. The biological material may be living or dead. The material may correspond to one or more cells from the organisms, in case the organism is a multicellular organism, the biological material may correspond to one or more cells from one or more tissues creating the multicellular organism. The biological material to be used according to the invention may be derived from particular organs or tissues of the multicellular organism, or from isolated cells obtained from a single or multicellular organism.

In obtaining the sample of nucleotides to be analysed from the biological material or biological source from which it is derived, the biological material or biological source may be subject to a number of different processing steps. Such steps might include tissue homogenisation, cell isolation and cytoplasma extraction, nucleic acid extraction and the like and such processing steps are generally well known for a person skilled in the art. Methods of isolating RNA from cells, tissues, organs or whole organisms are known to those skilled in the art and are described in Maniatis et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbour Press) (1989).

The term "organism" is intended to mean any single cell organism such as yeast or multicellular organism, including plants, fungi and animals, preferably mammals, such as humans, rats, pigs, cows, horses, dogs, guinea pigs, ferrets, rabbits, sheep, apes, monkeys and cats.

The term “target polynucleotides” is intended to mean polynucleotides present in the biological material of interest. If the target polynucleotide has a complementary polynucleotide present on the specific differential display array, it will hybridise thereto and thus give rise to a detectable signal.

The term “non-overlapping” is intended to mean that when the specific fragments used in the polynucleotide composition spots are obtained from the same polynucleotide, the regions are obtained from different parts of the polynucleotide and the different parts are located in such a manner that the regions not even overlap each other by a single nucleotide. In a polynucleotide of e.g. 1,000 nucleotides the regions 1-500 and 501-900 are non-overlapping. The non-overlapping polynucleotide regions may be located with a distance of one or more nucleotides from each other.

The term “primer” is intented to mean a polymer of 3-50 nucleotides.

The term “set of primers” is intended to mean one or more primers having the ability to amplify a polynucleotide region under suitable conditions. The length of the primers may be the same or different and dependent on the character of the polynucleotide region to be amplified. Design of such a set of primers is well known for a person skilled in the art. The set of primers having a sufficient length to specifically hybridise to a distinct polynucleotide in the sample and the length of the primers will be from about 3 to 50 nucleotides.

The term "stressed state and stressed" is intended to mean that the above described “biological material” and “biological source” is influenced compared to the normal condition. When an expression profile is obtained from a stressed biological material or source it is different compared to a non-stressed biological material or source. The biological material or source may be influenced by some kind of organic/inorganic compound, an environmental agent, a drug substance, pathogen, mutagen, mitogen, receptor mediated signal or the like. Normally, the biological material is influenced in such a manner that the expression profile of the polynucleotides in the biological material or source either directly or indirectly is affected resulting in at least one difference between the expression profile of the non-stressed biological material or source compared to the stressed biological material or source.

Methods of comparing the homology between different polynucleotides and/or parts of different polynucleotides are well known in the art. The polynucleotides may either belong to the same family or different families and/or being polynucleotides encoding the same polypeptide from the same or different species. Optimal alignment of nucleotides of a polynucleotide for comparison of the homologies may be conducted using the homology algorithm (Smith and Waterman, Adv. Appl. Math. 2: 482 (1981)), by the homology alignment algorithm of Needleman and Wunsch, J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson and Lipman, Proc. Natl. Acad. Sci., USA 85: 2444 (1988), by computerised implementations of these algorithms by using for example CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, California, GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG) 575 Science Dr., Madison, Wisconsin, USA. The above-mentioned algorithms and computer implementations should be regarded as examples and the invention not limited thereto.

EXAMPLES

EXAMPLE 1

Obtaining total RNA from a biological source

A. Obtaining total RNA from antidepressant treated mice

Male NMRI mice (Age: 12-14 weeks) were treated for 21 days with fluoxetine (10 mg/kg) or clomipramine (15 mg/kg) receiving one intraperitoneal injection per day. Control animals were treated with 0.9% saline in a similar fashion. On day 22, 24 hrs after the last injection the mice were sacrificed by cervical dislocation and the pituitary glands were rapidly isolated and snap-frozen in liquid nitrogen. Total RNA was purified from pituitary glands using TRI reagent according to standard procedure (Chomczynski and Sacchi, 1987).

B. Obtaining total RNA from the db/db and ob/ob mice

12 adult male C57BL/6J mice (X 6 weeks old; 22 ± 3 g; Mean ± SEM), db/db mice (6 weeks old 57 ± 3 g; Mean ± SEM), and ob/ob mice (6 weeks old; 57 ± 2 g; Mean ± SEM) (obtained from Møllegården, Ll. Skensved, Denmark and Charles River, Hamburg, Germany) were used in the experiment. Upon arrival at the laboratory animals were housed four per cage. The animals were kept under standard laboratory conditions (12:12 h light/dark cycle; lights on at 06.00 h) in a humidity- and

temperature controlled room with free access to standard food and water for another 6 weeks.

Food intake is measured every day. Body weight is measured every second day. After acclimatisation for 6 weeks, when the animals had reached an age of 12 weeks, days, food was withdrawn from all animals at 06.00 h. The animals were fasted overnight and decapitated. The liver was rapidly removed and thigh muscle, pancreas, white adipose tissue (omental fat) and hypothalumus were isolated and immediately frozen individually in dry ice and liquid nitrogen. The tissue was kept at -80°C until extraction.

The RNA was extracted using the phenol-chloroform extraction method (Chomczynski and Sacchi, 1987) from livers from three animals. The RNA was evaluated on a gel, and only RNA showing distinct bands on this gel was used for further analysis. Further, the RNA concentration was measured using OD. The RNA was then pooled from two of the animals.

EXAMPLE 2a

Identification of differentially regulated gene fragments using the RFDD-PCR technology.

Restriction fragment differential display analysis was carried out as described in US 6,261,770 (Azign Bioscience A/S). Five hundred (500) ng of total RNA isolated from mice treated as stated in Example 1aand 1b(pooled from 3 animals) was reverse-transcribed using a poly-dT primer, (T25V from DNA Technology, Århus, Denmark), and 100 units Reverse Transcriptase in 25 µl volume. The reaction was incubated at 42°C for 2 hours. For second-strand synthesis, a 50 µl mix consisting of buffer, dNTPs, 12 units DNA Polymerase I, and 0.8 units RNase H was added. The reaction mixture was incubated at 16°C for 2 hours, and subsequently phenol:chloroform extracted, precipitated by ethanol and dissolved in 20 µl H_20. Ten (10) µl cDNA was checked on an agarose gel for cDNA smear between 100 and 2000 bp.

The remaining 10 µl cDNA was digested with TaqI endonuclease at 65°C for 2 hours. The restriction fragment differential display PCR template was completed by ligating the digest with a standard adaptor and an EP-adaptor containing an extension protection group (EPG) at 37°C for 3 hours using T4 DNA ligase. The incubation was carried out at 37°C and not the standard 16°C to maintain the TaqI endonuclease active during the ligation. This prevented religation of cDNA TaqI-fragments but not adaptor ligation since the latter did not re-establish the TaqI site.

The template was PCR amplified using a 0-extension primer complementary to the EP-adaptor in combination with a 3-extension primer recognizing the standard adaptor and the three nucleotides adjacent to the TaqI site. The 0-extension primer was kinase-labelled with [^{33}P]dATP (3000 Ci mmol-1, equivalent to 1.11 x 1014 Bq mmol-1, ICN). All PCR reactions were carried out in 20 µl volume using 0.2 µl template and standard concentration of dNTPs and primers, using the following PCR-amplification profile: precycle 94 °C, 1 min., for the first 10 cycles: 94°C, 30 sec; 60°C with touchdown by 0.5°C for each cycle until 55°C is reached; 72°C, 1 min, for the last 25 cycles: 94°C, 30 sec; 55°C, 30 sec; 72°C, 1 min.

Samples from control and test animals were analysed in parallel and resolved on a standard 6% polyacrylamide sequencing gel by the ALF Express DNA sequencer (Amersham Pharmacia Biotech) and dried down onto Whatman paper. In addition a radio-labelled size marker (Research Genetics) was included on all gels.

Bands corresponding to differentially expressed genes (between samples from control and test animals) were identified and marked on the autoradiogram. The dried gel was lined up with markings on the film and the fragments were excised from the gel. The gene fragments were eluted (in TE buffer), and re-amplified using the same PCR conditions and primers as in the initial PCR reaction.

EXAMPLE 2b

Identification of differentially regulated gene fragments using the microarray technology.

Preparation of labelled samples

RNA from the biological material is obtained according to the procedures of Examples 1A or 1B

The RNA was preciptated by centrifugation and the RNA pellet was in 70% ethanol. The RNA was precipitated at 15,000 x g for 15 min. The supernatant was discarded and the pellet air-dried. The RNA was adjusted to a concentration of 1µg/µl with DEPC-H2O. In 2 separate tubes 25 µl total RNA (1µg/µl) was placed and 7 µl DEPC treated H2O added. 4 µl of oligo-dT (e.g. T25V primer) (1 µg/µl) was added to each tube. The tubes were incubated in a Thermal cycler at 65°C for 3 min.

Tube 1 was prepared by adding, 5 µl 10 x cDNA Buffer (500mM Tris-HCl, pH 8.3; 800 mM KCl; 100 mM MgCl2; 40mM DTT), 2 µl Cy3-dUTP (1mM, Cat. No. PA53022, Amersham Pharmabiotech), 5 µl 10 x dNTP (5mM dATP; 5 mM dCTP; 5 mM dGTP; 5mM dTTP). The contents were mixed and added 2 µl reverse transcriptase (100 U/µl). Tube 2 was prepared by adding, 5 µl 10 x cDNA Buffer (500mM Tris-HCl, pH 8.3; 800 mM KCl; 100 mM MgCl2; 40mM DTT), 2 µl Cy5-dUTP (1mM, Cat. No. PA55022, Amersham Pharmabiotech), 5 µl 10 x dNTP (5mM dATP; 5 mM dCTP; 5 mM dGTP; 5mM dTTP). The contents were mixed and added 2 µl reverse transcriptase (100 U/µl). Incubate tube 1 and 2 for 42 °C for 60 min and at

65°C for 15 min. The temperature was decreases to 42°C. Then added reverse transcriptase (100 U/µl) to each tube. Incubated at 42°C for 60 minutes and at 65°C for 15 min. The DNA was precipitate with 3M Na-acetate and 96% ethanol. The pellet in each tube was washed in 80% ethanol. Each pellet was resuspended in RNase Mix (10 mM Tris-HCl (pH 7.5), 0.1 mM EDTA (pH 8.0), RNAase A 100 mµ/ml). The tubes were incubated at 37°C for 60 min and then added 30 µl sterile H2O to each tube. The DNA was precipitated using 3M Na-Acetate (pH 6.0) and ice-cold 96% ethanol. The pellets were washed in 80% ethanol and air-dried. Each pellet was resuspended in 15 µl hybridization buffer (5 x SSC, 0.1% SDS, 100 µg/ml, blocking RNA). The two fluorescents probes were mixed 1:1 in a PCR tube. This is the Sample-Mix.

Hybridisation and detection of hybridisation pattern

The Sample-mix was denatured 100°C for 3 min and then transferred directly to 55°C for 30 sec. The Sample-Mix was placed on ice and then added to the microarray slide (Pan Array; MWG Germany). The microarray slide was placed in a box and inside the petri dish with the pre-wetted 3MM paper. The lid was replaced back onto the petri dish. The petri dish was placed in a plastic bag.

The petri dish incubated in a dark incubator at 65°C 12-16 hours. Slides was washed in Washing Buffer I (2xSSC). The slides were submerged in pre-warmed Washing Buffer II (2 x SSC, 0.1% SDS). Pre-warmed at 65°C for 1 hour in a volume of least 10 ml/slide to cover the slides. Incubated on an orbital shaker at 65°C for 10 min. The slides were washed in Washing Buffer III (0.2 x SSC) in a volume of least 10 ml/slide to cover the slides. Incubated on an orbital shaker at room temperature for 3 min. The slides were washed in Washing Buffer IV (0.1 x SSC). Washed in Washing Buffer V (0.5 x SSC). The washing with washing buffer V was repeated for additional 3 times. All Washing Buffer V was removed by centrifuging at 800rpm for 3 min. The slide was scanned and evaluation performed using Affymetrix 418 Scanner, Affymetrix 418 Scanner Software and the software ImaGene 4.0 (BioDiscovery) according to the user manual.

All the differentially regulated genes were identified on the microarrays and the genes selected for a bioinformatics analysis identifying the gene fragments comprising a non-conserved region with high inter species and low intra species homology. These selected gene fragments was used for the Specific Differential Display Array.

All gene expression data obtained from analysis carried out on the microarrays are stored in the database together with all the information about the following points.

8) Genes identified (their accession number and sequence)
9) The microarray and protocol used (e.g. Affymetrix, Pan Array from MWG)

10) Information about the Biological material use for the analysis e.g. mouse brain.

11) The sequence of the Gene fragments identified to go onto the Specific Differential Display Array (according to example 4)

12) The gene expression profile performed on the Biological Material

13) Disease description

14) Information about the treatment of the biological source e.g. mouse treated with the 10mg/kg antidepressant fluoxetine one time/day for 21 days.

15) Drug information: Drug name, Dose usage, chemical structure, mechanism of action, Drug Interaction, Clinical Trial information, Indications and usage, Contradictions, Drug/Laboratory test interactions, receptor binding studies results, ADME results, PK results, toxicological profile and Results from animal models

EXAMPLE 2c

Identification of differentially regulated proteins using the MDLC-MSMS e.g. ICAT.

Protocol for Protein Extraction previous to ICAT labelling

The procedure is performed on ice, unless stated otherwise.

Lysis Buffer:

50 mM Tris-HCL, pH 7.4, 150 mM NaCl, 0.1 % SDS, Protease Inhibitor Cocktail

Protease Inhibitor Cocktail (100x concentration):

20 mM AESBF, 1 mM Leupeptin, 100 µM Pepstatin

Tissue (Example1) was taken directly from –80°C and kept on ice.

1) Homogenization
   a) Weigh tissue
   b) Add 3-5 volumes (v/w) of icecold lysis buffer
   c) Homogenize on ice
   d) Incubate on ice for 15 min, every 5 min mix gently by inverting
   e) Spin 15.000 g for 15 min, 4°C
   f) Filtered supernatant through 1 layer of Miracloth to remove fat particles
2) Acetone Precipitation
   a) Add 5 volumes of cold acetone
   b) Incubated at -20°C for at least 2 hours, every 30 min mix gently by inverting
   c) Spin 13.000 g for 10 min, 4°C
   d) Aspirated supernatant
   e) Washed twice with cold acetone

f) Removed residual acetone by air drying or lyophilization
g) Redissolved the pellet in a small volumen of denaturing buffer (from ICAT kit)

3) Protein Concentration
   a) Measured protein concentration by Bradford Assay
   b) Diluted an aliquot until at least 100 µg protein/80 µl denaturing buffer/vial
4) Proceeded according to ICAT Protocol from Applied Biosystems
   a) Section 7.1.2 (in "Cleavable ICAT Reagent Kit"-Protocol from Applied Biosystems) and on.

<u>Protocol for MD-LC MALDI-MS/MS of ICAT™-modified proteins</u>

Protein preparations from mouse brain tissue were labeled with acidcleavable ICAT-reagent and trypsinized (last step according to Section 7.1.4).

Separations were performed on an "UltiMate™ Nano LC System" from LC Packings, with Switchos™, FAMOS™ and Probot™ modules.

High-Resolution Cation-Exchange Column Separation:

Procedure according to Section 7.1.6. Also confere to Section 7.1.5.

PolySULFOETHYL Aspartamide, 300 Å, 5 µm, 200 x 4.6 mm, from PolyLC.

Eluent A:	10 mM K-PO_4, pH 3.0, 25 % ACN
Eluent B:	10 mM K-PO_4, pH 3.0, 25 % ACN, 350 mM KCL
Linear gradient:	0-100 % B in 1 hour at flowrate 0.2 ml/min, room temp.

30 fractions were collected.

Purifying Biotinylated Peptides and Cleaving of Biotin

Procedure according to Section 7.2 through to 7.2.6 inclusive.

Capillary HPLC Column Separation:

Procedure according to Section 8.1 and 8.4.

RP-column, C18, 300 Å, 5 µm, 75 i.d. µm x 150 mm,

Eluent A:	0.1 % TFA in 5 % Acetonitril, 95 % H_2O (Milli-Q®)
Eluent B:	0.1 % TFA in 95 % Acetonitril, 5 % H_2O (Milli-Q®)
Linear gradient:	2-30 % B in 2 hours at flowrate 200 nl/min.

0.2 µl fractions were collected on the Probot micro fraction collector.

MALDI-MS and MS/MS on a "AB 4700 Proteomics Analyzer":

Procedure according to Section 8.4.

In MS analyses standard reflector method with 50 laser shots/subspectra and 20 randomly selected search pattern positions was used.

In MS/MS analyses standard MS/MS acquisition methods was used.

MS spectra was first recorded for quantitation. Peak pairs representing ICAT reagent labeled peptides was recognized by mass differences (multiplums of 9.03 m/z). Precursor ions for MS/MS analysis were selected. MS/MS analyses were runned in automated mode.
The differentialy expressed Proteins were identified using GPS Explorer^{TM} software from MS/MS spectra of ICAT reagent-labelled peptides.

All the differentially regulated proteins were identified on the MSMS spectrums and the corresponding genes selected for a bioinformatics analysis identifying the gene fragments <u>comprising a non-conserved region </u>with high inter species and low intra species homology. These selected gene fragments was used for the Specific Differential Display Array.

All protein expression data obtained from analysis carried out using the MDLC-MSMS method ICAT are stored in the database together with all the information about the following points.

1) Proteins identified (their Protein and Gene accession numbers and sequences)
2) Protein differential Display method and protocols used (e.g. ICAT or 2D-gel electrophoresis)
3) Information about the Biological material use for the analysis e.g. mouse brain.
4) Disease description
5) The Protein expression profile performed on the Biological Material
6) The sequence of the Gene fragments identified to go onto the Specific Differential Display Array (according to example 4)
7) Information about the treatment of the biological source e.g. mouse treated with the 10mg/kg antidepressant fluoxetine one time/day for 21 days.
8) Drug information: Drug name, Dose usage, chemical structure, mechanism of action, Drug Interaction, Clinical Trial information, Indications and usage, Contradictions, Drug/Laboratory test interactions, receptor binding studies results, ADME results, PK results, toxicological profile and Results from animal models

EXAMPLE 3

<u>PCR fragment cloning, sequencing, and bioinformatic analysis of differentially regulated genes</u>

The PCR fragments from example 2a were cloned using a commercially available cloning kit according to the manufacturers instructions. The selected clones were then sequenced using the Beckman CEQ-2000 sequencer. Sequences were converted to FASTA format and any vector sequence was masked using the 'cross_match' program running on a Linux platform. Each sequence was used to search for matches in the GenBank database and the EST subset of Genbank using the BLAST sequence similarity search program. For each sequence the most significant hit was reported as a bit score, length of alignment region, percentage identity, and the GenBank gene name annotation.

A. Differentially expressed gene in antidepressant treated mice

One differential expressed PCR-fragment was identified from analysis on RNA from Example 1A using the method desribed in Example 2a. This fragment corresponded to mus musculus corticotropin releasing hormone receptor [CRF-R, acc. no. NM_007762].

B. Differentially expressed gene in db/db and ob/ob mice

One differential expressed gene fragment was identified from analysis on RNA from Example 1B using the method described in Example 2a. This fragment corresponded to the growth arrest-specific gene (gas6, acc. No. X59846)

EXAMPLE 4

Identification of a fragment comprising a non-conserved region

A. Identification of an anti-depression related gene fragment

The cDNA sequence of NM_007762 (CRF-R) was downloaded from Genbank.

The alignments were performed to identify non-conserved regions of CRF-R regions having less than 75% sequence homology to other sequences within the species. The size of the region is preferably from 25 to 750 bases. The numbering of the nucleotide sequence start at the 5' end of the polynucleotide and correspond to a [atg]. The first homology search was done to generate a fragment that is highly specific for CRF-R. Secondly, interspecies homology-search was performed to ensure the fragment can be used for identification of CRF-R in other species. Using clustal alignment a 270 region was identified having 96% homology (interspecies) to rat CRF-R [Acc. No. NM_030999.1] and 91% homology (interspecies) to human CRF-R [Acc. No. XM_038681].

B. Identification of a diabetes/obesity related gene fragment

The cDNA sequence of mouse Gas6 (X59846) was downloaded from Genbank.

The alignments were performed to identify non-conserved regions of Gas6 having less than 75 % sequence homology to other sequences within the species. The size of the region is preferably from 25 to 750 bases. The numbering of the nucleotide sequence start at the 5'end of the polynucleotide and correspond to a [atg]. The first homology search was done to generate a fragment that is highly specific for Gas6. Secondly, interspecies homology-search was performed to ensure the fragment can be used for identification of Gas6 in other species. Using clustal alignment a 240 region was identified having 94 % homology (interspecies) to rat Gas6 [Acc. No. D42148] and 87% homology (interspecies) to human Gas6 [Acc. No. M33031]. The result of a few of the analysed gene fragments are listed below

Coding region in Mouse Gas6	Percent identity to Rat Gas6	Percent identity to Human Gas 6
314-553	92	83
554-793	92	83
794-1033	**94**	**87**
1034-1273	90	80
1274-1513	92	83
1514-1753	90	82
1754-2022	89	82

C. Storage of RFDD-PCR results in the Database

All gene expression data obtained from analysis carried out using the RFDD-PCR technology are stored in the database together with all the information about the following points.

1) Genes identified (their accession numbers and sequences, Expression windows were they were identified, Size of the RFDD-PCR fragments)
2) Information about the Biological material use for the analysis e.g. mouse brain.
3) Scanned Images of the gels stored in TIF and JPEG formats.
4) Information about the treatment of the biological source e.g. mouse treated with the 10mg/kg antidepressant fluoxetine one time/day for 21 days.
5) Disease description
6) Drug information: Drug name, Dose usage, chemical structure, mechanism of action, Drug Interaction, Clinical Trial information, Indications and usage, Contradictions, Drug/Laboratory test interactions, receptor binding studies results, ADME results, PK results, toxicological profile and Results from animal models

EXAMPLE 5

Amplification of the determined fragments

A. Amplification of the anti-depression related gene fragment

The CRF-R region 841-1110 was amplified using conventional PCR well known for a person skilled in the art using following primers:

Region	Primer 1	Primer 2
841-1110	CAT CTA CCA GGG CCC CAT GA	GAA GAA GCC CTG AAA GGA CTC CAG

Brain total RNA was purchased from Research Genetics (D6030-01 NTR Brain) The cDNA was synthesized using the Omniscript RT Kit (Qiagen, 205111).

B. Amplification of the diabetes/obesity related gene fragment

The Gas6 region 794-1033 was amplified using conventional PCR well known for a person skilled in the art using following primers:

Region	Primer 1	Primer 2
794-1033	TAA AAC TAT CCC CAG ACA TGG	GGA CAA TCC AGG TGC TGT C

Mouse liver mRNA was purchased from Clontech laboratories (cat # 6616-1) The cDNA was synthesized using the Omniscript RT Kit (Qiagen, 205111).

The gene fragments of Examples 5A and 5B were amplified according to the following procedure:

1-2 µl cDNA obtained above
250 µM dNTP (27-2035-01, Amersham Pharmacia)
0.5 µM of each PCR primers (see table above)
1.5 mM $MgCl_2$ final concentration (Y02016, Gibco BRL)
1X PCR buffer without $MgCl_2$ (Y02028, Gibco BRL)
2.5 UTaq polymerase (10966-026, Gibco BRL)
H_2O to 20 µl final volume

The PCR generated fragments were separated using on a conventional agarose gel and cloned into a suitable vector according to supplier's instruction and the nucleotide sequence was analysed using CEQ 2000 DNA Analysis System (Beckman Coulter, U.S.A.).

EXAMPLE 6

Preparation of Master Glycerol Stocks

The glycerol stocks were prepared in 96 wells-trays (Corning Cat. No. cci3793) from the bacterial tranformations of the cloned PCR fragments described in example 3 on a Biomek. 50µl glycerol media was transferred into each well of a plane 96-well tray (Corning Cat. No. cci3793). 50 µl bacterial culture was transferred into each well of the plane 96-well tray and mixed with 2 x 100µl. A Storage Mat-I lid (Corning, Cat. No. 3094) was placed on each tray and the trays stored at –80 °C.

EXAMPLE 7a

Preparation of plasmids for the specific differential display cDNA arrays.

Ampicillin (100mg/ml) was added to Circlegrow medium (obtained from Bio101, Carlsbad, CA 92008, U.S.A.). Circlegrow medium/ampicillin was added to each well in a 4 x 2 ml 96 deep well tray (Corning Cat.No. cci 3961). The glycerol stocks from example 6 were added to each well. The tray was sealed with sealing sheet (Merck Eurolab A/S, Denmark), and incubated with shaking at 37°C for 16 hours prior to plasmid purification. The plasmid purification was performed using Biomek or a conventional method according to Maniatis et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbour Press) (1989) well known for a person skilled in the art. Following plasmide purification the specific fragments were generated by standard PCR using specific primer sets. The fragments were made “spot-ready” by precipitation according to standard procedure and dissolving in spot buffer in a suitable concentration.

EXAMPLE 7b

Preparation of oligonucleotides for the specific differential display oligoarrays.

As an alternative to generating cDNA fragments for all the identified differentially regulated genes identified as described in Example 2a, 2b and 2c, oligonucleotide based Specific Differential Display Arrrays were produced. From the identified differentially expressed genes the bioinformatic analysis as described in Example 4 was carried out with the object of identifying 50bp regions with high interspecies homology and low intraspecies homology. The identified 50bp fragments were produced as long oligonucleotides with a primary amine in the 5’-end using oligonucleotide synthesizer (DNA Technology).

The oligonucleotides were made "spot-ready" by diluting the oligonucleotides to 20 µM according to standard procedure in spotting buffer as in Example 7. After spotting onto 3D-Link Amine-Binding slides (Motorola, USA) the microarray slides were treated as in Example 8.

EXAMPLE 8

Preparation of 3D-Link Amine-Binding slide (array)

3D-Link Amine-Binding slide (array) may be obtained from (Motorola USA). To have an orientation on the arrays visible after scanning of the slides the each of the arrays corners are marked by a double labelled Cy3 and Cy5 primer with a 5' end amino group (5'-Cy3/5) in a final concentration of 1 pmol/µl to enable the possibility to place a grid on the scanned array. After spotting, seal with sealing tape and store at -20°C until use.

3MM paper was pre-wetted with saturated NaCl solution. All the slides were placed in a slide box without a lid and then placed in a plastic bag containing the NaCl saturated 3MM paper. The plastic bag was closed and incubated at room temperature for 36 hours. After incubation, the slides were removed from the plastic bag. The slides were then placed in pre-warmed blocking solution (0.1% SDS, SurModics Blocking Solution. Incubated for 20 minutes at 50 °C. Washed in redistilled H_2O. Incubated in 4 x SSC, 0.1% SDS solution (50°C), and washed at room temperate in redistilled H_2O. The slides were boiled in redistilled H_2O for 2 min. Washed in redistilled H_2O at room temperature. The slide was incubated in pre-hybridisation buffer prewarmed to 50°C (50 ml 20 x SSC, 10 ml 100 x Denhardt solution, 2 ml 10% SDS, 4 ml Salmon Sperm DNA (10 mg/ml), 134 ml Redistilled H_2O) at 50°C for 30 min. Washed in redistilled H_2O. The slides were stored at room temperature; dry and dark until use.

EXAMPLE 9

Using the prepared specific differential display arrays for determination of an expression profile in a biological material

Preparation of labelled samples

RNA from the biological material is obtained according to the procedures of Examples 1A or 1B.

The RNA was preciptated by centrifugation and the RNA pellet was in 70% ethanol. The RNA was precipitated at 15,000 x g for 15 min. The supernatant was discarded and the pellet air-dried. The RNA was adjusted to a concentration of 1µg/µl with DEPC-H2O. In 2 separate tubes 25 µl total RNA (1µg/µl) was placed and and 7 µl

DEPC treated H2O added. 4 μl of oligo-dT (e.g. T25V primer) (1 μg/μl) was added to each tube. The tubes were incubated in a Thermal cycler at 65°C for 3 min.

Tube 1 was prepared by adding, 5 μl 10 x cDNA Buffer (500mM Tris-HCl, pH 8.3; 800 mM KCl; 100 mM MgCl2; 40mM DTT), 2 μl Cy3-dUTP (1mM, Cat. No. PA53022, Amersham Pharmabiotech), 5 μl 10 x dNTP (5mM dATP; 5 mM dCTP; 5 mM dGTP; 5mM dTTP). The contents were mixed and added 2 μl reverse transcriptase (100 U/μl). Tube 2 was prepared by adding, 5 μl 10 x cDNA Buffer (500mM Tris-HCl, pH 8.3; 800 mM KCl; 100 mM MgCl2; 40mM DTT), 2 μl Cy5-dUTP (1mM, Cat. No. PA55022, Amersham Pharmabiotech), 5 μl 10 x dNTP (5mM dATP; 5 mM dCTP; 5 mM dGTP; 5mM dTTP). The contents were mixed and added 2 μl reverse transcriptase (100 U/μl). Incubate tube 1 and 2 for 42 °C for 60 min and at 65°C for 15 min. The temperature was decreases to 42°C. Then added reverse transcriptase (100 U/μl) to each tube. Incubated at 42°C for 60 minutes and at 65°C for 15 min. The DNA was precipitate with 3M Na-acetate and 96% ethanol. The pellet in each tube was washed in 80% ethanol. Each pellet were resuspended in RNase Mix (10 mM Tris-HCl (pH 7.5), 0.1 mM EDTA (pH 8.0), RNAase A 100 mμ/ml). The tubes were incubated at 37°C for 60 min and thenadded 30 μl sterile H2O to each tube. The DNA was precipitated using 3M Na-Acetate (pH 6.0) and ice-cold 96% ethanol. the pellets were washed in 80% ethanol and air-dried. Each pellet were resuspended in 15 μl hybridization buffer (5 x SSC, 0.1% SDS, 100 μg/ml, blocking RNA). The two fluorescents probes were mixed 1:1 in a PCR tube. This is the Sample-Mix.

Hybridisation and detection of hybridisation pattern

The Sample-mix was denaturedat 100°C for 3 min and thentransferred directly to 55°C for 30 sec. The Sample-Mix was placed on ice and then added to the microarray slide. The microarray slide was placed in a box and inside the petri dish with the pre-wetted 3MM paper. The lid was replaced back onto the petri dish. The petri dish was placed in a plastic bag.

The petri dish incubated in a dark incubator at 65°C 12-16 hours. Tlides was washed in Washing Buffer I (2xSSC). The slides were submerged in pre-warmed Washing Buffer II (2 x SSC, 0.1% SDS). Pre-warmed at 65°C for 1 hour in a volume of least 10 ml/slide to cover the slides. Incubated on an orbital shaker at 65°C for 10 min. The slides were washed in Washing Buffer III (0.2 x SSC) in a volume of least 10 ml/slide to cover the slides. Incubated on an orbital shaker at room temperature for 3 min. The slides weres washed in Washing Buffer IV (0.1 x SSC). Washed in Washing Buffer V (0.5 x SSC). The washing with washing buffer V was repeated for additional 3 times. All Washing Buffer V was removed by centrifuging at 800rpm for 3 min. Theslide was scanned and evaluation performed using Affymetrix 418 Scanner, Affymetrix 418

Scanner Software and the software ImaGene 4.0 (BioDiscovery) according to the user manual.

EXAMPLE 10

All gene expression data obtained from analysis carried out on the Specific Differential Display Arrays are stored in the database together with all the information about the following points.

1) Genes identified (their accession number and sequence) and spotted onto the Specific Differential Display Array
2) Method the genes are identified with e.g. RFDD or MDLC-MSMS
3) Information about the Biological material use for the analysis e.g. mouse brain.
4) The gene and/or protein expression profile performed on the Biological Material
5) Information about the treatment of the biological source e.g. mouse treated with the 10mg/kg antidepressant fluoxetine one time/day for 21 days.
6) Drug information: Drug name, Dose usage, chemical structure, mechanism of action, Drug Interaction, Clinical Trial information, Indications and usage, Contradictions, Drug/Laboratory test interactions, receptor binding studies results, ADME results, PK results, toxicological profile and Results from animal models

Usage of Database

A gene expression profile is performed on brain tissue from a treatment of a mouse with a new NCE on a Specific Differential Display Array. This gene expression profile will then be stored in the database and used for a complex datamining against all the gene profiles stored in the database resulting in the possibility to elucidate e.g. the potential therapeutic and toxicological effects of a new tested drug and suggesting new modification of this NCE to make it a more efficient NCE towards this therapeutic field.

CLAIMS

1. A method for identifying a gene fragment that is differentially expressed in a non-stressed and stressed state of a biological source, said method comprising:
- identifying a gene that is differentially expressed in a non-stressed and stressed state;
- identifying a fragment of the identified gene, said fragment comprising a non-conserved region.

2. The method according to claim 1, wherein the identification of a gene that is differentially expressed in a non-stressed and stressed state of a biological source comprises:
- providing a first nucleotide sample from a first biological source being in a non-stressed state;
- providing a second nucleotide sample from a second biological source being in a stressed state;
- identifying a gene that is differentially expressed in the first and second biological source.

3. The method according to any one of claims 1-2, wherein the step of identifying a gene that is differentially expressed in the first and second biological source comprises using a differential display technique

4. The method according to claim 3, wherein the differential display technique used is restriction fragment differential display.

5. The method according to claim 3, wherein the differential display technique used is a microarray.

6. The method according to claim 1, wherein the identification of a gene that is differentially expressed in a non-stressed and stressed state of a biological source comprises:
- identifying a protein that is differentially expressed in a non-stressed and stressed state;
- identifying the gene corresponding to the differentially expressed protein; and
- identifying a fragment of the identified gene, said fragment comprising a non-conserved region.

7. The method according to claim 6, wherein the identification of a protein that is differentially expressed in a non-stressed and stressed state comprises

- providing a first protein sample from a first biological source being in a non-stressed state;
- providing a second protein sample from a second biological source being in a stressed state;
- identifying a protein that is differentially expressed in the first and second biological source.

8. The method according to any of claims 6-7, wherein the step of identifying a protein that is differentially expressed in the first and second biological source comprises using Multi-Dimensional-Liquid-Chromatography-Mass-Spectrometry MDLC-MS).

9. The method according to any of claims 6-7, wherein the step of identifying a protein that is differentially expressed in the first and second biological source comprises using ICAT technology, Protein 2D-gel electrophoresis or Direct Analysis of Large Protein Complexes (DALPC).

10. The method according to any one of claims 1-9, wherein the step of identifying a fragment of the identified gene comprising a non-conserved region comprises comparing by alignment the identified gene with other intraspecies genes.

11. The method according to any of claims 1-10, wherein the stressed state of the biological source is directly or indirectly related to a disease state, a chemical treatment, a biological treatment, environmental influence or any other physiological or pathophysiological change, or a combination thereof.

12. The method according to claim 11, wherein the disease state is selected from the group consisting of CNS related diseases, depression, diabetes type II and obesity.

13. The method according to claim 11, wherein the chemical treatment is treatment with one or more chemicals selected from the group consisting of naturally occurring chemical entities and synthetically derived chemical entities.

14. The method according to any one of claims 1-13, wherein the biological source is selected from the group consisting of tissues, organs, biological fluids or parts or combinations thereof.

15. The method according to any one of claims 1-14, wherein the non-conserved region has less than 85% intraspecies identity.

16. The method according to any one of claims 1-15, wherein the non-conserved region has at least 65% interspecies identity.

17. The method according to any one of claims 1-16, wherein the identified fragment has a length of from 25 to 500 nucleotides.

18. A database comprising a plurality of gene fragments as identified by the method according to any of claims 1-17.

19. The database according to claim 18, further comprising additional expressed genes that are substantially relevant to the stressed state of the biological source.

20. A specific differential display array comprising a plurality of polynucleotide spots associated with a surface of a solid support, wherein the individual polynucleotide spot comprises a gene fragment as identified by the method according to any of claims 1-17.

21. A specific differential display array according to claim 20 comprising polynucleotide spots identified partly by the method of any of claims 1-5 and partly by the method of any of claims 6-7.

22. A specific differential display array according to claim 20 or 21, wherein the individual polynucleotide spot is in single stranded or double stranded form.

23. A specific differential display array according to claims 20-22, wherein the polynucleotide in the individual polynucleotide spot is selected from the group consisting of DNA; RNA; cDNA; polynucleotides of natural, synthetic or semisynthetic origin; LNA; INA or PNA.

24. A specific differential display array according to any one of claims 20-23, wherein the solid support is made of a flexible or rigid material.

25. A specific differential display array according to any one of claims 20-24, wherein the array comprises from 2 to 50,000 polynucleotide spots.

26. A method for preparing a specific differential display array according to any one of claims 20-25 comprising

- generating a plurality of compositions each comprising a gene fragment as identified by the method according to any one of claims 1-17;
- stably associating each composition in one or more individual spots on a surface of a solid support.

27. A method according to claim 26, wherein the plurality of compositions comprises compositions identified partly by the method of any of claims 1-5 and partly by the method of any of claims 6-7.

28. A method according to claim 26 or 27, wherein the individual gene fragment is produced by one or more primers specific for said gene fragment.

29. A method for the determination of an expression profile in a biological material, said method comprising:

- obtaining a polynucleotide sample from the biological material,
- labelling said sample to obtain a labelled target polynucleotide sample,
- contacting at least one labelled target polynucleotide sample with a specific differential display array according to any of claims 20-25 under conditions which are sufficient to produce a hybridisation pattern, and
- detecting said hybridisation pattern to obtain the expression profile of the biological material.

30. A method for the determination of a difference in expression profiles from at least a first and a second biological material, said method comprising:

- obtaining a first expression profile of the first biological material according to the method of claim 29,
- obtaining a second expression profile of the second biological material according to the method of claim 29,

- comparing the first and the second expression profiles to identify any difference in the expression profiles from the first and second biological biological material.

31. The method according to claim 30, wherein the first and the second biological material are of the same kind of biological material.

32. The method according to claims 30 or 31, wherein the first biological material is in a non-stressed state and the second biological material is in a stressed state.

33. The method according to claim 32, wherein the stressed state of the second biological material is directly or indirectly related to a disease state, a chemical treatment, a biological treatment, environmental influence or any other physiological or pathophysiological change, or a combination thereof.

34. The method according to claim 33, wherein the chemical treatment is treatment with one or more chemicals selected from the group consisting of naturally occurring chemical entities or synthetically derived chemical entities.

35. The method according to claim 34, wherein the disease state is selected from depression, CNS related diseases, diabetes type II, or obesity.

36. The method according to any one of claims 32-35, wherein the stressed state of the second biological material is directly or indirectly related to the stressed state of the biological source used in the method according to any one of claims 1-17.

37. A method for identifying a therapeutic, prophylactic and/or toxic agent involved in a direct or indirect action on the expression profile in a biological material, said method comprises:

- obtaining a first expression profile of a first biological material according to the method of claim 29,
- obtaining a second expression profile of a second biological material according to the method of claim 29,
- applying a test compound to the second biological material and obtaining a third expression profile thereof according to the method of claim 22,
- comparing the first, second and third expression profiles, and

- identifying any differences in the expression profiles so as to identify any biological response of the test compound on the expression profile.

38. The method according to claim 37 further comprising:
- applying a test compound to the first biological material and obtaining a fourth expression profile thereof according to the method of claim 29,
- comparing the first, second, third and fourth expression profiles, and
- identifying any differences in the expression profiles so as to identify any biological response of the test compound on the expression profile.

39. The method according to claims 37 or 38, wherein the first and second biological material are of the same kind of biological material.

40. The method according to any one of claims 37-39, wherein the first biological material is in a non-stressed state and the second biological material is in a stressed state.

41. The method according to any one of claims 37-40, wherein the expression profile of the second biological material is direct or indirect measure of a disease state, a chemical treatment, a biological treatment, environmental influence or any other physiological or pathophysiological change, or a combination thereof.

42. The method according to any of claims 37-41, wherein the test compound is a chemical or biological derived compound such as compounds selected from the group consisting of therapeutic, prophylactic and/or toxic chemical entities, physiologically chemical entities, hormones, vitamins, nutrients, pesticides, fungicides, bacteriocides and any other organic chemical entity.

43. A diagnostic method for the determination of the differences of expression profiles between two biological materials, said method comprises:
- obtaining a first expression profile of a first biological material according to the method of claim 29,
- obtaining a second expression profile of a second biological material according to the method of claim 29,
- comparing the first and second expression profile, and identifying any difference in the expression profiles.

44. The diagnostic method according to claim 45, wherein the expression profile from more than two different biological materials are compared, such as biological materials, which are in different stages of a disease.

45. A kit for use in a hybridisation assay, said kit comprising a specific differential display array according to any one of claims 20-25.

46. The kit according to claim 45, wherein said kit further comprises reagents for generating a labelled target polynucleotide sample.

47. The kit according to claims 45 or 46, wherein said kit further comprises a hybridisation buffer.