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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 January 2018 (18.01.2018)

(10) International Publication Number
WO 2018/011107 A1

(51) International Patent Classification:
G01N 33/00 (2006.01) *G01N 33/574* (2006.01)

(21) International Application Number: PCT/EP2017/067196

(22) International Filing Date: 10 July 2017 (10.07.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16305873.8 11 July 2016 (11.07.2016) EP

(71) Applicants: INSERM (INSTITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MÉDICALE) [FR/FR]; 101, rue de Tolbiac, 75013 Paris (FR). **UNIVERSITÉ PAUL SABATIER TOULOUSE III** [FR/FR]; 118 route de Narbonne, 31400 Toulouse (FR). **CENTRE HOSPITALIER UNIVERSITAIRE DE TOULOUSE** [FR/FR]; 2, rue Viguerie Hôtel-Dieu Pont Neuf, 31300 TOULOUSE (FR).

(72) Inventors: LENFANT, Françoise; UMR1048 INSERM / UPS, CHU Rangueil, BP 84225, 31224 TOULOUSE Cedex (FR). **FONTAINE, Coralie**; UMR1048 INSERM / UPS, CHU Rangueil, BP 84225, 31224 TOULOUSE Cedex (FR). **ARNAL, Jean-François**; UMR1048 INSERM / UPS, CHU Rangueil, BP 84225, 31224 TOULOUSE Cedex (FR). **CHANTALAT, Elodie**; UMR1048 INSERM / UPS, CHU Rangueil, BP 84225, 31224 TOULOUSE Cedex (FR). **BOUDOU, Frédéric**; UMR1048 INSERM / UPS, CHU Rangueil, BP 84225, 31224 TOULOUSE Cedex (FR).

(74) Agent: COLLIN, Matthieu; Inserm Transfert, 7 rue Watt, 75013 PARIS (FR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: USE OF ER-ALPHA 46 IN METHODS AND KITS FOR ASSESSING THE STATUS OF BREAST CANCER

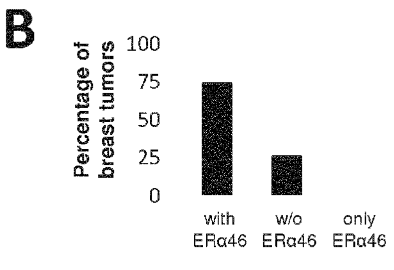


Figure 2

(57) Abstract: The present invention relates to methods and kits for the assessing status of breast cancer. Using approaches combining immunohistochemistry, western blotting and proteomics, the inventors first identified appropriate antibodies to detect ERα46 in breast cancer cell lines. They then analyzed the expression levels of ERα46 in a series of 116 ERα-positive human breast tumors. The inventors studied the mode of ERα46 generation and performed a systematic analysis of coregulator binding for both ERα isoforms. In particular, the present invention relates to a method for assessing the status of breast cancer in a subject comprising quantifying in a tumor tissue sample obtained from the subject the expression level of an ERα46 isoform comprising an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the position 21 to position 442 in SEQ ID NO:1.

WO 2018/011107 A1

USE OF ER-ALPHA 46 IN METHODS AND KITS FOR ASSESSING THE STATUS OF BREAST CANCER

FIELD OF THE INVENTION:

The present invention relates to methods and kits for the assessing status of breast cancer.

BACKGROUND OF THE INVENTION:

Breast cancer is a major public health concern since its incidence and mortality continues to rise. It is the second most common cancer overall and by far the most frequent cancer among women (1). The etiology of breast cancer is multifactorial and although the mechanisms of carcinogenesis remain poorly defined, the role of hormones is recognized as a major risk factor in breast cancer development, in particular 17β estradiol (E2) and its derivatives.

Estrogen receptor α (ERα) is one of two ERs and is involved in several key aspects of breast cancer diagnosis (2). Firstly, ERα protein immunoreactivity in the nucleus of mammary epithelial cells is systematically evaluated and quantified during anatomopathological diagnosis with 70% of breast cancers initially described as ERα-positive(2). Secondly, ERα expression in breast cancers correlates with improved survival rates and reduced risk of recurrence (3-5). Finally, the blockade of ERα activity represents a major targeted therapy for ERα-positive breast cancer, with tamoxifen and aromatase inhibitors having already benefitted millions of women (6). Despite the success of these treatments, 30 to 40% of patients develop resistance (7). This highlights the need for further in-depth characterization of ERα-positive tumors and a full understanding of the mechanisms underlying the disease in order to propose new therapeutic approaches.

In addition to the “classic” full length 66kDa ERα (ERα66) which harbors two important activation functions, AF-1 and AF-2, another isoform of 46kDa (ERα46) has been characterized which lacks the first 173 N-terminal amino acids. As ERα46 is devoid of the A/B domain, it lacks AF-1. It has been reported to be expressed in various cell types such as human osteoblasts(8), macrophages(9) and vascular endothelial cells(10), but also in cancer cells such as colorectal tumor tissues (11) and tamoxifen-resistant breast cancer cell lines(12). Mechanisms regulating both the expression of the ERα46 protein and its functions remain essentially unknown. It can be generated by either alternative splicing(14), proteolysis (13) or translation through an internal ribosome entry site (IRES) (14, 15). This IRES-mediated alternative initiation of translation thus generates two different proteins from a single RNA. A few studies have suggested ERα46 plays an inhibitory role in the growth of cancer cell lines, suggesting that ERα46 could affect tumor progression. The overexpression of ERα46 in proliferating MCF7 cells provoked cell cycle arrest in G0/G1 phase and inhibited ERα66-mediated estrogenic induction of the AF-1-sensitive reporters c-fos and cyclin D1, as well as estrogen-responsive element (ERE)-driven reporters(14, 16). It was also shown that ERα46inhibits growth and induces apoptosis in human HT-29 colon adenocarcinoma cells (17). This inhibition likely occurs through competition between ERα66 and ERα46 homodimers for binding to the ERE, leading to the specific recruitment of coregulators(16). The role of this truncated AF-1-deficient isoform has also been questioned *in vivo* using mice deficient in the ERαA/B domain (named *ERαAF-1*^{0}), which express only a short 49kDa isoform that isfunctionally similar to ERα46.These *ERαAF-1*^{0}mice revealed a complete infertility phenotype (18) that was associated with an altered proliferative effect of E2 on the uterine epithelium and a loss of its transcriptional response in this tissue (19).

Thus, the roles and functions of this ERα46 isoform appear to be different from those of full-length ERα66. The expression levels of this truncated isoform in human breast tumors remains unknown, even though the expression of a 47kDa isoform of ERα in human breast cancers was reported more than two decades ago (20).The diagnosis of ERα-positive breast tumors is now based on immunohistochemistry using antibodies that do not distinguish between the two isoforms due to their complete similar amino acid sequence. Currently, several antibodies are used for immunohistochemical detection of ERα66 in human breast tumors but many of them are not thoroughly characterized in terms of ERα46 recognition.

SUMMARY OF THE INVENTION:

The present invention relates to methods and kits for assessing the status of breast cancer. In particular, the present invention is defined by the claims.

DETAILED DESCRIPTION OF THE INVENTION:

To date, all studies conducted on breast cancer diagnosis have focused on the expression of Estrogen Receptor alpha (ERα). However, in contrast to the full-length 66kDa protein (ERα66), much less is known about a shorter 46kDa isoform (ERα46)that is expressed and can antagonize estrogen-induced proliferative breast cancer cell lines. Here, the inventors investigated the expression levels of ERα46 in breast tumors in relation to tumor grade and size and examined the mechanism of its generation and its specificities of coregulatory binding. Using approaches combining immunohistochemistry, western blotting and proteomics, the inventors first identified appropriate antibodies to detect ERα46 in breast cancer cell lines. They then analyzed the expression levels of ERα46 in a series of 116 ERα-positive human breast tumors. Finally, the inventors studied the mode of ERα46 generation and performed a systematic analysis of coregulator binding for both ERα isoforms. ERα46was expressed in over 70% of breast tumors at variable levels and was sometimes prominent over ERα66, especially in lower-grade and smaller-sized tumors. The inventors also found that ERα46 can be generated via internal ribosome entry site-mediated translation in the context of cell stress. Coregulator binding analysis of the ERα46 isoform was also found to be specific when compared to that of ERα66. Beside the full length ERα66 in breast tumors, an ERα46 isoform is also expressed in a majority of breast tumors. ERα46 should now represent an additional diagnosis factor to be considered in the pathophysiology, as it could influence both prognosis and treatment of breast cancer. In particular, this work indicates that the diagnosis of ERα-positive breast tumors must be reconsidered using appropriate antibodies which recognize ERα46. ERα46 expression correlates with lower-grade, smaller sized-tumors and it exhibits a different pattern of association with coregulators that should modulate response to common breast cancer treatments in the presence of estrogen compared to the ERα66 isoform. The frequent expression and functional characteristics of ERα46 indicate that ERα46 should now be taken into account in the diagnosis of breast cancer, with potential important implications for its prognosis and its treatment.

Accordingly, the present invention relates to a method for assessing the status of breast cancer in a subject comprising i) quantifying in a tumor tissue sample obtained from the subject the expression level of an ERα46isoform comprising an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the position 21 to position 442 in SEQ ID NO:1, ii) comparing the quantification value determined at step i) with a predetermined reference value and iii) detecting differences in the quantification value determined at step i) and the predetermined reference value is indicative of the status of breast cancer.

In some embodiments, the subject suffers from an estrogen-receptor positive breast cancer. A breast cancer patient having estrogen receptors present on many of the cancer cells is considered as a "estrogen-receptor-positive" (ER+) patient. ER-positive cancers rely on a source of estrogen to encourage proliferation (increase the number) of cancer cells. According to the invention, the cancer can be a low-grade cancer. Tumor "grade" is a system used to classify cancer cells in terms of how abnormal they appear and how quickly the tumor is likely to grow and spread. Many factors are considered when determining tumor grade, including the structure and growth pattern of the cells. More specifically, cancer cells are "low grade" if they appear similar to normal cells, and "high grade" if they appear poorly differentiated. For example, a G1 cancer would be classified as a low-grade cancer, whereas a G4 cancer would be classified as a high-grade cancer. Additionally or alternatively, the cancer can be an early-stage cancer. Cancer "stage" refers to the extent or severity of the cancer, based on factors such as the location of the primary tumor, tumor size, number of tumors, and lymph node involvement (spread of cancer into lymph nodes). For instance, an "early-stage" cancer is a single tumor, of small size, with a low degree of spread to regional lymph nodes. The stage of a cancer may for instance be determined using the TNM classification. For example, a T1/N0/M0 cancer would be classified as an early-stage cancer, whereas a T4/N3/M1 cancer would be classified as a late-stage cancer.

The method of the present invention is particular suitable at the diagnosis stage for assessing the status of the cancer. In particular, the inventors demonstrate that the expression of the ERα46 isoform correlates with lower-grade, smaller sized-tumors. Accordingly, the method of the present invention could contribute 1) to refine the prognosis and survival time of the patient, including overall survival (OS), progression-free survival (PFS) and/or the disease-free survival (DFS) of the cancer patient, 2) to adapt and further personalize the hormonal treatment, as the efficacy of tamoxifen or aromatase inhibitors, the two main hormonal treatment given to women diagnosed for an ER-positive breast cancer. Thus, the respective expression levels of the ERα66 and ERα46 isoforms could influence the therapeutic management of the patients.

As used herein, the term “tumor tissue sample” has its general meaning in the art and encompasses pieces or slices of tissue that have been removed including following a surgical tumor resection. The tumor tissue sample can be subjected to a variety of well-known post-collection preparative and storage techniques (e.g., fixation, storage, freezing, etc.) prior to determining the cell densities. Typically the tumor tissue sample is fixed in formalin and embedded in a rigid fixative, such as paraffin (wax) or epoxy, which is placed in a mould and later hardened to produce a block which is readily cut. Thin slices of material can be then prepared using a microtome, placed on a glass slide and submitted e.g. to immunohistochemistry (using an IHC automate such as BenchMark® XT, for obtaining stained slides). The tumour tissue sample can be used in microarrays, called as tissue microarrays

(TMAs). TMA consists of paraffin blocks in which up to 1000 separate tissue cores are assembled in array fashion to allow multiplex histological analysis. This technology allows rapid visualization of molecular targets in tissue specimens at a time, either at the DNA, RNA or protein level. TMA technology is described in WO2004000992, US8068988, Olli et al 2001 Human Molecular Genetics, Tzankov et al 2005, Elsevier; Kononen et al 1198; Nature Medicine.

As used herein, the term "ERα46 isoform" refers to any isoform of about 46kDa (ERα46) of estrogen receptor which lacks the first 173 N-terminal amino acids in comparison with the full length of "classic" full length 66kDa ERα (ERα66) which comprises the amino acid sequence from the position 21 to position 442 in SEQ ID NO:1. It can also referred to a related the ER protein which lacks the complete A domain and a part of B domain, whichcan be obtained by proteolysis of ERa66. As ERα46 is devoid of the A/B domain or part of B domain, it lacks AF-1.

SEQID NO:1:
SDNRRQGGRERLASTNDKGS**MAMESAKETRYCAVCNDYASGYHYGVWSC**
EGCKAFFKRSIQGHNDYMCPATNQCTIDKNRRKSCQACRLRKCYEVGM
MKGGIRKDRRGGRMLKHKRQRDDGEGRGEVGSAGDMRAANLWPSPLM
IKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTN
LADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSM
EHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF
VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTL
QQQHQRLAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLD
AHRLHAPTSRGGASVEETDQSHLATAGSTSSHSLQKYYITGEAEGFPATV

According to the present invention, the ERα46 protein isoformof the present is about 422 amino acids.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 20 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 19 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 18 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 17 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 16 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 15 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 14 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 13 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 12 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 11 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 10 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 9 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 8 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 7 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 6 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 5 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 4 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 3 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 2 to the amino acid at position 442 in SEQ ID NO:1.

In some embodiments, the ERα46 isoform comprises an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 1 to the amino acid at position 442 in SEQ ID NO:1.

According to the invention a first amino acid sequence having at least 90% of identity with a second amino acid sequence means that the first sequence has 90; 91; 92; 93; 94; 95; 96; 97; 98; 99 or 100% of identity with the second amino acid sequence. Sequence identity is frequently measured in terms of percentage identity (or similarity or homology); the higher the percentage, the more similar are the two sequences. Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith and Waterman, Adv. Appl. Math., 2:482, 1981; Needleman and Wunsch, J. Mol. Biol., 48:443, 1970; Pearson and Lipman, Proc. Natl. Acad. Sci. U.S.A., 85:2444, 1988; Higgins and Sharp, Gene, 73:237-244, 1988; Higgins and Sharp, CABIOS, 5:151-153, 1989; Corpet et al. Nuc. Acids Res., 16:10881-10890, 1988; Huang et al., Comp.ApplsBiosci., 8:155-165, 1992; and Pearson et al., Meth. Mol. Biol., 24:307-31, 1994).Altschul et al., Nat. Genet., 6:119-129, 1994, presents a detailed consideration of sequence alignment methods and homology calculations. By way of example, the alignment tools ALIGN (Myers and Miller, CABIOS 4:11-17, 1989) or LFASTA (Pearson and Lipman, 1988) may be used to perform sequence comparisons (Internet Program® 1996, W. R. Pearson and the University of Virginia, fasta20u63 version 2.0u63, release date December 1996). ALIGN compares entire sequences against one another, while LFASTA compares regions of local similarity. These alignment tools and their respective tutorials are available on the Internet at the NCSA Website, for instance. Alternatively, for comparisons of amino acid sequences of greater than about 30 amino acids, the Blast 2 sequences function can be employed using the default BLOSUM62 matrix set to default parameters, (gap existence cost of 11, and a per residue gap cost of 1). When aligning short peptides (fewer than around 30 amino acids), the alignment should be performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties). The BLAST sequence comparison system is available, for instance, from the NCBI web site; see also Altschul et al., J. Mol. Biol., 215:403-410, 1990; Gish. & States, Nature Genet., 3:266-272, 1993; Madden et al. Meth. Enzymol., 266:131-141, 1996; Altschul et al., Nucleic Acids Res., 25:3389-3402, 1997; and Zhang & Madden, Genome Res., 7:649-656, 1997.

Typically, the expression of the ERα46 protein isoform in the tumor tissue sample is determined by any well-known method in the art, for example, an immunoassay. Convenient immunoassays include, for example, enzyme linked immunosorbent assays (ELISAs), immunohistochemical techniques, and Western blots. In some embodiments, the expression of the ERα46 isoform in the tumor tissue sample is determined by immunohistochemistry. For example, the determination is performed by contacting the tumor tissue sample with a binding partner (e.g. an antibody) specific ERα46.The binding partner (e.g. antibody) has specificity for ERα46 and not for ERα66.As used herein, the term "specificity" refers to the ability of an antibody to detectably bind an epitope presented on an antigen, such as ERα46, while having relatively little detectable reactivity with non-ERα46 proteins or structures such as ERα66. Specificity can be relatively determined by binding or competitive binding assays, using, e.g., Biacore instruments, as described elsewhere herein. Specificity can be exhibited by, e.g., an about 10:1, about 20:1, about 50:1, about 100:1, 10.000:1 or greater ratio of affinity/avidity in binding to the specific antigen versus nonspecific binding to other irrelevant molecules. The term "affinity", as used herein, means the strength of the binding of an antibody to an epitope. The affinity of an antibody is given by the dissociation constant Kd, defined as [Ab] x [Ag] / [Ab-Ag], where [Ab-Ag] is the molar concentration of the antibody-antigen complex, [Ab] is the molar concentration of the unbound antibody and [Ag] is the molar concentration of the unbound antigen. The affinity constant Ka is defined by 1/Kd. Preferred methods for determining the affinity of mAbs can be found in Harlow, et al., Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1988), Coligan et al., eds., Current Protocols in Immunology, Greene Publishing Assoc. and Wiley Interscience, N.Y., (1992, 1993), and Muller, Meth. Enzymol. 92:589-601 (1983), which references are entirely incorporated herein by reference. One preferred and standard method well known in the art for determining the affinity of mAbs is the use of Biacore instruments.

Immunohistochemistry typically includes the following steps i) fixing the tumor tissue sample with formalin, ii) embedding said tumor tissue sample in paraffin, iii) cutting said tumor tissue sample into sections for staining, iv) incubating said sections with the binding partner specific for the immune checkpoint protein of interest, v) rinsing said sections, vi) incubating said section with a secondary antibody typically biotinylated and vii) revealing the antigen-antibody complex typically with avidin-biotin-peroxidase complex. Accordingly, the tumortissue sample is firstly incubated with the binding partners having for the immune checkpoint protein of interest. After washing, the labeled antibodies that are bound to the immune checkpoint protein of interest are revealed by the appropriate technique, depending of the kind of label is borne by the labeled antibody, e.g. radioactive, fluorescent or enzyme label. Multiple labelling can be performed simultaneously. Alternatively, the method of the present invention may use a secondary antibody coupled to an amplification system (to intensify staining signal) and enzymatic molecules. Such coupled secondary antibodies are commercially available, e.g. from Dako, EnVision system. Counterstaining may be used, e.g. Hematoxylin& Eosin, DAPI, Hoechst. Other staining methods may be accomplished using any suitable method or system as would be apparent to one of skill in the art, including automated, semi-automated or manual systems.

For example, one or more labels can be attached to the antibody, thereby permitting detection of the target protein (i.e the immune checkpoint protein). Exemplary labels include radioactive isotopes, fluorophores, ligands, chemiluminescent agents, enzymes, and combinations thereof. Non-limiting examples of labels that can be conjugated to primary and/or secondary affinity ligands include fluorescent dyes or metals (e.g. fluorescein, rhodamine, phycoerythrin, fluorescamine), chromophoric dyes (e.g. rhodopsin), chemiluminescent compounds (e.g. luminal, imidazole) and bioluminescent proteins (e.g. luciferin, luciferase), haptens (e.g. biotin). A variety of other useful fluorescers and chromophores are described in Stryer L (1968) Science 162:526-533 and Brand L and Gohlke J R (1972) Annu. Rev. Biochem. 41:843-868. Affinity ligands can also be labeled with enzymes (e.g. horseradish peroxidase, alkaline phosphatase, beta-lactamase), radioisotopes (e.g. ^{3}H, ^{14}C, ^{32}P, ^{35}S or ^{125}I) and particles (e.g. gold). The different types of labels can be conjugated to an affinity ligand using various chemistries, e.g. the amine reaction or the thiol reaction. However, other reactive groups than amines and thiols can be used, e.g. aldehydes, carboxylic acids and glutamine. Various enzymatic staining methods are known in the art for detecting a protein of interest. For example, enzymatic interactions can be visualized using different enzymes such as peroxidase, alkaline phosphatase, or different chromogens such as DAB, AEC or Fast Red. In some embodiments, the label is a quantum dot.For example, Quantum dots (Qdots) are becoming increasingly useful in a growing list of applications including immunohistochemistry, flow cytometry, and plate-based assays, and may therefore be used in conjunction with this invention. Qdotnanocrystals have unique optical properties including an extremely bright signal for sensitivity and quantitation; high photostability for imaging and analysis. A single excitation source is needed, and a growing range of conjugates makes them useful in a wide range of cell-based applications. QdotBioconjugates are characterized by quantum yields comparable to the brightest traditional dyes available. Additionally, these quantum dot-based fluorophores absorb 10-1000 times more light than traditional dyes. The emission from the underlying Qdot quantum dots is narrow and symmetric which means overlap with other colors is minimized, resulting in minimal bleed through into adjacent detection channels and attenuated crosstalk, in spite of the fact that many more colors can be used simultaneously.In other examples, the antibody can be conjugated to peptides or proteins that can be detected via a labeled binding partner or antibody. In an indirect IHC assay, a secondary antibody or second binding partner is necessary to detect the binding of the first binding partner, as it is not labeled.

In some embodiments, the resulting stained specimens are each imaged using a system for viewing the detectable signal and acquiring an image, such as a digital image of the staining. Methods for image acquisition are well known to one of skill in the art. For example, once the sample has been stained, any optical or non-optical imaging device can be used to detect the stain or biomarker label, such as, for example, upright or inverted optical microscopes, scanning confocal microscopes, cameras, scanning or tunneling electron microscopes, canning probe microscopes and imaging infrared detectors. In some examples, the image can be captured digitally. The obtained images can then be used for quantitatively or semi-quantitatively determining the amount of the immune checkpoint protein in the sample, or the absolute number of cells positive for the maker of interest, or the surface of cells positive for the maker of interest. Various automated sample processing, scanning and analysis systems suitable for use with IHC are available in the art. Such systems can include automated staining and microscopic scanning, computerized image analysis, serial section comparison (to control for variation in the orientation and size of a sample), digital report generation, and archiving and tracking of samples (such as slides on which tissue sections are placed). Cellular imaging systems are commercially available that combine conventional light microscopes with digital image processing systems to perform quantitative analysis on cells and tissues, including immunostained samples. See, e.g., the CAS-200 system (Becton, Dickinson & Co.). In particular, detection can be made manually or by image processing techniques involving computer processors and software. Using such software, for example, the images can be configured, calibrated, standardized and/or validated based on factors including, for example, stain quality or stain intensity, using procedures known to one of skill in the art (see e.g., published U.S. Patent Publication No. US20100136549). The image can be quantitatively or semi-quantitatively analyzed and scored based on staining intensity of the sample. Quantitative or semi-quantitative histochemistry refers to method of scanning and scoring samples that have undergone histochemistry, to identify and quantify the presence of the specified biomarker (i.e. immune checkpoint protein). Quantitative or semi-quantitative methods can employ imaging software to detect staining densities or amount of staining or methods of detecting staining by the human eye, where a trained operator ranks results numerically. For example, images can be quantitatively analyzed using a pixel count algorithms and tissue recognition pattern (e.g. Aperio Spectrum Software, Automated QUantitatative Analysis platform (AQUA® platform), or Tribvn with Ilastic and Calopix software), and other standard methods that measure or quantitate or semi-quantitate the degree of staining; see e.g., U.S. Pat. No. 8,023,714; U.S. Pat.No. 7,257,268; U.S. Pat.No. 7,219,016; U.S. Pat. No. 7,646,905; published U.S. Patent Publication No. US20100136549 and 20110111435; Camp et al. (2002) Nature Medicine, 8:1323-1327; Bacus et al. (1997) Analyt Quant CytolHistol, 19:316-328). A ratio of strong positive stain (such as brown stain) to the sum of total stained area can be calculated and scored. The amount of the detected biomarker (i.e. the immune checkpoint protein) is quantified and given as a percentage of positive pixels and/or a score. For example, the amount can be quantified as a percentage of positive pixels. In some examples, the amount is quantified as the percentage of area stained, e.g., the percentage of positive pixels. For example, a sample can have at least or about at least or about 0, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more positive pixels as compared to the total staining area. For example, the amount can be quantified as an absolute number of cells positive for the maker of interest. In some embodiments, a score is given to the sample that is a numerical representation of the intensity or amount of the histochemical staining of the sample, and represents the amount of target biomarker (e.g., the immune checkpoint protein) present in the sample. Optical density or percentage area values can be given a scaled score, for example on an integer scale.

In some embodiments, the expression level of the ERα46 isoform is determined at nucleic acid level. Typically, the expression level of the ERα46 isoform may be analysed by determining the quantity of the mRNA encoding for the ERα46 isoform. Methods for determining the quantity of mRNA are well known in the art. For example the nucleic acid contained in the samples (e.g., cell or tissue prepared from the subject) is first extracted according to standard methods, for example using lytic enzymes or chemical solutions or extracted by nucleic-acid-binding resins following the manufacturer's instructions. The extracted mRNA is then detected by hybridization (e. g., Northern blot analysis, in situ hybridization) and/or amplification (e.g., RT-PCR). Other methods of Amplification include ligase chain reaction (LCR), transcription-mediated amplification (TMA), strand displacement amplification (SDA) and nucleic acid sequence based amplification (NASBA).

Nucleic acids having at least 10 nucleotides and exhibiting sequence complementarity or homology to the mRNA of interest herein find utility as hybridization probes or amplification primers. It is understood that such nucleic acids need not be identical, but are typically at least about 80% identical to the homologous region of comparable size, more preferably 85% identical and even more preferably 90-95% identical. In certain embodiments, it will be advantageous to use nucleic acids in combination with appropriate means, such as a detectable label, for detecting hybridization.

Typically, the nucleic acid probes include one or more labels, for example to permit detection of a target nucleic acid molecule using the disclosed probes. In various applications, such as in situ hybridization procedures, a nucleic acid probe includes a label (e.g., a detectable label). A “detectable label” is a molecule or material that can be used to produce a detectable signal that indicates the presence or concentration of the probe (particularly the bound or hybridized probe) in a sample. Thus, a labeled nucleic acid molecule provides an indicator of the presence or concentration of a target nucleic acid sequence (e.g., genomic target nucleic acid sequence) (to which the labeled uniquely specific nucleic acid molecule is bound or hybridized) in a sample. A label associated with one or more nucleic acid molecules (such as a probe generated by the disclosed methods) can be detected either directly or indirectly. A label can be detected by any known or yet to be discovered mechanism including absorption, emission and/ or scattering of a photon (including radio frequency, microwave frequency, infrared frequency, visible frequency and ultra-violet frequency photons). Detectable labels include colored, fluorescent, phosphorescent and luminescent molecules and materials, catalysts (such as enzymes) that convert one substance into another substance to provide a detectable difference (such as by converting a colorless substance into a colored substance or vice versa, or by producing a precipitate or increasing sample turbidity), haptens that can be detected by antibody binding interactions, and paramagnetic and magnetic molecules or materials.

Particular examples of detectable labels include fluorescent molecules (or fluorochromes). Numerous fluorochromes are known to those of skill in the art, and can be selected, for example from Life Technologies (formerly Invitrogen), e.g., see, The Handbook—A Guide to Fluorescent Probes and Labeling Technologies). Examples of particular fluorophores that can be attached (for example, chemically conjugated) to a nucleic acid molecule (such as a uniquely specific binding region) are provided in U.S. Pat. No. 5,866, 366 to Nazarenko et al., such as 4-acetamido-4'-isothiocyanatostilbene-2,2'disulfonic acid, acridine and derivatives such as acridine and acridineisothiocyanate, 5-(2'-aminoethyl) aminonaphthalene-1-sulfonic acid (EDANS), 4-amino -N- [3 vinylsulfonyl)phenyl]naphthalimide-3,5 disulfonate (Lucifer Yellow VS), N-(4-anilino-1-naphthyl)maleimide, antl1ranilamide, Brilliant Yellow, coumarin and derivatives such as coumarin, 7-amino-4-methylcoumarin (AMC, Coumarin 120), 7-amino-4-trifluoromethylcouluarin (Coumarin 151); cyanosine; 4',6-diarninidino-2-phenylindole (DAPI); 5',5"dibromopyrogallol-sulfonephthalein (Bromopyrogallol Red); 7 -diethylamino -3 - (4'-isothiocyanatophenyl)-4-methylcoumarin; diethylenetriaminepentaacetate; 4,4'-diisothiocyanatodihydro-stilbene-2,2'-disulfonic acid; 4,4'-diisothiocyanatostilbene-2,2'-disulfor1ic acid; 5-[dimethylamino] naphthalene-1-sulfonyl chloride (DNS, dansyl chloride); 4-(4'-dimethylaminophenylazo)benzoic acid (DABCYL); 4-dimethylaminophenylazophenyl-4'-isothiocyanate (DABITC); eosin and derivatives such as eosin and eosin isothiocyanate; erythrosin and derivatives such as erythrosin B and erythrosinisothiocyanate; ethidium; fluorescein and derivatives such as 5-carboxyfluorescein (FAM), 5-(4,6dicl1lorotriazin-2-yDarninofluorescein (DTAF), 2'7'dimethoxy-4'5'-dichloro-6-carboxyfluorescein (JOE), fluorescein, fluorescein isothiocyanate (FITC), and QFITC Q(RITC); 2',7'-difluorofluorescein (OREGON GREEN®); fluorescamine; IR144; IR1446; Malachite Green isothiocyanate; 4-methylumbelliferone; orthocresolphthalein; nitrotyrosine; pararosaniline; Phenol Red; B-phycoerythrin; o-phthaldialdehyde; pyrene and derivatives such as pyrene, pyrene butyrate and succinimidyl 1-pyrene butyrate; Reactive Red 4 (Cibacron Brilliant Red 3B-A); rhodamine and derivatives such as 6-carboxy-X-rhodamine (ROX), 6-carboxyrhodamine (R6G), lissaminerhodamine B sulfonyl chloride, rhodamine (Rhod), rhodamine B, rhodamine 123, rhodamine X isothiocyanate, rhodamine green, sulforhodamine B, sulforhodamine 101 and sulfonyl chloride derivative of sulforhodamine 101 (Texas Red); N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA); tetramethylrhodamine; tetramethylrhodamineisothiocyanate (TRITC); riboflavin; rosolic acid and terbium chelate derivatives. Other suitable fluorophores include thiol-reactive europium chelates which emit at approximately 617 mn (Heyduk and Heyduk, Analyt. Biochem. 248:216-27, 1997; J. Biol. Chem. 274:3315-22, 1999), as well as GFP, LissamineTM, diethylaminocoumarin, fluorescein chlorotriazinyl, naphthofluorescein, 4,7-dichlororhodamine and xanthene (as described in U.S. Pat. No. 5,800,996 to Lee et al.) and derivatives thereof. Other fluorophores known to those skilled in the art can also be used, for example those available from Life Technologies (Invitrogen; Molecular Probes (Eugene, Oreg.)) and including the ALEXA FLUOR® series of dyes (for example, as described in U.S. Pat. Nos. 5,696,157, 6, 130, 101 and 6,716,979), the BODIPY series of dyes (dipyrrometheneborondifluoride dyes, for example as described in U.S. Pat. Nos. 4,774,339, 5,187,288, 5,248,782, 5,274,113, 5,338,854, 5,451,663 and 5,433,896), Cascade Blue (an amine reactive derivative of the sulfonatedpyrene described in U.S. Pat. No. 5,132,432) and Marina Blue (U.S. Pat. No. 5,830,912).

In addition to the fluorochromes described above, a fluorescent label can be a fluorescent nanoparticle, such as a semiconductor nanocrystal, e.g., a QUANTUM DOTTM (obtained, for example, from Life Technologies (QuantumDot Corp, Invitrogen Nanocrystal Technologies, Eugene, Oreg.); see also, U.S. Pat. Nos. 6,815,064; 6,682,596; and 6,649, 138). Semiconductor nanocrystals are microscopic particles having size-dependent optical and/or electrical properties. When semiconductor nanocrystals are illuminated with a primary energy source, a secondary emission of energy occurs of a frequency that corresponds to the handgap of the semiconductor material used in the semiconductor nanocrystal. This emission can he detected as colored light of a specific wavelength or fluorescence. Semiconductor nanocrystals with different spectral characteristics are described in e.g., U.S. Pat. No. 6,602,671. Semiconductor nanocrystals that can he coupled to a variety of biological molecules (including dNTPs and/or nucleic acids) or substrates by techniques described in, for example, Bruchez et al., Science 281 :20132016, 1998; Chan et al., Science 281:2016-2018, 1998; and U.S. Pat. No. 6,274,323. Formation of semiconductor nanocrystals of various compositions are disclosed in, e.g., U.S. Pat. Nos. 6,927, 069; 6,914,256; 6,855,202; 6,709,929; 6,689,338; 6,500,622; 6,306,736; 6,225,198; 6,207,392; 6,114,038; 6,048,616; 5,990,479; 5,690,807; 5,571,018; 5,505,928; 5,262,357 and in U.S. Patent Puhlication No. 2003/0165951 as well as PCT

Publication No. 99/26299 (published May 27, 1999). Separate populations of semiconductor nanocrystals can be produced that are identifiable based on their different spectral characteristics. For example, semiconductor nanocrystals can be produced that emit light of different colorsbased on their composition, size or size and composition. For example, quantum dots that emit light at different wavelengths based on size (565 mn, 655 mn, 705 mn, or 800 mn emission wavelengths), which are suitable as fluorescent labels in the probes disclosed herein are available from Life Technologies (Carlsbad, Calif.).Additional labels include, for example, radioisotopes (such as ^{3}H), metal chelates such as DOTA and DPTA chelates of radioactive or paramagnetic metal ions like Gd3+, and liposomes.Detectable labels that can be used with nucleic acid molecules also include enzymes, for example horseradish peroxidase, alkaline phosphatase, acid phosphatase, glucose oxidase, beta-galactosidase, beta-glucuronidase, or beta-lactamase.Alternatively, an enzyme can be used in a metallographic detection scheme. For example, silver in situ hybridization (SISH) procedures involve metallographic detection schemes for identification and localization of a hybridized genomic target nucleic acid sequence. Metallographic detection methods include using an enzyme, such as alkaline phosphatase, in combination with a water-soluble metal ion and a redox-inactive substrate of the enzyme. The substrate is converted to a redox-active agent by the enzyme, and the redoxactive agent reduces the metal ion, causing it to form a detectable precipitate. (See, for example, U.S. Patent Application Publication No. 2005/0100976, PCT Publication No. 2005/ 003777 and U.S. Patent Application Publication No. 2004/ 0265922). Metallographic detection methods also include using an oxido-reductase enzyme (such as horseradish peroxidase) along with a water soluble metal ion, an oxidizing agent and a reducing agent, again to form a detectable precipitate. (See, for example, U.S. Pat. No. 6,670,113).Probes made using the disclosed methods can be used for nucleic acid detection, such as ISH procedures (for example, fluorescence in situ hybridization (FISH), chromogenic in situ hybridization (CISH) and silver in situ hybridization (SISH)) or comparative genomic hybridization (CGH).

In situ hybridization (ISH) involves contacting a sample containing target nucleic acid sequence (e.g., genomic target nucleic acid sequence) in the context of a metaphase or interphase chromosome preparation (such as a cell or tissue sample mounted on a slide) with a labeled probe specifically hybridizable or specific for the target nucleic acid sequence (e.g., genomic target nucleic acid sequence). The slides are optionally pretreated, e.g., to remove paraffin or other materials that can interfere with uniform hybridization. The sample and the probe are both treated, for example by heating to denature the double stranded nucleic acids. The probe (formulated in a suitable hybridization buffer) and the sample are combined, under conditions and for sufficient time to permit hybridization to occur (typically to reach equilibrium). The chromosome preparation is washed to remove excess probe, and detection of specific labeling of the chromosome target is performed using standard techniques.For example, a biotinylated probe can be detected using fluorescein-labeledavidin or avidin-alkaline phosphatase. For fluorochrome detection, the fluorochrome can be detected directly, or the samples can be incubated, for example, with fluorescein isothiocyanate (FITC)-conjugated avidin. Amplification of the FITC signal can be effected, if necessary, by incubation with biotin-conjugated goat antiavidin antibodies, washing and a second incubation with FITC-conjugated avidin. For detection by enzyme activity, samples can be incubated, for example, with streptavidin, washed, incubated with biotin-conjugated alkaline phosphatase, washed again and pre-equilibrated (e.g., in alkaline phosphatase (AP) buffer). For a general description of in situ hybridization procedures, see, e.g., U.S. Pat. No. 4,888,278.Numerous procedures for FISH, CISH, and SISH are known in the art. For example, procedures for performing FISH are described in U.S. Pat. Nos. 5,447,841; 5,472,842; and 5,427,932; and for example, in Pir1kel et al., Proc. Natl. Acad. Sci. 83:2934-2938, 1986; Pinkel et al., Proc. Natl. Acad. Sci. 85:9138-9142, 1988; and Lichter et al., Proc. Natl. Acad. Sci. 85:9664-9668, 1988. CISH is described in, e.g., Tanner et al., Am. .1. Pathol. 157:1467-1472, 2000 and U.S. Pat. No. 6,942,970. Additional detection methods are provided in U.S. Pat. No. 6,280,929.Numerous reagents and detection schemes can be employed in conjunction with FISH, CISH, and SISH procedures to improve sensitivity, resolution, or other desirable properties. As discussed above probes labeled with fluorophores (including fluorescent dyes and QUANTUM DOTS®) can be directly optically detected when performing FISH. Alternatively, the probe can be labeled with a nonfluorescent molecule, such as a hapten (such as the following non-limiting examples: biotin, digoxigenin, DNP, and various oxazoles, pyrrazoles, thiazoles, nitroaryls, benzofurazans, triterpenes, ureas, thioureas, rotenones, coumarin, courmarin-based compounds, Podophyllotoxin, Podophyllotoxin-based compounds, and combinations thereof), ligand or other indirectly detectable moiety. Probes labeled with such non-fluorescent molecules (and the target nucleic acid sequences to which they bind) can then be detected by contacting the sample (e.g., the cell or tissue sample to which the probe is bound) with a labeled detection reagent, such as an antibody (or receptor, or other specific binding partner) specific for the chosen hapten or ligand. The detection reagent can be labeled with a fluorophore (e.g., QUANTUM DOT®) or with another indirectly detectable moiety, or can be contacted with one or more additional specific binding agents (e.g., secondary or specific antibodies), which can be labeled with a fluorophore.In other examples, the probe, or specific binding agent (such as an antibody, e.g., a primary antibody, receptor or other binding agent) is labeled with an enzyme that is capable of converting a fluorogenic or chromogenic composition into a detectable fluorescent, colored or otherwise detectable signal (e.g., as in deposition of detectable metal particles in SISH). As indicated above, the enzyme can be attached directly or indirectly via a linker to the relevant probe or detection reagent. Examples of suitable reagents (e.g., binding reagents) and chemistries (e.g., linker and attachment chemistries) are described in U.S. Patent Application Publication Nos. 2006/0246524; 2006/0246523, and 2007/ 01 17153.

It will he appreciated by those of skill in the art that by appropriately selecting labelled probe-specific binding agent pairs, multiplex detection schemes can he produced to facilitate detection of multiple target nucleic acid sequences (e.g., genomic target nucleic acid sequences) in a single assay (e.g., on a single cell or tissue sample or on more than one cell or tissue sample). For example, a first probe that corresponds to a first target sequence can he labelled with a first hapten, such as biotin, while a second probe that corresponds to a second target sequence can be labelled with a second hapten, such as DNP. Following exposure of the sample to the probes, the bound probes can he detected by contacting the sample with a first specific binding agent (in this case avidin labelled with a first fluorophore, for example, a first spectrally distinct QUANTUM DOT®, e.g., that emits at 585 mn) and a second specific binding agent (in this case an anti-DNP antibody, or antibody fragment, labelled with a second fluorophore (for example, a second spectrally distinct QUANTUM DOT®, e.g., that emits at 705 mn). Additional probes/binding agent pairs can he added to the multiplex detection scheme using other spectrally distinct fluorophores. Numerous variations of direct, and indirect (one step, two step or more) can he envisioned, all of which are suitable in the context of the disclosed probes and assays.

Probes typically comprise single-stranded nucleic acids of between 10 to 1000 nucleotides in length, for instance of between 10 and 800, more preferably of between 15 and 700, typically of between 20 and 500. Primers typically are shorter single-stranded nucleic acids, of between 10 to 25 nucleotides in length, designed to perfectly or almost perfectly match a nucleic acid of interest, to be amplified. The probes and primers are “specific” to the nucleic acids they hybridize to, i.e. they preferably hybridize under high stringency hybridization conditions (corresponding to the highest melting temperature Tm, e.g., 50 % formamide, 5x or 6x SCC. SCC is a 0.15 M NaCl, 0.015 M Na-citrate).The nucleic acid primers or probes used in the above amplification and detection method may be assembled as a kit. Such a kit includes consensus primers and molecular probes. A preferred kit also includes the components necessary to determine if amplification has occurred. The kit may also include, for example, PCR buffers and enzymes; positive control sequences, reaction control primers; and instructions for amplifying and detecting the specific sequences.

In some embodiments, the methods of the invention comprise the steps of providing total RNAs extracted from and subjecting the RNAs to amplification and hybridization to specific probes, more particularly by means of a quantitative or semi-quantitative RT-PCR.

In some embodiments, the expression level is determined by DNA chip analysis. Such DNA chip or nucleic acid microarray consists of different nucleic acid probes that are chemically attached to a substrate, which can be a microchip, a glass slide or a microsphere-sized bead. A microchip may be constituted of polymers, plastics, resins, polysaccharides, silica or silica-based materials, carbon, metals, inorganic glasses, or nitrocellulose. Probes comprise nucleic acids such as cDNAs or oligonucleotides that may be about 10 to about 60 base pairs. To determine the expression level, a sample from a test subject, optionally first subjected to a reverse transcription, is labelled and contacted with the microarray in hybridization conditions, leading to the formation of complexes between target nucleic acids that are complementary to probe sequences attached to the microarray surface. The labelled hybridized complexes are then detected and can be quantified or semi-quantified. Labelling may be achieved by various methods, e.g. by using radioactive or fluorescent labelling. Many variants of the microarray hybridization technology are available to the man skilled in the art (see e.g. the review by Hoheisel, Nature Reviews, Genetics, 2006, 7:200-210).

In some embodiments, the nCounter® Analysis system is used to detect intrinsic gene expression. The basis of the nCounter® Analysis system is the unique code assigned to each nucleic acid target to be assayed (International Patent Application Publication No. WO 08/124847, U.S. Patent No. 8,415,102 and Geiss et al. Nature Biotechnology. 2008. 26(3): 317-325; the contents of which are each incorporated herein by reference in their entireties). The code is composed of an ordered series of colored fluorescent spots which create a unique barcode for each target to be assayed. A pair of probes is designed for each DNA or RNA target, a biotinylated capture probe and a reporter probe carrying the fluorescent barcode. This system is also referred to, herein, as the nanoreporter code system. Specific reporter and capture probes are synthesized for each target. The reporter probe can comprise at a least a first label attachment region to which are attached one or more label monomers that emit light constituting a first signal; at least a second label attachment region, which is non-over-lapping with the first label attachment region, to which are attached one or more label monomers that emit light constituting a second signal; and a first target- specific sequence. Preferably, each sequence specific reporter probe comprises a target specific sequence capable of hybridizing to no more than one gene and optionally comprises at least three, or at least four label attachment regions, said attachment regions comprising one or more label monomers that emit light, constituting at least a third signal, or at least a fourth signal, respectively. The capture probe can comprise a second target-specific sequence; and a first affinity tag. In some embodiments, the capture probe can also comprise one or more label attachment regions. Preferably, the first target-specific sequence of the reporter probe and the second target- specific sequence of the capture probe hybridize to different regions of the same gene to be detected. Reporter and capture probes are all pooled into a single hybridization mixture, the "probe library". The relative abundance of each target is measured in a single multiplexed hybridization reaction. The method comprises contacting the tumor sample with a probe library, such that the presence of the target in the sample creates a probe pair - target complex. The complex is then purified. More specifically, the sample is combined with the probe library, and hybridization occurs in solution. After hybridization, the tripartite hybridized complexes (probe pairs and target) are purified in a two-step procedure using magnetic beads linked to oligonucleotides complementary to universal sequences present on the capture and reporter probes. This dual purification process allows the hybridization reaction to be driven to completion with a large excess of target-specific probes, as they are ultimately removed, and, thus, do not interfere with binding and imaging of the sample. All post hybridization steps are handled robotically on a custom liquid-handling robot (Prep Station, NanoStringTechnologies). Purified reactions are typically deposited by the Prep Station into individual flow cells of a sample cartridge, bound to a streptavidin-coated surface via the capture probe,electrophoresed to elongate the reporter probes, and immobilized. After processing, the sample cartridge is transferred to a fully automated imaging and data collection device (Digital Analyzer, NanoString Technologies). The expression level of a target is measured by imaging each sample and counting the number of times the code for that target is detected. For each sample, typically 600 fields-of-view (FOV) are imaged (1376 X 1024 pixels) representing approximately 10 mm2 of the binding surface. Typical imaging density is 100- 1200 counted reporters per field of view depending on the degree of multiplexing, the amount of sample input, and overall target abundance. Data is output in simple spreadsheet format listing the number of counts per target, per sample. This system can be used along with nanoreporters. Additional disclosure regarding nanoreporters can be found in International Publication No. WO 07/076129 and WO07/076132, and US Patent Publication No. 2010/0015607 and 2010/0261026, the contents of which are incorporated herein in their entireties. Further, the term nucleic acid probes and nanoreporters can include the rationally designed (e.g. synthetic sequences) described in International

Publication No. WO 2010/019826 and US Patent Publication No.2010/0047924, incorporated herein by reference in its entirety.

Expression level of a gene may be expressed as absolute expression level or normalized expression level. Typically, expression levels are normalized by correcting the absolute expression level of a gene by comparing its expression to the expression of a gene that is not a relevant for determining the cancer stage of the subject, e.g., a housekeeping gene that is constitutively expressed. Suitable genes for normalization include housekeeping genes such as theactin gene ACTB, ribosomal 18S gene, GUSB, PGK1 and TFRC. This normalization allows the comparison of the expression level in one sample, e.g., a subject sample, to another sample, or between samples from different sources.

In some embodiments, the method of the present invention further comprises determining the expression level of at least one further isoform. In some embodiments, the further isoform is ERα66. In some embodiments, when the expression levels of ERα46 and ERα66 isoforms are determined in the tumor tissue sample, the ratio of the respective levels is determined and is compared to the predetermined reference value wherein detecting differences in the determined ratio and the predetermined reference valueis indicative of the status of breast cancer.

In some embodiments, the predetermined reference value is a threshold value or a cut-off value. Typically, a "threshold value" or "cut-off value" can be determined experimentally, empirically, or theoretically. A threshold value can also be arbitrarily selected based upon the existing experimental and/or clinical conditions, as would be recognized by a person of ordinary skilled in the art. For example, retrospective measurement of expression level of ERα46 in properly banked historical subject samples may be used in establishing the predetermined reference value. The threshold value has to be determined in order to obtain the optimal sensitivity and specificity according to the function of the test and the benefit/risk balance (clinical consequences of false positive and false negative). Typically, the optimal sensitivity and specificity (and so the threshold value) can be determined using a Receiver Operating Characteristic (ROC) curve based on experimental data. For example, after determiningtheexpression level of ERα46 in a group of reference, one can use algorithmic analysis for the statistic treatment of the measured expression levels of the gene(s) in samples to be tested, and thus obtain a classification standard having significance for sample classification. The full name of ROC curve is receiver operator characteristic curve, which is also known as receiver operation characteristic curve. It is mainly used for clinical biochemical diagnostic tests. ROC curve is a comprehensive indicator that reflects the continuous variables of true positive rate (sensitivity) and false positive rate (1-specificity). It reveals the relationship between sensitivity and specificity with the image composition method. A series of different cut-off values (thresholds or critical values, boundary values between normal and abnormal results of diagnostic test) are set as continuous variables to calculate a series of sensitivity and specificity values. Then sensitivity is used as the vertical coordinate and specificity is used as the horizontal coordinate to draw a curve. The higher the area under the curve (AUC), the higher the accuracy of diagnosis. On the ROC curve, the point closest to the far upper left of the coordinate diagram is a critical point having both high sensitivity and high specificity values. The AUC value of the ROC curve is between 1.0 and 0.5. When AUC>0.5, the diagnostic result gets better and better as AUC approaches 1. When AUC is between 0.5 and 0.7, the accuracy is low. When AUC is between 0.7 and 0.9, the accuracy is moderate. When AUC is higher than 0.9, the accuracy is quite high. This algorithmic method is preferably done with a computer. Existing software or systems in the art may be used for the drawing of the ROC curve, such as: MedCalc 9.2.0.1 medical statistical software, SPSS 9.0, ROCPOWER.SAS, DESIGNROC.FOR, MULTIREADER POWER.SAS, CREATE-ROC.SAS, GB STAT VI0.0 (Dynamic Microsystems, Inc. Silver Spring, Md., USA), etc.

Depending on the breast statusdetermined by the method of the present invention, the physician can determined the appropriate form of therapy which typically includes anti-estrogen therapy.As used herein, the term "anti-estrogen therapy" refers to a therapy using an estrogen receptor (ER) antagonist or an estrogen depleting agent. As used herein, the term "estrogen antagonist" refers to a compound that binds to an estrogen receptor, and preferably to the alpha isoform of the estrogen receptor, (either in homodimeric or heterodimeric form), thereby inhibiting or substantially reducing the effect of the respective agonist (estrogen). The estrogen antagonist can be a competitive or noncompetitive antagonist. A competitive estrogen antagonist competes with estrogen (or other agonists) for an estrogen receptor. Competitive or non-competitive estrogen antagonists can be found by common estrogen receptor binding assays, such as those described in the National Institutes of Health (NIH) publication no. 03-4504 (2002) including the protocols provided in the appendix. In some embodiments, the estrogen antagonist is selected from the group consisting of Tamoxifen (purchasable e.g. as Novaldex from Astra Zeneca, or from other manufactures, for example under the trade names Jenoxifen, Kessar, Nourytam, Tamobeta, Tamofen, Tamokadin, Tamoxasta, Tamox-GRY, Tamoxifen AL, Tamoxifen- biosyn, Tamoxifen cell pharm, TamoxifenHeumann, TamoxifenHexal, Tamoxifenmedac, Tamoxifen-ratiopharm, Tamoxigenat, Ta- moximerck, Tamoxistad, Zemide, and the like) ,Raloxifene (purchasable e.g. as Revista from Eli Lilly), Clomifene (purchasable e.g. as Clomhexal from Hexal), Toremifene (purchasable as Fareston from GTx Inc.), Trilostane (purchasable as Modrenal from Bioenvision, UK only) or functional derivatives thereof. In some embodiments, the estrogen antagonist is Tamoxifen or a functional derivative thereof. As used herein, the term "tamoxifen" refers to (Z)-2-[4-(l,2-Diphenyl-l-butenyl)phenoxy]-N,N-dimethylethanamine. In some embodiments, Tamoxifen is to be administered in combination with a hormone such as progesterone for treating proliferative endometrial diseases such as endometrial cancer in order to maximize therapeutic efficacy. Functional derivates are generally obtained from the above compounds by chemical modification. In the case of Tamoxifen, such derivates comprise for example 4- hydroxy-tamoxifen and 4-hydroxy-N-desmethyl-tamoxifen (En- doxifen). Estrogen depleting agents reduce serum estradiol levels in the patient. The aromatase (estrogensyntase) inhibitors, which inhibit the enzyme that converts androgens to estrogens, are an especially important class of estrogen depleting agent. Aromatase inhibitors useful according the present invention include steroidal aromatase inhibitors, such as formestane and exemestane, and non-steroidal aromatase inhibitors, such as anastrozole, vorozole, letrozole and aminoglutethimide. Typically a non-steroidal aromatase inhibitor, such as anastrozole or letrozole, is used. The use of such aromatase inhibitors for the treatment of hormone sensitive breast cancer is known to those of skill in the art, for example anastrozole is administered at a dose of one mg daily and letrozole is administered at a dose of 2.5 mg daily.The anti-estrogentherapy may be performed as a monotherapy or in a combination therapy together with chemotherapy and/or radiation. In some embodiments, the anti-estrogentherapyis to be performed as a monotherapy. In some embodiments, an anti-estrogen therapy is performed as an adjuvant therapy in patients afflicted with endometrium or ovary cancer in order to prevent metastases. In the context of an adjuvant therapy, chemotherapy is commonly performed in accordance with the so-called CNF scheme, using the substances cyclophosphomid, metotrexat and 5- fluoruracil. Alternatively, chemotherapeutic regimens based on anthracycline-containing agents are also available. Other frequently used chemotherapeutic agents include alkylating agents, e.g. ethylenimines and methylamelamines, such as thio- tepa, altretamine, triethylenemelamine, trietylenephosphoramide, triethylenethiophosphaoramide and trimethylolo- melamine; alkyl sulfonates such as busulfanandpiposulfan; nitrogen mustards such as ifosfamide, chlorambucil, estramus- tine, chlornaphazine, cholophosphamide, mechlorethamine, mech- lorethamine oxide hydrochlonde, novembichin, phenesterine, prednimustine, trofosfamide; nitrosureas such as fotemustine, lomustine, carmustine, chlorozotocin, nimustine, ranimustine; aziridines such as carboquone, benzodopa, meturedopa, uredopa; purine analogs such as 6-mercaptopurine, fludarabine, thiogua- nine; pyrimidine analogs such as ancitabine, azacitidine, 6- azauridine, carmofur, cytarabine, dideoxyuridine, doxifluri- dine, enocitabine, floxuridine, 5-FU; and nitrosureas such as carmustine, chlorozotocin, fotemustine, lomustine, nimustine, ranimustine and the like.

A further object relates to a kit suitable for performing the method of the present invention which comprises a binding partner specific (e.g. antibody) for ERα46. In some embodiments, the kit of the present comprises binding partner specific for ERα46 and a binding partner specific for ERα66. In some embodiments, said binding partners are antibodies as described above. In some embodiments, these antibodies are labelled as described above. Typically, the kits described above will also comprise one or more other containers, containing for example, wash reagents, and/or other reagents capable of quantitatively detecting the presence of bound antibodies. Typically compartmentalised kit includes any kit in which reagents are contained in separate containers, and may include small glass containers, plastic containers or strips of plastic or paper. Such containers may allow the efficient transfer of reagents from one compartment to another compartment whilst avoiding cross-contamination of the samples and reagents, and the addition of agents or solutions of each container from one compartment to another in a quantitative fashion. Such kits may also include a container which will accept the tumor tissue sample, a container which contains the antibody(s) used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, and like), and containers which contain the detection reagent.

The invention will be further illustrated by the following figures and examples. However, these examples and figures should not be interpreted in any way as limiting the scope of the present invention.

FIGURES:

Figure 1: Evaluation of antibodies used for human breast cancer diagnosis for the specific recognition of ERα isoforms. A) Schematic representation of the ERα66 and ERα46 isoforms. The location of the epitopes used for the generation of antibodies is indicated where known. B) Representative picture of western blot assays performed using the SP1 antibody to test for the expression of the two ERα isoforms in both MDA-MB-231 (MDA) cells transfected with plasmids encoding either the ERα46 (MDA-46kDa) or the ERα66 (MDA-66kDa) isoforms or MCF7 cells which express both isoforms. C) The different antibodies used in breast cancer diagnosis (1D5,6F11, SP1 and EP1) were tested for their ability to

recognize either ERα66, ERα46 or both isoforms in immunohistochemistry experiments performed in MCF7, MDA-46kDa and MDA-66kDa cells. D) Representative picture of western blot experiments evaluating the expression of both ERα isoforms in MCF-7 cells, as determined by the different antibodies indicated. HC-20 antibody commonly used on this procedure was added. Protein extracts prepared from MDA cells were used here as an ERα-negative control.

Figure 2: Evaluation of the relative expression of the ERα46 and ERa66 isoforms in human ERα-positive breast tumors. To determine the relative expression of the ERα46 and ERα66 isoforms in breast cancers, western blot experiments were performed with the SP1 anti-ERα antibody on protein extracts prepared from 116 ERα-positive and 20 ERα-negative human breast tumors. A) Representative image of western blot of human ERα-positive breast tumors samples blotted with an anti-ERα antibody (SP1). One result representative of an ERα-negative tumor is also shown. The MDA-ER66 cell line that co-expresses ERα66 and ERα46 was used as a positive control. The ratio of expression of ER46 compared to ER66 expression is indicated below each lane. B) The results obtained in these experiments were summarized and expressed as a percentage of tumors tested expressing or not ERα46 in combination with ERα66. C) Distribution of the ERα46/ERα66 ratio of expression values within the 116 tumor samples. D) The number of tumors expressing ERα46 or (E) the abundance of ERα46 expression over ERα66 expression were classified depending upon tumor grade (I, II or III). F) The ERα46/ERα66 ratio of expression was correlated as a function of tumor size or G) classified depending on whether the size of the tumor was < or > 2cm.

Figure 3: Proteomic analysis of the 46kDa protein detected in tumor samples. A) Western blot with the SP1 antibody obtained after immunoprecipitation of ERα in two human tumors overexpressing the putative ERα46 isoform. Bands corresponding to the 46kDa and 66kDa isoforms were then excised and used for proteomic analysis. MCF7 lysates were used as positive controls. B) Proteomic analysis results from the different cell lines and the four breast tumor samples.

Figure 4: IRES-mediated generation of ERα46. A) MDA cells were transduced with constructs expressing either ERα46 or wild-type ERα66, or an ERα66 cDNA harbouring mutated ATG2/3 codons (ERα46°). A representative image of western blot experiments is shown here, illustrating the loss of expression of the ERα46 isoform in cells transduced with the vector encoding ERα46°. B) Scheme illustrating the construction of a bi-cistronic pTRIP-CRF-L2 vector allowing CMV-driven expression of the LucR gene separated from the

LucF reporter by the subcloned A/B domains of *ESR1*. C) The A/B domain of the *ESR1* gene includes an IRES that responds to stress. MDA cells were transduced with lentivirus containing the LucR-IRES-LucF construct. These cells were then subjected to 6 hours of stress with DTT (3, 5 or 8mM) or thapsigargin (10, 50 or 500nM) or to a 6 hour treatment with vehicle control (PBS for DTT and DMSO for thapsigargin). Both luciferase reporter gene activities were measured and the results of these experiments were expressed as the LucF/LucR ratio. **: significantly different from control (Kruskal-Wallis).* D) Representative western blot of the increased expression of ERα46 observed in stressed MDA cells transfected with cDNA expressing ERα66.

Figure 5: Possible mechanisms to generate the ERα46 and epitope mapping of the antibodies used in this work. From the *ESR1* gene, ERα46 can be generated either by alternative splicing of mRNA (1), by IRES-dependent translation from the longer ESR1 transcript (2), or by proteolysis of ERα66 by as-yet unknown proteases (3). It should be emphasized that these three different mechanisms could concur to generate the ERα46 isoform - a unique feature to the best of our knowledge.

Figure 6: ERα46 limits mammary tumorigenesis. Kaplan-Meir curve at the age at first detection of MMTV-PYMT-ERαAF1+/+ (n=16) and MMTV-PYMT-ERαAF1+/° (n=23) mammary tumors.

EXAMPLE 1:

Methods:

Cell culture and Immunohistochemistry

The ERα-negative MDA-MB-231(MDA) cell line stably expressing either the ERα46 or ERα66 isoforms (MDA-46kDa and MDA-66kDa) were kindly obtained from G. Flouriot and R. Métivier(21). We also transiently transfected MDA cells with either the pcDNA3.1 plasmid, encoding the cDNA of ERα-46, or a mutated version, ERα-46kDa°, using the *TransIT*®-BrCa Transfection Reagent (Euromedex, France). This mutated plasmid was obtained by subcloning a fragment from ER66 that contains an AUG/UCG mutation at position 174 and an AUG/AUA mutation at position 176 (15).Both the ERα-positive MCF7 cell line and MDA cells expressing the different ERα isoforms were maintained in DMEM (4.5g/l glucose) supplemented with 10% fetal bovine serum and antibiotics (100U/ml penicillin, 100μg/ml streptomycin with an additional0.5mg/ml hygromycin for transfected MDA cells), at 37°C in a 5% CO_2 humidified atmosphere. For immunohistohemistry (IHC), cells were formalin-fixed and paraffin-embedded using the Shandon™ Cytoblock™ Cell Block Preparation System, according to the manufacturer's protocol. Immunohistochemistry was performed with a DakoAutostainer Link 48 on 3µm sections. The panel of antibodies tested is provided in Table 1. Antigen retrieval was performed using aDako PT Link pressure cooker in pH 6.0 citrate buffer. An EnVision™ system was used for antibody detection. The panel of antibodies used wasSP1(Abcam), HC20 (SantaCruz), 6F11 (Novocastra), 1D5 and EP1 (Dako).

Human breast cancer sample collection

The retrospective study used tumors samples from patients diagnosed with ERα-positive invasive breast carcinoma established on a previously performed biopsy. The diagnosis was performed with the 6F11 antibody, between 2011 and 2014.These samples were obtained from the "CRB cancer des Hôpitaux de Toulouse; BB-0033-00014" collection and has been approved by Institutional and national ethics committees. Patient samples were obtained after authorization by the French Ministry of Higher Education and Research (declaration DC 2009-989; DC-2011-1388; transfer agreement AC-2008-820; AC-2011-130) and from the "Comité de Protection des PersonnesSudOuest et OutreMer- l'AgenceRégionale de Santé Midi-Pyrénées, Toulouse". Clinical and biological annotations were consistent with the Comité National Informatique etLibertés guidelines. Tumors included on this retrospective studyhad a size greater than or equal to 1centimeter to enable tissues not needed for diagnosis to be included. They were frozen in 1.5ml cryotubes using the SnapFrostII™ (Excilone, France) and stored at -80°C. Patients with an ipsilateral recurrence of breast cancer, treated with neoadjuvant chemotherapy, who have received chemotherapy treatment for another disease or who received thoracic radiation therapy (recurrence or another pathology) were excluded. All tumors were classified by the anatomopathologist (HER2 status, tumor size, ERα overexpression, lymph node involvement, histological type)and were graded according to WHOguidelines. The analysis was performed on a series of 116 ERα-positive invasive ductal orlobular breast carcinoma (22 grade I, 60 grade II and 34 grade III).

Western blots on tumor samples:

5 mm of each tumorwas lysed in cold lysis buffer (150mM NaCl, 50mM Tris-HCl pH7.5, 0.1% SDS, 1mM EDTA, 5mMNaF, 1mM orthovanadate, 0.5mM DTT, and protease inhibitors) using a PrecellysHomogeniser(Bertin Technologies, France). Proteins were separated on a 10% SDS-PAGE gel and transferred to a nitrocellulose membrane. After blocking, membranes were incubated overnight with the primary antibody (SP1, Santa Cruz). Subsequently, blots were incubated with an HRP-conjugated secondary antibody (anti-rabbit, Cell Signaling) and visualized by ECL detection, according to the manufacturer's instructions

(Amersham Biosciences, CT), using a ChemiDoc™ Imaging System (Bio-Rad). Bands corresponding to ERα46 and ERα66 were quantified using ImageJ densitometry and the ratio of the band intensities was calculated.

Endoplasmic reticulum stress

MDA-66kDa cells were seeded in 6-well plates. At 80% confluence, the cells were subjected to 6h of stress with DTT (Euromedex) or Thapsigargin (Sigma) as indicated, then rinsed in PBS and lysed by sonication in SDS buffer (5% SDS, 10% glycerol, 80mM Tris pH6.8). Western blot analyses were then performed on 4–20% denaturing polyacrylamide “stain-free” gels (BioRad).

Immunoprecipitation and proteomic analysis

ERα-enriched protein fractionsfrom tumor protein extracts were obtainedthrough immunoprecipitation using the anti-human ERα primary HC20 antibody. Following their purification using Protein G sepharose beads, a first western blot was performed to check the efficiency of the immunoprecipitation. In parallel, the immunoprecipitate was diluted with Laemmli buffer, then separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using a short and low-voltage electrophoretic migration. After Instant Blue staining, the bands corresponding to ERα46 and ERα66 were respectively excised from the gel. Proteins were in-gel digested by trypsin, and resulting peptides were extracted from the gel and analyzed by nano-LC–MS/MS using an ultimate 3000 system (Dionex, Amsterdam, Netherlands) coupled to an LTQ-OrbitrapVelos mass spectrometer (Thermo Scientific, Bremen, Germany).

Plasmids, lentiviral production and luciferase assay

cDNA coding for the A/B (aa2-173) domain of the human *ESR1* gene encoding ERα was amplified by PCR and cloned into the *Spe*I and *Nco*I sites of the pTRIP CRF1AL2 bicistronic mRNA vector that encodes both theRenilla luciferase (LucR) and Firefly luciferase (LucF) genes separated by this putative IRES-ERα sequence (22). The final construct was verified by sequencing. In such a transgene, LucR expression is cap-dependent whereas LucExpression is IRES-dependent, thusthe level of IRES activity can be deduced from the LucF/LucR ratio.The production of lentiviral particles was performed in HEK293 cells. Transduced MDA-MB 231 cells (MDA-A/B) were subjected to ER stress as indicated. After a PBS wash, cells were frozen at -80°C. Luciferase measurements were performed with a LB960 luminometer (Berthold) using the dual reporter assay kit (E1960) (Promega) according to the manufacturer’s recommendations.

Coregulator-peptide interaction profiling

Ligand-mediated modulation of the interactions between the ERα46 and ERα66 proteins and their coregulatorswas characterized by a MARCoNI (Microarray Assay for Real-time Coregulator-Nuclear receptor Interaction, PamGene International BV, the Netherlands). This method has been describedpreviously(23, 24).

Statistical analyses

Categorical variables are reported in terms of frequencies and percentages, continuous variables are presented by median and range values. Comparisons between groups were performed using the Mann-Whitney rank sum test for continuous variables. Correlations between continuous variables were evaluated using the Spearman's rank correlation test. All p-valuesare two-sided. For all statistical tests, differences were considered significant at the 5% level. Statistical analyses were performed using the STATA 13.0 software (STATA Corp, College Station, TX).

Results:

Characterization of the anti-ERα antibodies commonly used for breast tumor diagnosis

The ERα46 isoform is devoid of the A/B domain and thus lacks one of the major activation functions (AF-1).The rest of its sequence from amino acids 173 to 595 at the C-terminalis completely identical to ERα66 (Fig. 1A). Therefore, to characterize the expression of ERα46 in breast tumors an antibody must be used that is directed against the C-terminal domain. One of the first monoclonal antibodies to be used against ERα for tumor diagnosis was 1D5(25). However, it is directed against the A/B domain so is unable to recognize the ERα46 isoform (26) (Fig. 1A). Later on, the respective murine and rabbit monoclonal antibodies 6F11 and SP1 were extensively used for diagnosis, with improved specificities compared to the 1D5 clone (27, 28). More recently, the monoclonal rabbit antibody EP1has been commercialized. However, whereas SP1 is known to have been obtained after immunization with a peptide directed against the C-terminal domain, the abilities of the 6F11 and EP1 antibodies to recognize ERα46 have not, to our knowledge, been reported.

To test this, we used control ERα-negative MDA-MB-231cells, or MDA cells engineered to stably express either the ERα46 or ERα66 isoform, alongsideMCF7 cells co-expressing both proteins (Fig. 1B). Interestingly, a small amount of ERα46 expression was found in MDA-66kDa cells. As previously suggested, this is presumably due to translation through the IRES in the first exon (15). Immunohistochemistry performed on these four cell lines demonstrated that, among the four tested antibodies (1D5, 6F11, SP1 and EP1), only SP1 was able to specifically detect the ERα46 isoform in the MDA-ER46kDa cells (Fig. 1C). The immunoreactivity of the different antibodies were also tested via western blot alongside the HC-20 antibody (commonly used on this procedure but not used for diagnosis as it is polyclonal rabbit antibody) (Fig. 1D).The HC-20 and SP1 antibodies recognized both the ERα46 and ERα66 isoforms whereas the EP1 and 6F11 antibodies were unable to detect ERα46, in line with the immunohistochemistry results. We also noticed that the 1D5 antibody had a quasi-undetectable reactivity when used in this procedure. Altogether, these data demonstrate that from the set of antibodies commonly used for breast cancer diagnosis, the SP1 antibody is the only one able to recognize the 46kDa isoform by immunohistochemistry.

Detection of ERα46 expression and abundance in human breast carcinoma cells

We next examined whether the ERα46 isoform could be detected in a series of 116 ERα-positive breast tumor samples (characterized with the 6F11 antibody). We also compared the relative abundance of the ERα46 and ERα66 isoforms by western blotting. Most of the breast tumors(70%) expressed both the ERα46 and ERα66 isoforms, at varying levels (Fig. 2A and 2B). The ERα46/ERα66 ratio varied from 0 to 3.48, with an average of 0.37. Furthermore, even though the vast majority of tumors expressed lower levels of ERα46 than ERα66, 10% of the tumors tested expressed predominantly the shorter isoform (Fig 2C).

We next analyzed the relationship between clinical parameters (grade and size of tumor) and ERα46 expression. We found that high-grade tumors correlated with lower ERα46 expression since91% of Grade I tumors expressed ERα46,whereas this figure was 75 and 62% for Grades II and III, respectively (Fig. 2D). Moreover, the ERα46/ERα66 ratio of the relative expression of these isoforms was also significantly higher in low-grade tumors which are differentiated compared to tumors of Grades II and III which are poorlydifferentiated ($p=0.0024$ and $p=0.005$ respectively; see Fig. 2E). The abundance of ERα46 was also inversely correlated with tumor size (Fig. 2F). Finally, we classified our samples using a size parameter usually used by the American Joint Committee on Cancer (AJCC) to characterize tumor evolution, which is set at a 2cm cut-off. Using this classification, we found that ERα46 expression was higher in small-sized tumors compared to tumors greater than 2cm in diameter ($p=0.039$; Fig. 2G).

Altogether, this data obtained by analyzing the expression of the ERα46 polypeptide in a panel of 116ERα-positive breast tumors highlights the fact that ERα46 was expressed in more than 70% of cases. Furthermore, although the expression of this short isoform was highly variable, it correlated with the tumor evolution stage with a higher expression in low-grade tumors and lower expression in tumors that were larger, less differentiated and of higher-grade.

Expression of ERα46 was not detected in breast tumors initially diagnosed as ERα-negative

A few studies have shown that 8% of tumors diagnosed as ERα-negative using the 1D5 antibody were actually positive for ERα when tested with next-generation antibodies such as SP1(27-29). Importantly, the authors of these studies did not take into account the presence of ERα46, which cannot be detected by the 1D5 antibody. Therefore, we hypothesized that these tumors may have indeed lost the expression of ERα66 but not ERα46. We explored this possibility by evaluating the expression of the ERα46 isoform in a series of 19 tumors identified as ERα-negative using the 6F11 antibody. However, none of these samples was found to express the short ERα46 isoform. A representative sample is shown in Fig. 2A.

Characterization of the ERα46 isoform

Although the results obtained by western blot analysis show bands at the expected sizes, this in itself does not definitively prove that the observed signal at 46kDa in tumors was indeed the short isoform of ERα. Hence, we excised and sequenced the proteins in each band.

To reach these aims, we purified the ERα proteins from MCF7 cells and from lysates of four tumor samples, by immune precipitating the two ERα isoforms using the anti-human ERα primary HC20 antibody(Fig. 3A). After separation by SDS-PAGE, the gel bands corresponding to the 46kDa and 66kDa proteins were excised and further digested for proteomic analysis. In MCF7 cells (Fig. 3B),24.4% of the ERα66 sequence was detected, including a peptide covering the N-terminal domain of amino acids 9-32. Importantly, and as expected, although 23.3% of the ERα46 sequence was detected, no peptide from the N-terminal A/B domain was identified. Proteomic analysis of immune precipitated ERα proteins from four tumor samples respectively identified up to 25% of the ERα66 sequence and 15.3% of the ERα46 sequence (Fig. 3B). Again, although peptides 184-206, 402-412 and 450-457 were found in the 46kDa-sized band, no peptides located before the second ATG (position 174) were detected. These were detected in the 66kDa-sized band from these tumor samples. Therefore, although we were unable to characterize the start codon of ERα46, we confirm for the first time that the 46kDa band identified in Western blot analyses of ERα-positive tumors is without doubt a shorter isoform of ERα.

ERα-46kDa can be expressed following IRES translation in response to stress

A number of cancer-related proteins are now known to be translated via an Internal Ribosomal Entry Site (IRES), and accumulating evidence suggests that IRES-mediated translation may be of particular importance to malignant cells (30, 31). It has already been proposed that this mechanism participates in ERα46 generation in transfected COS cells(15).

Our results also suggest that IRES-mediated translation likely occurs in MDA cells transfected with full-length ERα66 since a 46kD ERα–positive protein was also detected by western blot analysis (Fig. 1B). This hypothesis was further supported by transfecting MDA cells with the ERα46° plasmid in which the two potential initiation codons (AUG174/176) are mutated.ERα46 expression was lost in these cells (Fig.4A).

We next sought to determine how this putative IRES sequence upstream of the AUG174/176 can be stimulated. Activation of IRES sites in tumor cells has been shown in cells continually subjected to diverse stress conditions of the tumor microenvironment (31). Therefore, we investigated whether stress conditions could activate such a mechanism for translation of ERα46. A series of experiments was carried out using a bicistronic plasmid encoding both the Renilla luciferase (LucR) and Firefly luciferase (LucF) genes, between which we inserted the cDNA encoding the A/B domain (aa 2-173) that contains this putative IRES-ERα sequence (Fig. 4B). In this construct, LucF expression is used to report levels of translation induced by sequences within the A/B domain, thus a higher LucF/LucR ratio indicates a higher induction of expression. MDA cells transduced with this construct were submitted to two well-characterized tress inducers, dithiothreitol (DTT) and thapsigargin (a more specific inducer of endoplasmic reticulum stress) (32). Interestingly, transient exposure to DTT induced a 3-fold increase in the LucF/LucR ratio (Fig. 4C) which correlated with an increase in ERα46 protein expression (Fig. 4D). Similar results were obtained with thapsigargin, whose treatment resulted in a 2.5-fold increase in the LucF/LucR activity ratio and an increase in ERα46 expression (Fig. 4C and 4D). As a control, we used siRNA directed against LucR which diminished LucF activity, demonstrating the absence of either an internal promoter in the intervening sequence or stress-induced cryptic alternative splicing that could have shunted the LucRcistron. Taken together, these data showed that ERα46 can be produced by stress inducers via an IRES-dependent mechanism.

Study of cofactors that differentially interact with the ERα66 and ERα46 isoforms

Previous reports have suggested that ERα46 antagonizes the proliferative action of ERα66 in different cancer cell lines treated with E2 (8, 16,33). This inhibition may occur through the differential recruitment of coregulators by the ERα46 and ERα66 isoforms in the cellular responses induced by E2. To test this hypothesis, we used a MARCoNI assay to characterize the interaction of the two ERα isoforms with their coregulators and any modulation of this by binding to their ligand (E2). This assay allowed us to analyze the binding of both receptor isoforms to 154 unique coregulator-derived motifs, both in their unliganded (apo) conformation or with concentrations of E2 ranging from 10^{-12} to 10^{-7}, corresponding to full ligand saturation and receptor activation (23). The resulting overall binding patterns indicated that, qualitatively, the receptors bind to the same subset of co regulators, with a clear response of the ERα46 isoform to E2. However, an isoform-selective difference in the binding levels of both apo and fully-activated receptors was clearly observed (Fig. 5A). Further analysis of the E2 response curves evidenced that: *(i)* both isoforms behave similarly for some interactions (with BRD8 for instance); *(ii)*some peptides bind better toone of the isoforms, for exampleNCOA3(also named SRC-3) binds better to ERα66 and PRGC1 to ERα46; and *(iii)* some cofactors bind to both isoforms equally in their apo conformation, but increasing E2 concentrations favor their association with one or the other, as observed for the binding of EP300 to ERα66 or NROB2 to ERα46.The hierarchical clustering of ligand-induced modulation of coregulator interactions was then performed to look for differences and was quantified as the log-fold change in binding (modulation index, MI). This analysis confirmed that although qualitatively the overall responses looked generally quite similar, there is a quantitative differential modulation with some selective preference to certain coregulators. Upon E2 binding, an overall increased preference for cofactor binding to ERα66 over ERα46 was observed when the EC50 values were plotted.

We then investigated the potency of the ERαantagonists 5-OH-tamoxifen and fulvestrant in inhibiting cofactor binding to ERα46 and ERα66 in the presence of E2. The profile of the EC_{50}valuesfor 4-OH-tamoxifen and fulvestrant clearly showed a better efficacy of 4-OH-tamoxifen than fulvestrant in inhibiting E2-induced binding of the receptor isoforms to coregulators. However, the potencies of these antagonists to inhibit binding to ERα46 and ERa66 were comparable. Altogether, these data clearly demonstrate that the two isoforms show some specificity and heterogeneity in terms of their binding to coregulators.

Discussion

The work reported here aimed to analyze the expression levelsand characteristics of the hidden ERα46 isoform in breast tumor samples. We have clearly shown that ERα46 is expressed in the majority of human breast tumors tested (more than 70%). We have also established that the expression levels of ERα46 are highly variable, sometimes even prominent over those of ERα66.Importantly, the ERα46/ERα66 expression ratio negatively correlated with tumor grade: poorly-differentiated tumors (of higher grade and larger size)presented lower amounts of ERα46. These data indicate that this shorter isoform, previously ignored in diagnoses and analyses, is significantly expressed in breast tumors and may have an impact on diagnosis. Unfortunately, this retrospective study started in 2011 so it is too early to further analyze any correlation between the abundance of ERα46 and overall survival or recurrence of disease. This criterion requires a time period of15-20years due to delayed tumor relapses of ERα-positive tumors (32).

A short study previously reported the expression of a 47kDa isoform in human breast cancer that is able to bind to radioactive tamoxifen aziridine, which could be the same as the 46kDa ERα isoform described here(22). Using electrophoresis with radiolabeled tamoxifen, they found that 49% of tumor samples were found to express the two isoforms, whereas 36% contained only the longest form. Our proteomic analysis is the first to definitively identify the band detected by western blotas an ERα isoform deprived of the A/B domain. In our present study, based on western blotting, we found that the majority of tumors (70%) expressed the 46kDa isoform, potentially a consequence of the greater sensitivity of our approach.

Our results also show that the antibodies currently used for the diagnosis of breast cancer are unable to detect the ERα46 isoform. Indeed, the 1D5, 6F11 and EP1 antibodies are directed against the A/B domain, which is absent in the ERα46 isoform (Fig. 5). As a consequence, the hormone-dependent characterization of the tumor is only based on the level of expression of the ERα66 isoform. Using a short series of 20 breast tumors that were initially classified as ERα-negative using the 6F11 antibody, we found no evidence of tumors expressing only the ERα46 isoform, but this study should be extended to a larger sample series (in process). Furthermore, we found that ERα46 expression level was related to tumor size, suggesting that the expression of the 46kDa isoform in breast tumor could be associated with a limited tumor growth. Such a hypothesis is supported by previous studies demonstrating that ERα46 antagonizes the proliferative effects induced by ERα66 activation both *in vitro* in MCF7 cells (16)and SaOS osteosarcoma cells (8), as well as in colorectal tumor tissues (33). Its expression could therefore maintain a low tumor volume by driving an apoptosis mechanism related to cell cycle arrest (34).Importantly, other studies have linked the N-terminal region of ERα with cell proliferation. Merot et al. (35)used *in vitro* systems to show that the respective contribution of AF-1 and AF-2 towards ERα transcriptional activity varies upon thes tage of cell differentiation, with ERα AF-1 required for E2-induced proliferation. This key role of AF-1 was also demonstrated physiologically in the uterus, a tissue that is highly sensitive to the proliferative actions of E2 *in vivo*. Indeed, it was shown that E2 had no proliferative action on uterine epithelial cells in ERα-AF1° mice, which express only a 49kDa ERα isoform with its AF-1 domain deleted (19).ERαhas been described as being at the crossroads of paracrine or autocrine growth factor and endocrine estrogenic signaling (33), and its activity can be controlled in the absence of E2 through phosphorylation cascades induced by IGF-1, EGF or FGF-2. Importantly, all residues of ERα that have so far been implicated in these E2-independent responses or in the modulation of ERα activities in response to growth factor signaling are located within the N-terminal region of the protein (36). Altogether these data support the hypothesis that AF-1 is the region of ERα required for cell proliferation, and that its absence in the ERα46 isoform is likely to confer specific properties to this protein compared to the ERα66 isoform. This may be of critical importance for breast cancer diagnosis and design of treatment.

In our study, we also analyzedthe ability of the two isoforms to bind cofactors using theMARCoNI assay. We clearly demonstrate that the two isoforms showsome heterogeneity in terms of binding to coregulators.These data emphasize the fact that, as already reported byPenot et al (16), ERα46 and ERα66 recruit specific coregulators even though ERα-66 and ERα46 homodimers similarly activated the transcription of an ERE-driven reporter gene in the presence of E2. Specific cofactor recruitment has also been suggested by the characterization of two AF-1-specific coactivators, the p68 and p72 RNA helicases (13, 24). The ability of the two ERα isoforms to recruit transcription factors to the *pS2/TFF1* promoter was also compared by Re-ChIP-experiments(21).This study identified that ERα46 specifically recruited components of the Sin3 repressive complex (NCOR/SMRT) to the *TFF1* promoter in the absence of E2. This was associated with specific inhibition of the basal transcription of this gene by the ERα46 isoform. More recently, the quaternary structure of a biologically-active ERα-coactivator complex on DNA has been determined by cryoelectron microscopy (37). This study showed the location of the AF-1 domain in the complex, which supports a role in the recruitment of the coactivator SRC-3. Interestingly, in our assay we also found a stronger binding of ERα66 to some peptides derived from SRC-3 (Fig. 5B). In contrast, the ERα46 isoform was found to bind to NRB02 better than the ERα66 isoform. NRB02 acts as a negative regulator of receptor-dependent signaling pathways. These data therefore underline the importance of the AF-1 domain for full transcriptional activation of the ERα-coactivator complex.

Our observations raise the hypothesis that the presence of the short ERα46 isoform in breast tumors could indicate a more favorable prognosis. Such an assessment is also supported by the study by Klinge et al which indicated that almost 40% of patients developing a secondary tamoxifen resistance exhibit a reduced expression of ERα46 (12). This supports the idea that endocrine resistance is associated with a decreased expression of ERα46 and thus with poor breast cancer prognosis. Subtle interactions between these isoforms could influence the action of Selective Estrogen Receptors Modulators (SERMs) against tumor growth and metastasis.

Interestingly, tamoxifen antagonizes the AF-2 of both the ERα66 and ERα46 isoforms, but at the same time it partially activates the AF-1 of ERα66, and cannot elicit such an action on the AF-1-deficient ERα46 isoform.As shown in our analysis with a MARCoNI assay, ERα46 appears to be as potent as the ERα66 in dissociating coregulatory binding in response to tamoxifen or fulvestrant. However, since these interactions are very complex, further investigations are needed in tumor samples *in vivo*.

Altogether, these data point out the importance of the expression of both ERα isoforms in breast tumors. Hence, their characterization could help inrefining both prognosis and therapy. Although the exact mechanisms accounting for the expression of the ERα46 isoform still remain to be clarified, three potential processes have been identified: (*i*) alternative splicing could generate an mRNA deficient in the nucleotide sequence corresponding to exon 1 encoding the A/B domain(14); *(ii)* proteolysis, as has also been suggested in human breast tumors (13) and in the mouse uterus (38); and*(iii)* anIRES located within the full length mRNA could allow the initiation of translation at a downstream ATG which encodes methionine 174 in the human ERα66(15) (Fig. 6). Unfortunately, our proteomic approach did not identify peptides close to this initiation codon. One possible reason that could explain the absence of such peptide in our analysis may be the presence of many lysine and aspartic acid residues present within the 171-190 sequence surrounding the two potential initial codons (KGS**M**A**M**ESAK). Trypsin proteolysis targets these residues, so could have led to the generation of peptides shorter than 8 residues which would then have been difficult to identify during proteomic analysis. At this stage, it is not possible to determine the respective roles of the different mechanisms of ERα46 generation. However, we have clearly demonstrated that under stress the IRES within the ERα mRNA can be activated and lead to the generation of ERα46. Defects in the cellular transcription program have been already described in tumor cells, leading to the activation of such IRES-mediated initiation (39). Moreover, 4E-BP1, a negative regulator of cap-dependent mRNA translation, was found to be overexpressed in breast tumors compared to healthy epithelium, suggesting that translational mechanisms such as IRES might be active(40).

In conclusion, this study has demonstrated that a shorter ERα46 isoform is frequently coexpressed in ERα-positive breast tumors, as revealed by western blot analysis. This is currently the only technique that can detect and evaluate the relative amounts of each isoform. More careful attention must therefore be taken in the choice of antibodies used for immunohistochemistry to consider the expression of the ERα46 isoform. Although the prominent mechanisms of ERα46generation remain to be defined, we have clearly demonstrated that this shorter isoform can differentially bind to co-regulators in response to E2 or SERMs. It should be emphasized that three different mechanisms can concur to generate the ERα46 isoform, a unique feature to the best of our knowledge (Fig. 5). This probably indicates the importance of this shorter isoform in E2 signaling and its proliferative actions. Our results suggest an inhibitory effect of ERα46 on E2-induced proliferation which is in accordance with the withdrawal of the AF-1region of ERα, a domain that is very important for mediating the proliferative action of several growth factors which can activate ERαeven in the absence of estrogens. Thus, due to the importance of ERα and hormonal treatments in the management of breast cancers, the ERα46isoform could represent an additional prognosis factor and could have an influence on anti-hormonal treatment.

EXAMPLE 2:

Methods:

B6.FVB- Tg(MMTV-PyVT634Mul/LellJ) mice were purchased from the Jackson Laboratories. These mice were crossed with the ERα-AF1°/° mice previously described (Billon et al., 2009)to generate the PYMT-WT mice and the PyMT-AF1+/° mice (expressing one allele of ERα WT(ERα-66kDa) and one allele of ERα-AF1° (ERα-46kDa)). The MMTV-PyMT transgene was carried by male mice in all breeding pairs. PyMT -WT and PyMT-AF1+/° were on a pure C57BL/6J background. All mice were cared for according to guidelines established by the INSERM Institute.

MMTV-PYMT positive female mice were palpated for tumors twice a week starting at 7 weeks of age. Tumor onset were measured upon detection of any palpable tumor mass. Tumor size were measured by calliper and tumor endpoint was defined at 25 weeks when all mice were sacrified and tumor weights were analyzed. Kaplan Meyer curve of the age at first detection and the log-rank statistical test using a Mantel-Cox Log-rank test were used to determine differences in age at first detection.

Results:

Our previous studies have demonstrated that ERα-46kDa expression is sometimes prominent over ERα66, in particular in differentiated, smaller grade and smaller size tumors in human breast tumor samples. Moreover, expression of ERα-46KDa isoform in human breast tumor cell lines support an inhibitory action of ERα46 on E2-dependent proliferative response of ERα66.

However, direct in vivo action of the ERα-46 kDa isoform in carcinogenesis has been lacking. To this end, the ERα-AF1°/° mice previously described (Billon et al., 2009) was introduced, through interbreeding, into a widely used transgenic mouse model of mammary tumorigenesis [MMTV-PyMT (PyMT)] that represent well-accepted models of human breast cancers . Early tumor development was assessed by palpation and age at first detection was found to be delayed in mice harboring the ERα-46KDa expression in addition to the ERα66KDa (Time= 130 days as opposed to Time=110 days for WT mice, see Figure 6).

Together our data provide the first in vivo evidence that ERα46 plays a critical role in limiting the breast tumor appearance, identifying a new marker for human breast tumor.

REFERENCES:

Throughout this application, various references describe the state of the art to which this invention pertains. The disclosures of these references are hereby incorporated by reference into the present disclosure.

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer Journal international du cancer. 2015;136:E359-86.
2. Johnston SR. New strategies in estrogen receptor-positive breast cancer. Clin Cancer Res. 2010;16:1979-87.
3. Fisher B, Anderson S, Redmond CK, Wolmark N, Wickerham DL, Cronin WM. Reanalysis and results after 12 years of follow-up in a randomized clinical trial comparing total mastectomy with lumpectomy with or without irradiation in the treatment of breast cancer. The New England journal of medicine. 1995;333:1456-61.
4. Howell A. Adjuvant aromatase inhibitors for breast cancer. Lancet. 2005;366:431-3.
5. Silvestrini R, Daidone MG, Luisi A, Boracchi P, Mezzetti M, Di Fronzo G, et al. Biologic and clinicopathologic factors as indicators of specific relapse types in node-negative breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1995;13:697-704.
6. Polyak K, Vogt PK. Progress in breast cancer research. Proceedings of the National Academy of Sciences of the United States of America. 2012;109:2715-7.
7. Jordan VC, O'Malley BW. Selective estrogen-receptor modulators and antihormonal resistance in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25:5815-24.
8. Denger S, Reid G, Kos M, Flouriot G, Parsch D, Brand H, et al. ERalpha gene expression in human primary osteoblasts: evidence for the expression of two receptor proteins. Molecular endocrinology. 2001;15:2064-77.

9. Murphy AJ, Guyre PM, Wira CR, Pioli PA. Estradiol regulates expression of estrogen receptor ERalpha46 in human macrophages. PloS one. 2009;4:e5539.

10. Li L, Haynes MP, Bender JR. Plasma membrane localization and function of the estrogen receptor alpha variant (ER46) in human endothelial cells. Proceedings of the National Academy of Sciences of the United States of America. 2003;100:4807-12.

11. Jiang H, Teng R, Wang Q, Zhang X, Wang H, Wang Z, et al. Transcriptional analysis of estrogen receptor alpha variant mRNAs in colorectal cancers and their matched normal colorectal tissues. The Journal of steroid biochemistry and molecular biology. 2008;112:20-4.

12. Klinge CM, Riggs KA, Wickramasinghe NS, Emberts CG, McConda DB, Barry PN, et al. Estrogen receptor alpha 46 is reduced in tamoxifen resistant breast cancer cells and re-expression inhibits cell proliferation and estrogen receptor alpha 66-regulated target gene transcription. Molecular and cellular endocrinology. 2010;323:268-76.

13. Maaroufi Y, Lacroix M, Lespagnard L, Journe F, Larsimont D, Leclercq G. Estrogen receptor of primary breast cancers: evidence for intracellular proteolysis. Breast cancer research : BCR. 2000;2:444-54.

14. Flouriot G, Brand H, Denger S, Metivier R, Kos M, Reid G, et al. Identification of a new isoform of the human estrogen receptor-alpha (hER-alpha) that is encoded by distinct transcripts and that is able to repress hER-alpha activation function 1. Embo J. 2000;19:4688-700.

15. Barraille P, Chinestra P, Bayard F, Faye JC. Alternative initiation of translation accounts for a 67/45 kDa dimorphism of the human estrogen receptor ERalpha. BiochemBiophysRes Commun. 1999;257:84-8.

16. Penot G, Le Peron C, Merot Y, Grimaud-Fanouillere E, Ferriere F, Boujrad N, et al. The human estrogen receptor-alpha isoform hERalpha46 antagonizes the proliferative influence of hERalpha66 in MCF7 breast cancer cells. Endocrinology. 2005;146:5474-84.

17. Jiang HP, Teng RY, Wang Q, Zhang X, Wang HH, Cao J, et al. Estrogen receptor alpha variant ERalpha46 mediates growth inhibition and apoptosis of human HT-29 colon adenocarcinoma cells in the presence of 17beta-oestradiol. Chinesemedical journal. 2008;121:1025-31.

18. Billon-Gales A, Fontaine C, Filipe C, Douin-Echinard V, Fouque MJ, Flouriot G, et al. The transactivating function 1 of estrogen receptor alpha is dispensable for the vasculoprotective actions of 17beta-estradiol. Proceedings of the National Academy of Sciences of the United States of America. 2009;106:2053-8.

19. Abot A, Fontaine C, Raymond-Letron I, Flouriot G, Adlanmerini M, Buscato M, et al. The AF-1 activation function of estrogen receptor alpha is necessary and sufficient for uterine epithelial cell proliferation in vivo. Endocrinology. 2013;154:2222-33.

20. Jozan S, Julia AM, Carretie A, Eche N, Maisongrosse V, Fouet B, et al. 65 and 47 kDa forms of estrogen receptor in human breast cancer: relation with estrogen responsiveness. Breast cancer research and treatment. 1991;19:103-9.

21. Metivier R, Penot G, Carmouche RP, Hubner MR, Reid G, Denger S, et al. Transcriptional complexes engaged by apo-estrogen receptor-alpha isoforms have divergent outcomes. Embo J. 2004;23:3653-66.

22. Ainaoui N, Hantelys F, Renaud-Gabardos E, Bunel M, Lopez F, Pujol F, et al. Promoter-Dependent Translation Controlled by p54nrb and hnRNPM during Myoblast Differentiation. PloS one. 2015;10:e0136466.

23. Koppen A, Houtman R, Pijnenburg D, Jeninga EH, Ruijtenbeek R, Kalkhoven E. Nuclear receptor-coregulator interaction profiling identifies TRIP3 as a novel peroxisome proliferator-activated receptor gamma cofactor. Molecular & cellular proteomics : MCP. 2009;8:2212-26.

24. Houtman R, de Leeuw R, Rondaij M, Melchers D, Verwoerd D, Ruijtenbeek R, et al. Serine-305 phosphorylation modulates estrogen receptor alpha binding to a coregulator peptide array, with potential application in predicting responses to tamoxifen. Molecular cancer therapeutics. 2012;11:805-16.

25. al Saati T, Clamens S, Cohen-Knafo E, Faye JC, Prats H, Coindre JM, et al. Production of monoclonal antibodies to human estrogen-receptor protein (ER) using recombinant ER (RER). International journal of cancer Journal international du cancer. 1993;55:651-4.

26. Pavao M, Traish AM. Estrogen receptor antibodies: specificity and utility in detection, localization and analyses of estrogen receptor alpha and beta. Steroids. 2001;66:1-16.

27. Cheang MC, Treaba DO, Speers CH, Olivotto IA, Bajdik CD, Chia SK, et al. Immunohistochemical detection using the new rabbit monoclonal antibody SP1 of estrogen receptor in breast cancer is superior to mouse monoclonal antibody 1D5 in predicting survival. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2006;24:5637-44.

28. Bogina G, Zamboni G, Sapino A, Bortesi L, Marconi M, Lunardi G, et al. Comparison of anti-estrogen receptor antibodies SP1, 6F11, and 1D5 in breast cancer: lower

1D5 sensitivity but questionable clinical implications. American journal of clinical pathology. 2012;138:697-702.

29. Madeira KP, Daltoe RD, Sirtoli GM, Rezende LC, Carvalho AA, GuimaraesIdos S, et al. Comparison of immunohistochemical analysis with estrogen receptor SP1 and 1D5 monoclonal antibodies in breast cancer. Pathology, research and practice. 2012;208:657-61.

30. Faye MD, Holcik M. The role of IRES trans-acting factors in carcinogenesis. Biochimicaetbiophysicaacta. 2015;1849:887-97.

31. Leprivier G, Rotblat B, Khan D, Jan E, Sorensen PH. Stress-mediated translational control in cancer cells. Biochimicaetbiophysicaacta. 2015;1849:845-60.

32. Komar AA, Hatzoglou M. Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. Cell cycle. 2011;10:229-40.

33. Fox EM, Andrade J, Shupnik MA. Novel actions of estrogen to promote proliferation: integration of cytoplasmic and nuclear pathways. Steroids. 2009;74:622-7.

34. Jeng MH, Jiang SY, Jordan VC. Paradoxical regulation of estrogen-dependent growth factor gene expression in estrogen receptor (ER)-negative human breast cancer cells stably expressing ER. Cancer letters. 1994;82:123-8.

35. Merot Y, Metivier R, Penot G, Manu D, Saligaut C, Gannon F, et al. The relative contribution exerted by AF-1 and AF-2 transactivation functions in estrogen receptor alpha transcriptional activity depends upon the differentiation stage of the cell. The Journal of biological chemistry. 2004;279:26184-91.

36. Kato S, Endoh H, Masuhiro Y, Kitamoto T, Uchiyama S, Sasaki H, et al. Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. Science. 1995;270:1491-4.

37. Yi P, Wang Z, Feng Q, Pintilie GD, Foulds CE, Lanz RB, et al. Structure of a biologically active estrogen receptor-coactivator complex on DNA. Molecular cell. 2015;57:1047-58.

38. Horigome T, Ogata F, Golding TS, Korach KS. Estradiol-stimulated proteolytic cleavage of the estrogen receptor in mouse uterus. Endocrinology. 1988;123:2540-8.

39. Faye JC, Fargin A, Bayard F. Dissimilarities between the uterine estrogen receptor in cytosol of castrated and estradiol-treated rats. Endocrinology. 1986;118:2276-83.

40. Morfoisse F, Kuchnio A, Frainay C, Gomez-Brouchet A, Delisle MB, Marzi S, et al. Hypoxia induces VEGF-C expression in metastatic tumor cells via a HIF-1alpha-independent translation-mediated mechanism. Cell reports. 2014;6:155-67.

CLAIMS:

1. A method for assessing the status of breast cancer in a subject comprising i) quantifying in a tumor tissue sample obtained from the subject the expression level of an ERα46 isoform comprising an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the position 21 to position 442 in SEQ ID NO:1, ii) comparing the quantification value determined at step i) with a predetermined reference value and iii) detecting differences in the quantification value determined at step i) and the predetermined reference value is indicative of the status of breast cancer.

2. The method of claim 1 wherein the subject suffers from an estrogen-receptor positive breast cancer.

3. The method of claim 1 wherein the ERα46 isoform comprises:

   - an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 20 to the amino acid at position 442 in SEQ ID NO:1, or

   - an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 19 to the amino acid at position 442 in SEQ ID NO:1 or,

   - an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 18 to the amino acid at position 442 in SEQ ID NO:1 or,

   - an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 17 to the amino acid at position 442 in SEQ ID NO:1 or,

   - an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 16 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 15 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 14 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 13 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 12 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 11 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 10 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 9 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 8 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 7 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 6 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 5 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 4 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 3 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 2 to the amino acid at position 442 in SEQ ID NO:1 or,

- an amino acid sequence having at least 90% of identity with the amino acid sequence ranging from the amino acid at position 1 to the amino acid at position 442 in SEQ ID NO:1.

4. The method of claim 1 wherein the expression of the ERα46 isoform in the tumor tissue sample is determined by immunohistochemistry.

5. The method of claim 1 wherein the expression level of the ERα46 isoform in the tumor tissue sample is determined by quantifying the quantity of the mRNA encoding for the ERα46 isoform.

6. The method of claim 1 which further comprises determining the expression level of ERα66 isoform.

7. The method of claim 6 wherein the expression levels of ERα46 and ERα66 isoforms are determined in the tumor tissue sample, the ratio of the respective levels is determined and is compared to the predetermined reference value wherein detecting differences in

the determined ratio and the predetermined reference value is indicative of the status of breast cancer.

8. A kit suitable for performing the method of claim 6 which comprises a binding partner specific for ERα46 and a binding partner specific for ERα66.

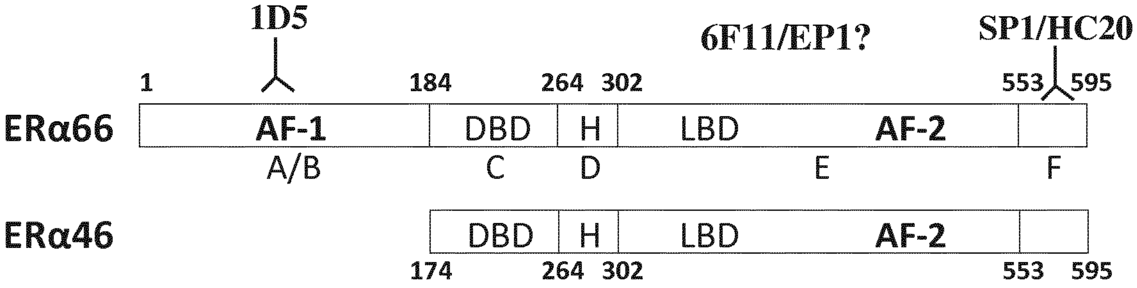


Figure 1A

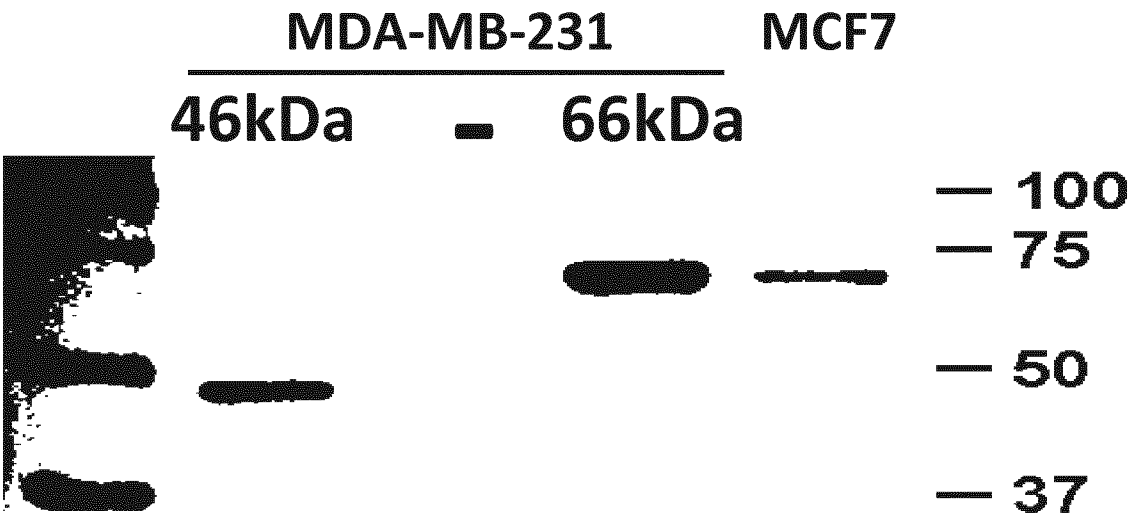


Figure 1B

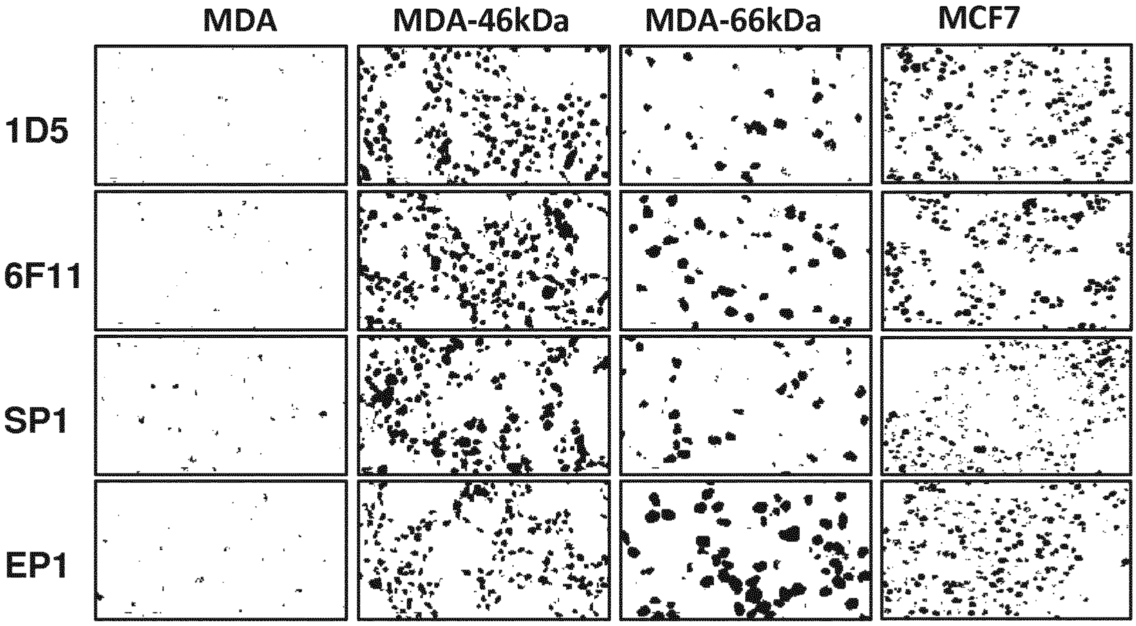


Figure 1C

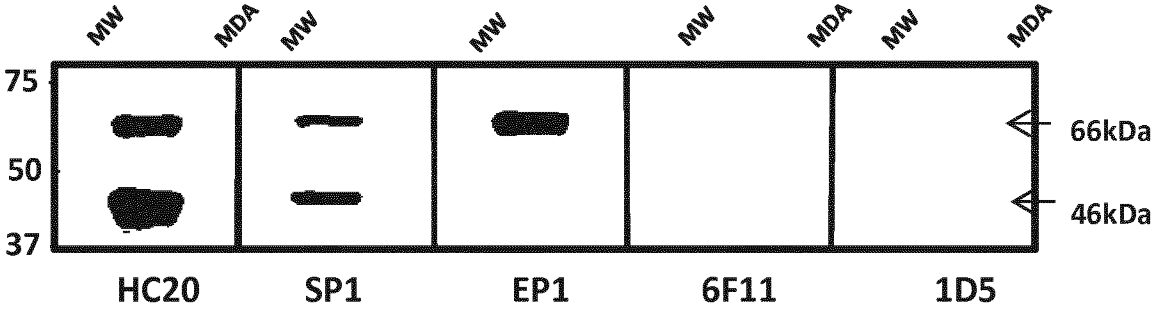
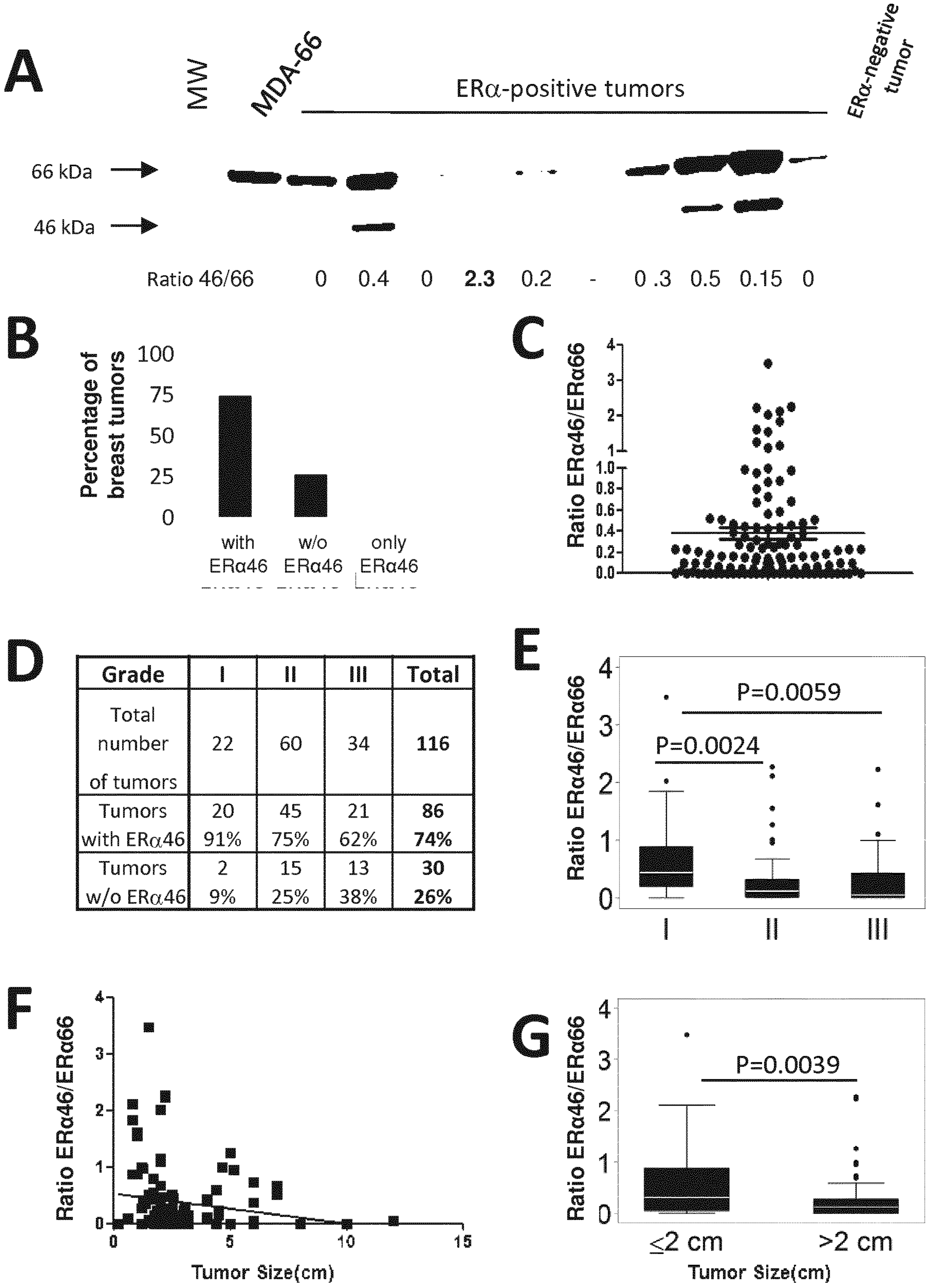


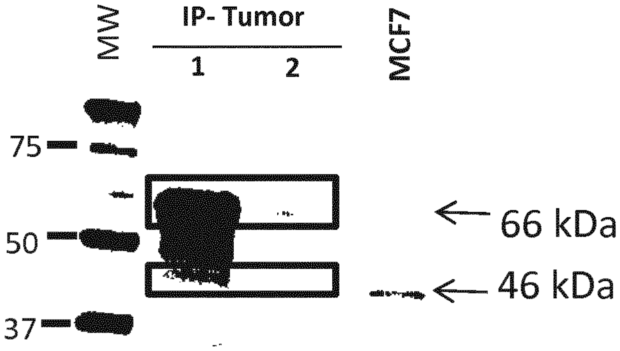
Figure 1D



Grade	I	II	III	Total
Total number of tumors	22	60	34	**116**
Tumors with ERα46	20 91%	45 75%	21 62%	**86** **74%**
Tumors w/o ERα46	2 9%	15 25%	13 38%	**30** **26%**

Figure 2

A



B

Tumor	Sample	Mascot Score	Cov. (%) / isoform	# Identified peptides	First N-ter peptide Identified and validated	Peptide identified before Met 174	Last C-ter peptide Identified and validated
MCF7	IP46	759	23.3	14	212 - 231	no	556-581
	IP66	440	24.4	12	009 - 032	yes	556-581
Tumor 1	IP46	96	8.0	5	402 - 412	no	556 - 581
	IP66	52	4.4	1	556 - 581	no	556 - 581
Tumor 2	IP 46	228	15.3	5	184-206	no	556-581
	IP 66	522	25.0	10	38-48	yes	556-581
Tumor 3	IP 46	65	3.0	1	450 - 467	no	450 - 467
Tumor 4	IP 46	35	1.8	1	402-412	no	402-412

Figure 3

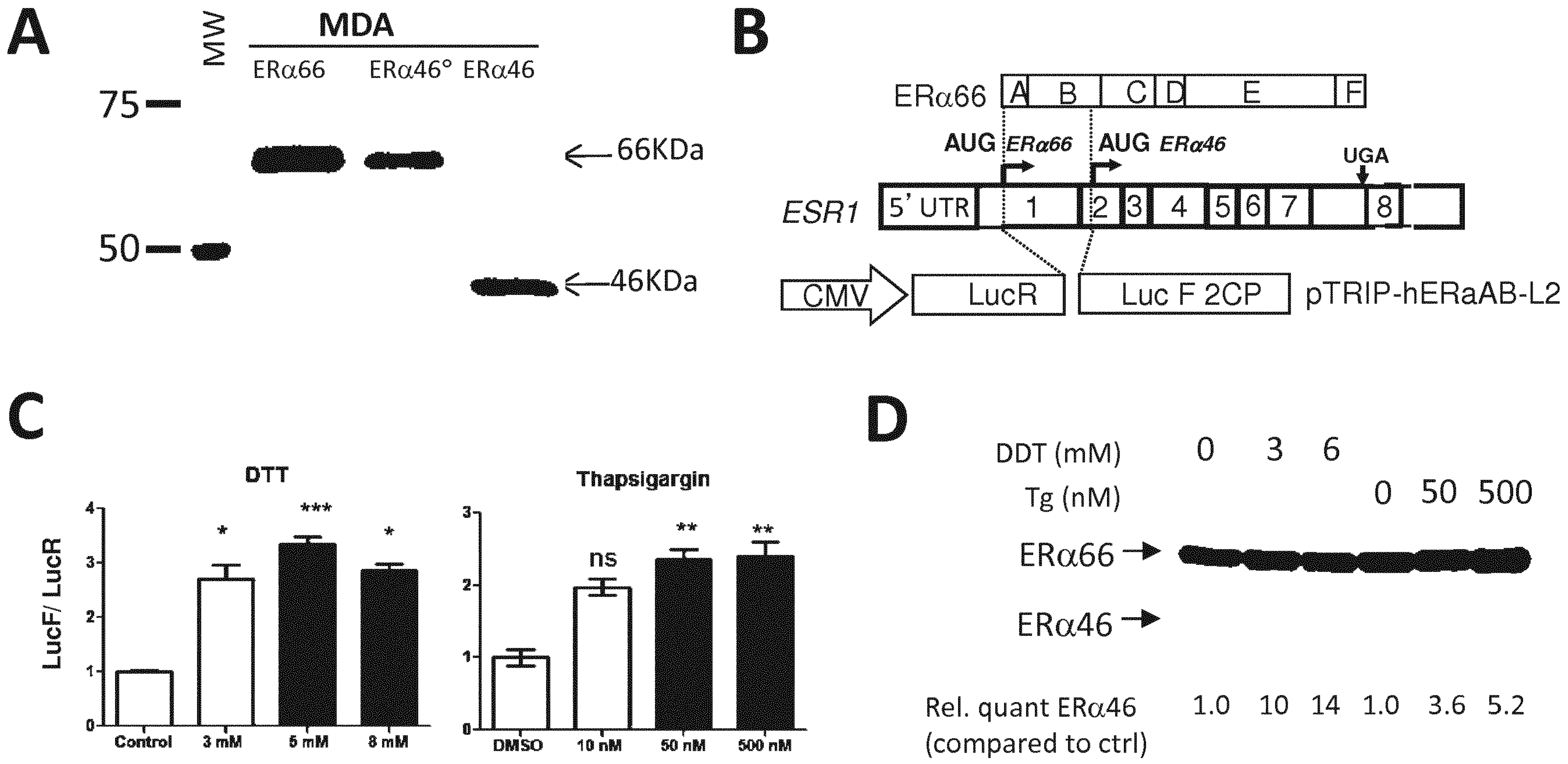

A
MW
MDA
ERα66
ERα46°
ERα46
75
50
66KDa
46KDa
B
ERα66
A
B
C
D
E
F
AUG ERα66
AUG ERα46
UGA
ESR1
5' UTR
1
2
3
4
5
6
7
8
CMV
LucR
Luc F 2CP
pTRIP-hERaAB-L2
C
DTT
Thapsigargin
LucF/ LucR
0
1
2
3
4
*
***
*
ns
**
**
Control
3 mM
5 mM
8 mM
DMSO
10 nM
50 nM
500 nM
D
DDT (mM)
0
3
6
Tg (nM)
0
50
500
ERα66
ERα46
Rel. quant ERα46 (compared to ctrl)
1.0
10
14
1.0
3.6
5.2

Figure 4

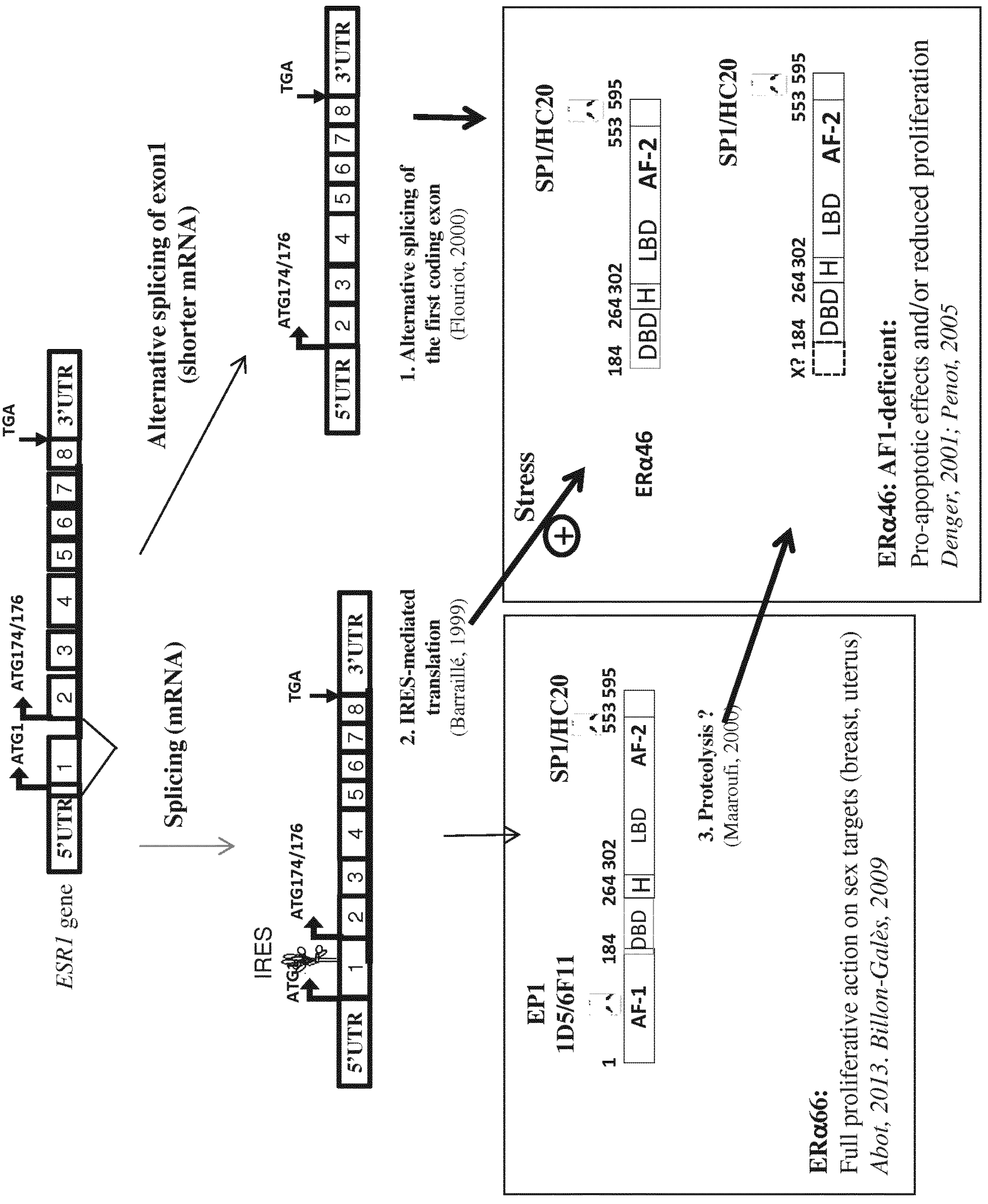

ESR1 gene
ATG1
ATG174/176
TGA
5'UTR
1
2
3
4
5
6
7
8
3'UTR
Splicing (mRNA)
Alternative splicing of exon1
(shorter mRNA)
IRES
ATG1
ATG174/176
TGA
5'UTR
1
2
3
4
5
6
7
8
3'UTR
ATG174/176
TGA
5'UTR
2
3
4
5
6
7
8
3'UTR
1. Alternative splicing of the first coding exon
(Flouriot, 2000)
2. IRES-mediated translation
(Barraillé, 1999)
3. Proteolysis ?
(Maaroufi, 2000)
Stress
+
ERα46
184
264 302
553 595
SP1/HC20
DBD
H
LBD
AF-2
X? 184
264 302
553 595
SP1/HC20
DBD
H
LBD
AF-2
ERα46: AF1-deficient:
Pro-apoptotic effects and/or reduced proliferation
Denger, 2001; Penot, 2005
EP1
1D5/6F11
1
184
264 302
553 595
SP1/HC20
AF-1
DBD
H
LBD
AF-2
ERα66:
Full proliferative action on sex targets (breast, uterus)
Abot, 2013, Billon-Galès, 2009

Figure 5

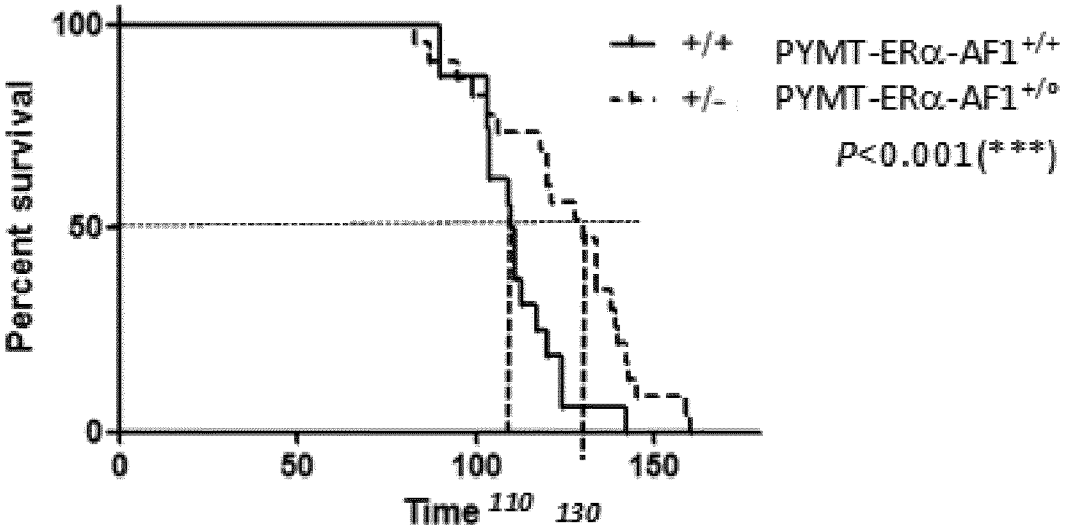

Percent survival
100
50
0
0
50
100
150
Time 110 130
+/+ PYMT-ERα-AF1+/+
+/- PYMT-ERα-AF1+/°
P<0.001(***)

Figure 6

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/067196

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01N33 G01N33/574
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, INSPEC, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KLINGE C M ET AL: "Estrogen receptor alpha 46 is reduced in tamoxifen resistant breast cancer cells and re-expression inhibits cell proliferation and estrogen receptor alpha 66-regulated target gene transcription", MOLECULAR AND CELLULAR ENDOCRINOLOGY, ELSEVIER IRELAND LTD, IE, vol. 323, no. 2, 29 July 2010 (2010-07-29), pages 268-276, XP027054423, ISSN: 0303-7207 [retrieved on 2010-03-17]	8
A	the whole document In particular: Title; Abstract; Materials and methods section. ----- -/--	1-7

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search 18 August 2017	Date of mailing of the international search report 28/08/2017
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer C.F. Angioni

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	US 2008/146658 A1 (LI JIN [US] ET AL) 19 June 2008 (2008-06-19) the whole document In particular: Figs. 1, 2, 7 and 8; Paragraphs 20, 21, 27, 28, 161, 165, 166 and 171-173. -----	8 1-7
A	YU GU ET AL: "The therapeutic target of estrogen receptor-alpha36 in estrogen-dependent tumors", JOURNAL OF TRANSLATIONAL MEDICINE, BIOMED CENTRAL, LONDON, GB, vol. 12, no. 1, 21 January 2014 (2014-01-21), page 16, XP021174590, ISSN: 1479-5876, DOI: 10.1186/1479-5876-12-16 the whole document In particular: Page 10 of 12, conclusion. -----	1-8

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2017/067196

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2008146658 A1	19-06-2008	CN 101528038 A	09-09-2009
		EP 2096916 A2	09-09-2009
		JP 5265556 B2	14-08-2013
		JP 2010508284 A	18-03-2010
		US 2008146658 A1	19-06-2008
		WO 2008052005 A2	02-05-2008