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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 December 2008 (24.12.2008)

PCT

(10) International Publication Number
WO 2008/155661 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/IB2008/002334

(22) International Filing Date: 16 April 2008 (16.04.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/923,690 16 April 2007 (16.04.2007) US

(71) Applicants *(for all designated States except US)*: **IPSOGEN** [FR/FR]; 163 avenue de Luminy, F-13009 Marseille (FR). **INSERM- INSTITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MEDICALE** [FR/FR]; 101 rue de Tolbiac, F-75654 Paris Cedex 13 (FR). **INSTITUT PAOLI-CALMETTES** [FR/FR]; 232 boulevard Sainte-Marguerite, F-13009 Marseille (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BERTUCCI, François** [FR/FR]; 25 avenue Georges Bizet, F-13470 Carnoux-en--Provence (FR). **BIRNBAUM, Daniel** [FR/FR]; La Croix-du-Sud, 9, rue Baldacchini, F-13009 Marseille (FR). **VIENS, Patrice** [FR/FR]; La Palmeraie Borely, Entrée A, 70, avenue d'Haifa, F-13008 Marseille (FR). **FERT, Vincent** [FR/FR]; 1 rue du Chalet, F-13008 Marseille (FR). **HERMITTE, Fabienne** [FR/FR]; Les Grands Pins, Avenue de la Gare, F-13600 La Ciotat (FR). **DEBONO, Stéphane** [FR/FR]; Résidence Le Clos des Lavandes, 53 Traverse du Tonkin, F-13010 Marseille (FR). **DERACO, Stéphane** [FR/FR]; 45 avenue Lespinet, F-31400 Toulouse (FR). **BORIE, Nathalie** [FR/FR]; 13 rue Croix Léonardoux, F-63000 Clermont-Ferrand (FR). **PIETTE, Fanny** [BE/BE]; 13 avenue Henri Dietrich, B-1200 Brussels (BE).

(74) Agent: **BREDEMA**; 38 avenue de l'Opéra, F-75002 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, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

(54) **Title:** METHODS OF ASSESSING A PROPENSITY OF CLINICAL OUTCOME FOR A FEMALE MAMMAL SUFFERING FROM BREAST CANCER

(57) **Abstract:** The present invention relates to a method of assessing a propensity of clinical outcome for a female mammal suffering from breast cancer in view of the expression of specific nucleic acids sequence in a biological sample.

WO 2008/155661 A2

METHODS OF ASSESSING A PROPENSITY OF CLINICAL OUTCOME FOR A FEMALE MAMMAL SUFFERING FROM BREAST CANCER

FIELD OF THE INVENTION

[0001] The present invention relates to methods of assessing a propensity of the clinical outcome of a female mammal suffering from breast cancer, preferably after said female mammal has been treated with chemotherapy, for example anthracycline-based chemotherapy.

BACKGROUND

[0002] Breast cancer is the most common nonskin malignancy in women and the second leading cause of female cancer mortality (FEAR *et al.*, *IEEE Potentials*, vol. 22 (1), p :12-18, 2003).

[0003] Worldwide, breast cancer is the most common cancer in women. It is estimated than in the year 2000, there were 350.000 new breast cancer cases in Europe, while the number of deaths from breast cancer was estimated at 130.000. Breast cancer is responsible for 26.5% of all new cancer cases among women in Europe, and 17.5% of cancer deaths. The highest incidence rates for the year 2000 were in Western Europe, with France in third position (42.000 new cases and 12.000 deaths). Despite these high rates of incidence and mortality, the survival of women diagnosed with breast cancer increased in Europe and in France since the end of the 1970s. This improvement is probably in relation with early diagnosis and screening programs and with adjuvant systemic therapy.

[0004] Adjuvant chemotherapy (CT) for breast cancer has undergone major changes over the past two decades. Results from the published update of the overview analysis by the Early Breast Cancer Trialists' Collaborative group indicated that administration of adjuvant CT significantly reduced the risk of recurrence by 23.5% and the risk of death by 15.3% . According to the same overview, the 10-year recurrence-free survival for node-positive patients

treated with adjuvant CT was 47.6% for patients younger than 50 years and 43.6% for those 50 to 69 years of age. The 10-year overall survival (OS) was 53.8% and 48.6% respectively. This overview analysis also demonstrated that, as compared with standard combination of cyclophosphamide, methotrexate and 5FU (CMF), regimens that contained anthracyclins reduced the annual risk of recurrence of breast cancer by 12% and the annual risk of death by 11%. Such regimens are significantly ($2p=0.0001$ for recurrence, $2p<0.00001$ for breast cancer mortality) more effective than CMF.

[0005] The most commonly used anthracycline-based adjuvant CT regimen in USA consists of four cycles of doxorubicin plus cyclophosphamide (AC) administrated every 21 days. Six cycles of FAC (cyclophosphamide, doxorubicin, and fluorouracil) every 3 weeks were also accepted as appropriate adjuvant regimen. Since epirubicin is less cardiotoxic than doxorubicin at an equimolar dose (recommended cumulative doses of doxorubicin and epirubicin are 550 mg/m^2 and 1.000 mg/m^2, respectively), several groups introduced epirubicin. A National Cancer Institute of Canada study showed that six cycles of cyclophosphamide, epirubicin, fluorouracil (CEF) were superior to six cycles of CMF . The *Groupe Français d'Etudes Adjuvantes* (GFEA; The French Adjuvant Trial Group) has studied epirubicin in the treatment of breast cancer for several years. The FEC regimen (fluorouracil, epirubicin, cyclophosphamide) has been evaluated in the trial setting lymph node-positive patients. Six cycles of adjuvant FEC 50 (epirubicin 50 mg/m^2) are better than 3 cycles. Subsequently a trial in patients less than 65 years of age, with node-positive operable breast cancer, compared FEC 50 *versus* FEC 100 (epirubicin 100 mg/m^2). Six cycles of FEC 100 was associated with improved relapse rates and better survival. Thus, 6 cycles of FEC every three weeks were generally accepted a few years ago in France as appropriate and "standard" adjuvant regimens for early breast cancer.

[0006] Recently, taxanes have emerged as potent agents for the adjuvant treatment of breast cancer. Studies involving more than 20.000 patients have been reported or are ongoing. Recent published adjuvant trials with taxanes

(paclitaxel, docetaxel) in node-positive breast cancer have demonstrated an additional benefit (as compared with regimen without taxanes), ranging from 2 to 7% in absolute difference in disease-free survival (DFS) or overall survival (OS) at 5 years. Two trials showed the benefit of incorporating sequentially 4 courses of paclitaxel after 4 cycles of AC: CALGB 9344 and NSABP B-28 . Two trials showed the benefit of incorporating docetaxel: BCIRG 01 study, which compared the FAC regimen (6 cycles) to the TAC regimen (docetaxel, doxorubicin, and fluorouracil, 6 cycles) , and PACS 01 study. The PACS 01 study (1.999 patients included) was promoted by the French Federation of Anti-Cancer Centers (FNCLCC). It compared the FEC 100 regimen (6 cycles) to a sequential regimen, 3 cycles of FEC100 followed by 3 cycles of docetaxel administered at the dose of 100 mg/m^2 every 3 weeks in node-positive patients. At a median follow-up of 60 months, adjuvant CT with 3 cycles of FEC100 followed by 3 cycles of docetaxel improved recurrence-free survival (reduction in the hazard rate of recurrence, 17%, p=0.04) and OS (reduction in the hazard rate of death, 23% p=0.005) (13). The 5-year DFS are 78.3% (3 FEC100 - 3 docetaxel arm) *vs* 73.2% (6 FEC100 arm) and the 5-year OS are 90.7 *vs* 86.7 respectively. In comparison with the BCIRG study, the incidence of febrile neutropenia, infection and cardiac dysfunction is very low especially in the sequential arm. As a consequence of these trials, the combination of anthracyclin and taxane has become the new standard of adjuvant CT for node-positive breast cancer. Several other trials promoted by the FNCLCC (PACS) investigated the optimal scheme of combination eprubicin-docetaxel: the PACS 04 study compared the FEC 100 regimen (6 cycles) to the combination epirubicin 75 mg/m^2 + docetaxel 75 mg/m^2 every 3 weeks in node-positive patients. Follow-up is ongoing with 3.015 patients included (end of inclusions in August 2004). The PACS 06 compared FEC 100 x 3 cycles every 2 weeks followed by docetaxel 100 mg/m^2 x 3 cycles every 2 weeks, in association with G-CSF, with either a 2-week or a 4-week interval between FEC and docetaxel. The primary endpoint was to define the rate of patients with any toxicity requiring dose reduction or treatment delay by more than one week over the 6 courses. As May 2005, the recruitment was stopped after 74 inclusions with the following conclusion, FEC 100 x 3 cycles every 2 weeks

followed by docetaxel 100 mg/m^2 x 3 cycles every 2 weeks, with a 2-week interval between FEC and docetaxel is not feasible due to an excess of skin / hand-foot syndrome severe toxicities.

[0007] Currently, adjuvant CT in early breast cancer is indicated according classical prognostic factors such the axillary lymph node status, the pathological size and grading of tumour, the hormonal receptor expression, and age of patients. These factors remain insufficient for reflecting the whole heterogeneity of disease, and none of them has been validated for selecting the optimal regimen of CT, resulting in the delivery of a combination of anthracyclin-taxane to all node-positive patients. However, recent studies have shown that in sub-groups of patients the addition of taxanes did not provide benefit as compared to FAC or FEC and that these classical regimens without taxanes might provide long survival in certain patients. Altogether with the potential toxicity and cost of the combination of anthracyclin-taxane, as well as the ongoing introduction/development of new drugs in adjuvant regimens (CT such as capecitabine, targeted therapy such as trastuzumab , hormone therapy such as anti-aromatases , diphosphonates), these data call for the identification of parameters predictive of clinical outcome (prognostic and/or predictive of response to CT) after given regimen of adjuvant CT.

[0008] A lot of research, mainly retrospective, has been performed to find predictive biological factors of adjuvant CT effectiveness, but, presently, there is still no individual admitted factor. The current prognostic factors evaluate only poorly the heterogeneous clinical behavior of disease. In consequence, many N- patients are subjected to unnecessary anthracycline-based adjuvant CT, and all N+ patients receive regimens based on anthracyclines and taxanes (Piccart et al. The Breast 14:439-445, 2005). However, taxanes are not yet universally accepted as standard treatment (Colozza et al. Oncologist 11:111-125, 2006). Recent randomized studies (Buzdar et al. Clin Cancer Res 8:1073-1079, 2002; Henderson et al. J Clin Oncol 21:976-983, 2003; Mamounas et al. J Clin Oncol 23:3686-3696, 2005; Martin et al. N Engl J Med 352:2302-2313, 2005; Roche et al. J Clin Oncol 24:5664-5671, 2006) have shown that the

addition of taxanes provides a significant but small benefit (3 to 7%) in 5-year survival. This suggests that a majority of patients do not benefit from the anthracycline-taxane combination. The availability of new drugs in adjuvant setting and the heterogeneity of breast cancer render necessary the tailoring of treatment without systematically associating all drugs. This challenge supposes to better assess the metastatic risk after CT. No biological factor predictive of anthracycline-based adjuvant CT efficacy (Hayes, The Breast 14:493-499, 2005) has yet been validated and introduced in routine use.

[0009] A predictive factor will be of a tremendous interest to select patients who benefit or who do not benefit from a specific regimen of adjuvant CT. Breast cancer is a complex genetic disease characterized by the accumulation of multiple molecular alterations. Pathological and clinical factors are insufficient to capture the complex cascade of events which drive the heterogeneous clinical behaviour of tumours.

[0010] High-throughput molecular technologies provide novel tools to tackle this complexity. In particular, DNA microarrays allow the simultaneous and quantitative analysis of the mRNA expression levels of thousands of genes in a single assay. The first research results are promising; comprehensive gene expression profiles of breast tumours are revealing new sub-groups of tumour in groups *a priori* identical, but with different outcome.

[0011] Several retrospective studies confirm the prognostic potential of DNA microarrays in breast cancer (Bertucci et al. Omics 10:429-443, 2006). Most studies focused on survival without any adjuvant systemic therapy (van de Vijver et al. N Engl J Med 347:1999-2009, 2002; van 't Veer et al. Nature 415:530-536, 2002; Wang et al. Lancet 365:671-679, 2005; Foekens et al. J Clin Oncol 24:1665-1671, 2006) after adjuvant HT (Ma et al. Cancer Cell 5:607-616, 2004; Paik et al. N Engl J Med 351:2817-2826, 2004; Oh et al. J Clin Oncol 24:1656-1664, 2006) and after neo-adjuvant CT (Sorlie et al. Proc Natl Acad Sci U S A 98:10869-10874, 2001; Sorlie et al. Proc Natl Acad Sci U S A 100:8418-8423, 2003). A few studies directly analyzed the response to primary CT (Ayers et al. J Clin Oncol 22:2284-2293, 2004; Bertucci et al.

Cancer Res 64:8558-8565, 2004; Chang et al. Lancet 362:362-369, 2003; Hannemann et al. J Clin Oncol. 23:3331-3342, 2005). Only few data with small (Bertucci al. Lancet 360:173-174; discussion 174, 2002; Bertucci et al. Hum Mol Genet 9:2981-2991, 2000) or heterogeneous series (Pawitan et al. Breast Cancer Res 7:R953-964, 2005) are available regarding outcome after adjuvant CT.. In all these studies, the prognostic and/or predictive multigenic signatures appeared more performing than individual molecular and pathoclinical parameters.

[0012] There is a need of adapting adjuvant CT in patients that are candidate to CT. The ongoing introduction of new drugs in adjuvant setting - in general associated to a low and heterogeneous benefit and a morbid and financial cost - necessitates refining the assessment of the metastatic risk after a given CT regimen and the decision regarding what CT regimen to use.

[0013] After exhausting testing we have identified gene marker sets that predict clinical outcome after CT, and methods of use thereof. This represents a step towards molecular tailoring by guiding patients towards the most beneficial CT regimen. This would allow moving away from the "one shoe fits all" strategy used in oncology for many years and from the ongoing therapeutic escalation.

SUMMARY OF THE INVENTION

[0014] The invention relates to a method for assessing the clinical outcome of a female mammal suffering from breast cancer, comprising the step of:

[0015] a) generating a metagene adjusted value underER by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 10 nucleic acid sequences selected in the group comprising or consisting of: SEQ ID No:374 (nm_000212), SEQ ID No:1027 (nm_007365), SEQ ID No:598 (nm_000636), SEQ ID No:717 (nm_024598), SEQ ID No:573 (nm_001527), SEQ ID No:83 (nm_015065), SEQ ID No:12 (nm_002964), SEQ ID No:405 (nm_000852), SEQ ID No:856 (nm_005564), SEQ ID No:384

(nm_002466), SEQ ID No:167 (nm_002627), SEQ ID No:51 (nm_198433), SEQ ID No:999 (nm_145290), SEQ ID No:979 (nm_004414), SEQ ID No:2 (nm_005245), SEQ ID No:98 (nm_016267), SEQ ID No:751 (nm_002423), SEQ ID No:696 (nm_001428), SEQ ID No:1050 (BC034638), SEQ ID No:488 (nm_002979), SEQ ID No:262 (nm_005194), SEQ ID No:1020 (nm_000359), SEQ ID No:1106 (BC015969), SEQ ID No:952 (nm_003878), SEQ ID No:675 (nm_001512), SEQ ID No:289 (nm_020179), SEQ ID No:553 (nm_004701), SEQ ID No:579 (nm_001814), SEQ ID No:760 (nm_005746), SEQ ID No:805 (nm_014624), SEQ ID No:361 (nm_002906), SEQ ID No:448 (nm_198569), SEQ ID No:170 (nm_002428), SEQ ID No:878 (nm_002774), SEQ ID No:1117, SEQ ID No:612 (nm_032515), SEQ ID No:540 (nm_003159), SEQ ID No:823 (nm_000100), SEQ ID No:131 (nm_145280), SEQ ID No:705 (nm_005596), SEQ ID No:31 (nm_005558), and SEQ ID No:199 (nm_024323), fragments, derivatives or complementary sequences thereof.

[0016] Preferably, at least 20 nucleic acid sequences selected in said group, and more preferably at least 25 nucleic acid sequences selected in said group.

[0017] In one embodiment, said metagene adjusted value underER is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 20 nucleic acid sequences selected in the group consisting of : SEQ ID No:374 (nm_000212); SEQ ID No:1027 (nm_007365); SEQ ID No:598 (nm_000636); SEQ ID No:573 (nm_001527); SEQ ID No:83 (nm_015065); SEQ ID No:12 (nm_002964); SEQ ID No:405 (nm_000852); SEQ ID No:856 (nm_005564); SEQ ID No:167 (nm_002627); SEQ ID No:51 (nm_198433); SEQ ID No:98 (nm_016267); SEQ ID No:751 (nm_002423); SEQ ID No:696 (nm_001428); SEQ ID No:262 (nm_005194); SEQ ID No:1020 (nm_000359); SEQ ID No:579 (nm_001814); SEQ ID No:760 (nm_005746); SEQ ID No:805 (nm_014624); SEQ ID No:878 (nm_002774); and SEQ ID No:612 (nm_032515), fragments, derivatives or complementary sequences thereof.

[0018] In another embodiment, said metagene adjusted value underER is generated by comparing the expression level, in a biological sample from said

female mammal and in a control, of the 27 nucleic acid sequences selected in the group consisting of : SEQ ID No:374 (nm_000212); SEQ ID No:1027 (nm_007365); SEQ ID No:598 (nm_000636); SEQ ID No:573 (nm_001527); SEQ ID No:83 (nm_015065); SEQ ID No:12 (nm_002964); SEQ ID No:405 (nm_000852); SEQ ID No:856 (nm_005564); SEQ ID No:167 (nm_002627); SEQ ID No:51 (nm_198433); SEQ ID No:98 (nm_016267); SEQ ID No:751 (nm_002423); SEQ ID No:696 (nm_001428); SEQ ID No:262 (nm_005194); SEQ ID No:1020 (nm_000359); SEQ ID No:579 (nm_001814); SEQ ID No:760 (nm_005746); SEQ ID No:805 (nm_014624); SEQ ID No:878 (nm_002774); SEQ ID No:612 (nm_032515); SEQ ID No:384 (nm_002466); SEQ ID No:2 (nm_005245); SEQ ID No:1050 (BC034638); SEQ ID No:952 (nm_003878); SEQ ID No:361 (nm_002906); SEQ ID No:31 (nm_005558); and SEQ ID No:199 (nm_024323), fragments, derivatives or complementary sequences thereof.

[0019] b) generating a metagene adjusted value underPR by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 6 nucleic acid sequences selected in the group comprising or consisting of : SEQ ID No:598 (nm_000636), SEQ ID No:1122, SEQ ID No:364 (nm_002253), SEQ ID No:387 (nm_006563), SEQ ID No:34 (nm_001229), SEQ ID No:657 (nm_000633), SEQ ID No:384 (nm_002466), SEQ ID No:451 (nm_001110), SEQ ID No:999 (nm_145290), SEQ ID No:1056 (AK126297), SEQ ID No:15 (nm_003243), SEQ ID No:1090 (AK125808), SEQ ID No:1120, SEQ ID No:12 (nm_002964), SEQ ID No:743 (nm_006875), SEQ ID No:414 (nm_000546), SEQ ID No:374 (nm_000212), SEQ ID No:711 (nm_002291), SEQ ID No:663 (nm_006928), SEQ ID No:1102 (AK124587), SEQ ID No:237 (nm_002644), SEQ ID No:60 (nm_022640), SEQ ID No:361 (nm_002906), SEQ ID No:119 (nm_004730) (or SEQ ID No:1109 (NM_002019)), SEQ ID No:167 (nm_002627), SEQ ID No:339 (nm_144970), SEQ ID No:333 (nm_145037), SEQ ID No:83 (nm_015065), SEQ ID No:330 (nm_018291), SEQ ID No:1024 (nm_030666), SEQ ID No:229 (nm_004586), SEQ ID No:925 (nm_005257), SEQ ID No:788 (nm_001005369), SEQ ID No:1104 (AK128524), SEQ ID No:1103 (BX108410), SEQ ID No:66

(nm_000416), SEQ ID No:1030 (nm_024007), SEQ ID No:1119, SEQ ID No:1068 (AK024670), SEQ ID No:241 (nm_000801), SEQ ID No:398 (nm_003084), SEQ ID No:74 (nm_000878), SEQ ID No:1087 (AK074131), SEQ ID No:955 (nm_001986), SEQ ID No:71 (nm_004633), SEQ ID No:1105 (BC072392), SEQ ID No:856 (nm_005564), SEQ ID No:231 (nm_006678), SEQ ID No:593 (nm_001511), SEQ ID No:384 (nm_002466), SEQ ID No:519 (nm_020125), SEQ ID No:579 (nm_001814), SEQ ID No:1039 (nm_006209), SEQ ID No:31 (nm_005558), SEQ ID No:327 (nm_173825), SEQ ID No:573 (nm_001527), SEQ ID No:98 (nm_016267), SEQ ID No:1059 (AK091113), SEQ ID No:886 (nm_000075), SEQ ID No:1032 (nm_005688), SEQ ID No:1091 (XM_378178), SEQ ID No:233 (nm_178155), SEQ ID No:938 (nm_003012), SEQ ID No:264 (nm_152862), SEQ ID No:546 (nm_005874), SEQ ID No:1099 (BC066343) SEQ ID No:1037 (nm_023068), SEQ ID No:550 (nm_004848), SEQ ID No:1027 (nm_007365), SEQ ID No:1005 (nm_014938), SEQ ID No:820 (nm_000593), and SEQ ID No:370 (nm_000106), fragments, derivatives or complementary sequences thereof.

[0020] Preferably, at least 10 nucleic acid sequences selected in said group, as an example at least 20 nucleic acid sequences or at least 30 nucleic acid sequences, and more preferably at least 36 nucleic acid sequences selected in said group.

[0021] In one embodiment, said metagene adjusted value underPR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 6 nucleic acid sequences selected in the group consisting of : SEQ ID No:364 (nm_002253); SEQ ID No:34 (nm_001229); SEQ ID No:657 (nm_000633); SEQ ID No:339 (nm_144970); SEQ ID No:229 (nm_004586); SEQ ID No:1119, fragments, derivatives or complementary sequences thereof.

[0022] In another embodiment, said metagene adjusted value underPR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 36 nucleic acid sequences selected in the group consisting of : SEQ ID No:364 (nm_002253); SEQ ID No:34

(nm_001229); SEQ ID No:657 (nm_000633); SEQ ID No:339 (nm_144970); SEQ ID No:229 (nm_004586); SEQ ID No:1119; SEQ ID No:387 (nm_006563); SEQ ID No:1056 (AK126297); SEQ ID No:15 (nm_003243); SEQ ID No:1120; SEQ ID No:414 (nm_000546); SEQ ID No:374 (nm_000212); SEQ ID No:711 (nm_002291); SEQ ID No:663 (nm_006928); SEQ ID No:237 (nm_002644); SEQ ID No:60 (nm_022640); SEQ ID No:119 (nm_004730) ; SEQ ID No:330 (nm_018291); SEQ ID No:1024 (nm_030666); SEQ ID No:925 (nm_005257); SEQ ID No:1104 (AK128524); SEQ ID No:1103 (BX108410); SEQ ID No:66 (nm_000416); SEQ ID No:1068 (AK024670); SEQ ID No:374 (nm_000212); SEQ ID No:74 (nm_000878); SEQ ID No:231 (nm_006678); SEQ ID No:593 (nm_001511); SEQ ID No:384 (nm_002466); SEQ ID No:1039 (nm_006209); SEQ ID No:327 (nm_173825); SEQ ID No:886 (nm_000075); SEQ ID No:1032 (nm_005688); SEQ ID No:264 (nm_152862); SEQ ID No:1037 (nm_023068); and SEQ ID No:1005 (nm_014938), fragments, derivatives or complementary sequences thereof.

[0023] c) generating a metagene adjusted value underEGFR by comparing the level, in a biological sample from said female mammal and in a control, of at least 10 nucleic acid sequences selected in the group comprising or consisting of : SEQ ID No:1071 (NM_001033047), SEQ ID No:254 (nm_005581), SEQ ID No:6 (nm_003225), SEQ ID No:883 (nm_000125), SEQ ID No:543 (nm_005080), SEQ ID No:681 (nm_020974), SEQ ID No:63 (nm_001002295), SEQ ID No:212 (nm_024852), SEQ ID No:635 (nm_001002029), SEQ ID No:535 (nm_003226), SEQ ID No:1125, SEQ ID No:109 (nm_000662), SEQ ID No:342 (nm_001846), SEQ ID No:927 (nm_004703), SEQ ID No:1124, SEQ ID No:124 (nm_014899), SEQ ID No:280 (nm_020764) (or SEQ ID No:1110 (nm_024522)), SEQ ID No:297 (nm_016463), SEQ ID No:791 (nm_016835), SEQ ID No:210 (nm_178840), SEQ ID No:827 (nm_152499), SEQ ID No:1064 (nm_000767), SEQ ID No:147 (nm_014675), SEQ ID No:323 (nm_001014443), SEQ ID No:106 (nm_004619), SEQ ID No:181 (nm_000848), SEQ ID No:376 (nm_057158), SEQ ID No:116 (nm_014034), SEQ ID No:252 (nm_000758), SEQ ID No:797 (nm_022131), SEQ ID No:911 (nm_000168), SEQ ID No:720 (nm_004726),

SEQ ID No:889 (nm_000561), SEQ ID No:250 (nm_000930), SEQ ID No:179 (nm_004747), SEQ ID No:786 (nm_033388), SEQ ID No:177 (nm_015996), SEQ ID No:1047 (BC012900), SEQ ID No:301 (nm_004326), SEQ ID No:207 (nm_003940), SEQ ID No:936 (nm_003462), SEQ ID No:916 (nm_001453) (or SEQ ID No:1116 (nm_004040)), SEQ ID No:1052 (BX096026), SEQ ID No:159 (nm_000224), SEQ ID No:1096 (AK127274), SEQ ID No:28 (nm_021800), SEQ ID No:1054 (AK123264), SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (nm_053279)), SEQ ID No:825 (nm_024704), SEQ ID No:145 (nm_017786), SEQ ID No:491 (nm_004374), SEQ ID No:485 (nm_003834), SEQ ID No:1072 (AY007114), SEQ ID No:274 (nm_032108), SEQ ID No:258 (nm_080545), SEQ ID No:292 (nm_014371), SEQ ID No:803 (nm_183047), SEQ ID No:349 (nm_031946), SEQ ID No:1123, SEQ ID No:763 (nm_014585), SEQ ID No:438 (nm_001759), SEQ ID No:94 (nm_014315), SEQ ID No:845 (nm_001089), SEQ ID No:1084 (BX648964), SEQ ID No:734 (nm_025137), SEQ ID No:943 (nm_002141), SEQ ID No:1085 (nm_000720), and SEQ ID No:276 (nm_012202), fragments, derivatives or complementary sequences thereof.

[0024] Preferably, at least 20 nucleic acid sequences selected in said group, as an example at least 24 nucleic acid sequences or at least 30 nucleic acid sequences, and more preferably at least 37 nucleic acid sequences selected in said group.

[0025] In one embodiment, said metagene adjusted value underEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 24 nucleic acid sequences selected in the group consisting of : SEQ ID No:1071 (nm_001033047); SEQ ID No:254 (nm_005581); SEQ ID No:6 (nm_003225); SEQ ID No:883 (nm_000125); SEQ ID No:543 (nm_005080); SEQ ID No:681 (nm_020974); SEQ ID No:63 (nm_001002295); SEQ ID No:212 (nm_024852); SEQ ID No:635 (nm_001002029); SEQ ID No:535 (nm_003226); SEQ ID No:1125); SEQ ID No:1124; SEQ ID No:297 (nm_016463); SEQ ID No:791 (nm_016835); SEQ ID No:827 (nm_152499); SEQ ID No:207 (nm_003940); SEQ ID No:916

(nm_001453) (or SEQ ID No:1116 (nm_004040)); SEQ ID No:1052 (BX096026) ; SEQ ID No:159 (nm_000224); SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (nm_053279)); SEQ ID No:845 (nm_001089); and SEQ ID No:1085 (nm_000720), fragments, derivatives or complementary sequences thereof.

[0026] In another embodiment, said metagene adjusted value underEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 37 nucleic acid sequences selected in the group consisting of : SEQ ID No:1071 (nm_001033047); SEQ ID No:254 (nm_005581); SEQ ID No:6 (nm_003225); SEQ ID No:883 (nm_000125); SEQ ID No:543 (nm_005080); SEQ ID No:681 (nm_020974); SEQ ID No:63 (nm_001002295); SEQ ID No:212 (nm_024852); SEQ ID No:635 (nm_001002029); SEQ ID No:535 (nm_003226); SEQ ID No:1125; SEQ ID No:1124; SEQ ID No:297 (nm_016463); SEQ ID No:791 (nm_016835); SEQ ID No:827 (nm_152499); SEQ ID No:207 (nm_003940); SEQ ID No:916 (nm_001453) (or SEQ ID No:1116 (nm_004040)); SEQ ID No:1052 (BX096026) ; SEQ ID No:159 (nm_000224); SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (NM_053279)); SEQ ID No:845 (nm_001089); SEQ ID No:1085 (NM_000720) ; SEQ ID No:109 (nm_000662); SEQ ID No:342 (nm_001846); SEQ ID No:927 (nm_004703); SEQ ID No:280 (nm_020764) (or SEQ ID No:1110 (NM_024522)); SEQ ID No:210 (nm_178840); SEQ ID No:181 (nm_000848); SEQ ID No:116 (nm_014034); SEQ ID No:250 (nm_000930); SEQ ID No:177 (nm_015996); SEQ ID No:825 (nm_024704); SEQ ID No:145 (nm_017786); and SEQ ID No:276 (nm_012202), fragments, derivatives or complementary sequences thereof.

[0027] d) generating a score (S_C) from said metagene adjusted values using a mathematical method establishing a relation between the combined metagene values and the clinical outcome of said female mammal.

[0028] In one embodiment, the mathematical method used in step d) comprises a Cox regression analysis (Wright et al., Proc. Natl. Acad. Sci. USA,

vol.100 (17), p. 9991-9996, 2003) or a CART analysis (Breiman et al Classification and Regression Trees, Chapman & Hall 1984).

[0029] In a particular embodiment, the mathematical method is a Cox regression analysis and the score (S_C) is generated according to the following formula: S_C = a x underER + b x underPR + c x under EGFR, wherein "a" is comprised in the interval [-6,26 ; +0,49] , "b" is comprised in the interval [-2,65; +0,29] and "c" is comprised in the interval [-6,69; + 1,65].

[0030] For example the formula is: S_C = - 2.90279 x underER – 1.47423 x underPR – 4.17198 x under EGFR.

[0031] The invention further relates to a method for assessing the clinical outcome of a female mammal suffering from breast cancer, comprising the step of:

[0032] a) generating a metagene adjusted value underEGFR by comparing the expression level, in a biological sample from said female mammal and in a control, of at least one nucleic acid sequence selected in the group consisting of: SEQ ID No:1071 (NM_001033047), SEQ ID No:254 (nm_005581), SEQ ID No:6 (nm_003225), SEQ ID No:883 (nm_000125), SEQ ID No:543 (nm_005080), SEQ ID No:681 (nm_020974), SEQ ID No:63 (nm_001002295), SEQ ID No:212 (nm_024852), SEQ ID No:635 (nm_001002029), SEQ ID No:535 (nm_003226), SEQ ID No:1125, SEQ ID No:109 (nm_000662), SEQ ID No:342 (nm_001846), SEQ ID No:927 (nm_004703), SEQ ID No:1124, SEQ ID No:124 (nm_014899), SEQ ID No:280 (nm_020764) (or SEQ ID No:1110 (nm_024522)), SEQ ID No:297 (nm_016463), SEQ ID No:791 (nm_016835), SEQ ID No:210 (nm_178840), SEQ ID No:827 (nm_152499), SEQ ID No:1064 (NM_000767), SEQ ID No:147 (nm_014675), SEQ ID No:323 (nm_001014443), SEQ ID No:106 (nm_004619), SEQ ID No:181 (nm_000848), SEQ ID No:376 (nm_057158), SEQ ID No:116 (nm_014034), SEQ ID No:252 (nm_000758), SEQ ID No:797 (nm_022131), SEQ ID No:911 (nm_000168), SEQ ID No:720 (nm_004726), SEQ ID No:889 (nm_000561), SEQ ID No:250 (nm_000930), SEQ ID No:179 (nm_004747), SEQ ID No:786

(nm_033388), SEQ ID No:177 (nm_015996), SEQ ID No:1047 (BC012900), SEQ ID No:301 (nm_004326), SEQ ID No:207 (nm_003940), SEQ ID No:936 (nm_003462), SEQ ID No:916 (nm_001453) (or SEQ ID No:1116 (NM_004040)), SEQ ID No:1052 (BX096026), SEQ ID No:159 (nm_000224), SEQ ID No:1096 (AK127274), SEQ ID No:28 (nm_021800), SEQ ID No:1054 (AK123264), SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (nm_053279)), SEQ ID No:825 (nm_024704), SEQ ID No:145 (nm_017786), SEQ ID No:491 (nm_004374), SEQ ID No:485 (nm_003834), SEQ ID No:1072 (AY007114), SEQ ID No:274 (nm_032108), SEQ ID No:258 (nm_080545), SEQ ID No:292 (nm_014371), SEQ ID No:803 (nm_183047), SEQ ID No:349 (nm_031946), SEQ ID No:1123, SEQ ID No:763 (nm_014585), SEQ ID No:438 (nm_001759), SEQ ID No:94 (nm_014315), SEQ ID No:845 (nm_001089), SEQ ID No:1084 (BX648964), SEQ ID No:734 (nm_025137), SEQ ID No:943 (nm_002141), SEQ ID No:1085 (nm_000720), and SEQ ID No:276 (nm_012202), fragments, derivatives or complementary sequences thereof.

[0033] Preferably, said nucleic acid sequence is SEQ ID No:681 (nm_020974), fragments, derivatives or complementary sequences thereof.

[0034] Preferably, at least 10 nucleic acid sequences selected in said group, as an example at least 20 nucleic acid sequences or at least 24 nucleic acid sequences, and more preferably at least 37 nucleic acid sequences selected in said group.

[0035] In one embodiment, said metagene adjusted value underEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 24 nucleic acid sequences selected in the group consisting of : SEQ ID No:1071 (nm_001033047); SEQ ID No:254 (nm_005581); SEQ ID No:6 (nm_003225); SEQ ID No:883 (nm_000125); SEQ ID No:543 (nm_005080); SEQ ID No:681 (nm_020974); SEQ ID No:63 (nm_001002295); SEQ ID No:212 (nm_024852); SEQ ID No:635 (nm_001002029); SEQ ID No:535 (nm_003226); SEQ ID No:1125); SEQ ID No:1124; SEQ ID No:297 (nm_016463); SEQ ID No:791 (nm_016835); SEQ ID No:827 (nm_152499); SEQ ID No:207 (nm_003940); SEQ ID No:916

(nm_001453) (or SEQ ID No:1116 (nm_004040)); SEQ ID No:1052 (BX096026) ; SEQ ID No:159 (nm_000224); SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (NM_053279)); SEQ ID No:845 (nm_001089); and SEQ ID No:1085 (NM_000720), fragments, derivatives or complementary sequences thereof.

[0036] In another embodiment, said metagene adjusted value underEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 37 nucleic acid sequences selected in the group consisting of : SEQ ID No:1071 (nm_001033047); SEQ ID No:254 (nm_005581); SEQ ID No:6 (nm_003225); SEQ ID No:883 (nm_000125); SEQ ID No:543 (nm_005080); SEQ ID No:681 (nm_020974); SEQ ID No:63 (nm_001002295); SEQ ID No:212 (nm_024852); SEQ ID No:635 (nm_001002029); SEQ ID No:535 (nm_003226); SEQ ID No:1125; SEQ ID No:1124; SEQ ID No:297 (nm_016463); SEQ ID No:791 (nm_016835); SEQ ID No:827 (nm_152499); SEQ ID No:207 (nm_003940); SEQ ID No:916 (nm_001453) (or SEQ ID No:1116 (nm_004040)); SEQ ID No:1052 (BX096026) ; SEQ ID No:159 (nm_000224); SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (NM_053279)); SEQ ID No:845 (nm_001089); SEQ ID No:1085 (NM_000720) ; SEQ ID No:109 (nm_000662); SEQ ID No:342 (nm_001846); SEQ ID No:927 (nm_004703); SEQ ID No:280 (nm_020764) (or SEQ ID No:1110 (NM_024522)); SEQ ID No:210 (nm_178840); SEQ ID No:181 (nm_000848); SEQ ID No:116 (nm_014034); SEQ ID No:250 (nm_000930); SEQ ID No:177 (nm_015996); SEQ ID No:825 (nm_024704); SEQ ID No:145 (nm_017786); and SEQ ID No:276 (nm_012202), fragments, derivatives or complementary sequences thereof.

[0037] b) generating a metagene adjusted value overEGFR by comparing the expression level, in a biological sample from said female mammal and in a control, of at least one nucleic acid sequences selected in the group consisting of SEQ ID No:405 (nm_000852), SEQ ID No:374 (nm_000212), SEQ ID No:1122, SEQ ID No:598 (nm_000636), SEQ ID No:262 (nm_005194), SEQ ID No:1099 (BC066343), SEQ ID No:696 (nm_001428), SEQ ID No:1059

(AK091113), SEQ ID No:751 (nm_002423), SEQ ID No:1121, SEQ ID No:286 (nm_002417), SEQ ID No:244 (nm_199002), SEQ ID No:18 (nm_001880), SEQ ID No:121 (nm_014553), SEQ ID No:1107 (BC073775), SEQ ID No:103 (nm_003619), SEQ ID No:1118, SEQ ID No:42 (nm_000757), and SEQ ID No:1067 (AK123784), fragments, derivatives or complementary sequences thereof.

[0038] Preferably, said nucleic acid sequence is SEQ ID No: 1107 (BC073775) or SEQ ID No: 1099 (BC066343), fragments, derivatives or complementary sequences thereof.

[0039] More preferably, at least 5 nucleic acid sequences selected in said group, as an example at least 10 nucleic acid sequences, and more preferably at least 12 nucleic acid sequences selected in said group.

[0040] In one embodiment, said metagene adjusted value overEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 5 nucleic acid sequences selected in the group consisting of: SEQ ID No:1122; SEQ ID No:598 (nm_000636); SEQ ID No:696 (nm_001428); SEQ ID No:1059 (AK091113); and SEQ ID No:121 (nm_014553), fragments, derivatives or complementary sequences thereof.

[0041] In another embodiment, said metagene adjusted value overEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 12 nucleic acid sequences selected in the group consisting of : SEQ ID No:1122; SEQ ID No:598 (nm_000636); SEQ ID No:696 (nm_001428); SEQ ID No:1059 (AK091113); SEQ ID No:121 (nm_014553); SEQ ID No:262 (nm_005194); SEQ ID No:1099 (BC066343); SEQ ID No:751 (nm_002423); SEQ ID No:1121; SEQ ID No:286 (nm_002417); SEQ ID No:103 (nm_003619); and SEQ ID No:1118, fragments, derivatives or complementary sequences thereof.

[0042] c) generating a score (S_C) from said metagene adjusted values using a mathematical method establishing a relation between the combined metagene values and the clinical outcome of said female mammal.

[0043] In one embodiment, the mathematical method used in step c) comprises a Cox regression analysis or a CART analysis.

[0044] In another embodiment, the mathematical method is a Cox regression and the score (S_C) to the following formula: S_C = a x overEGFR + b x underEGFR, wherein "a" is comprised in the interval [-1,85; +0,81] and "b" is comprised in the interval [-3,86; +0,70]

[0045] For example the formula is: S_C = – 1.33 x over EGFR – 2.28 x under EGFR.

[0046] The invention further relates to a method of assessing the clinical outcome of a female mammal suffering from breast cancer, comprising the steps of:

[0047] a) generating a metagene adjusted value underER by comparing the expression level, in a biological sample from said female mammal and in a control, of at least two genes, e.g. by using nucleic acid sequences selected in the group of Affymetrix® Probe Sets, of table IX or XII, preferably table XII (described below),

[0048] b) generating said metagene adjusted value underPR by comparing the expression level, in a biological sample from said female mammal and in a control, of at least two genes, e.g. by using nucleic acid sequences selected in the group of Affymetrix® Probe Sets, of table X or XIII, preferably table XIII (described below),

[0049] c) generating said metagene adjusted value underEGFR by comparing the expression level, in a biological sample from said female mammal and in a control, of at least two genes, e.g. by using the nucleic acid sequences selected in the group of Affymetrix® Probe Sets, of table XI or XIV, preferably table XIV (described below),

[0050] d) generating a score (S_C) from said metagene adjusted values using a mathematical method establishing a relation between the combined metagene values and the clinical outcome of said female mammal.

[0051] In one embodiment, the mathematical method used in step d) comprises a Cox regression or CART analysis.

[0052] In another embodiment, the mathematical method used in step d) is a Cox regression and the score (S_C) is generated according to the following formula: S_C = a x underER + b x underPR + c x under EGFR, wherein "a" is comprised in the interval [-6,26 ; +0,49] , "b" is comprised in the interval [-2,65; +0,29] and "c" is comprised in the interval [-6,69; + 1,65].

[0053] For example, the formula is : S_C = - 2.90279 x underER – 1.47423 x underPR – 4.17198 x under EGFR.

[0054] Preferably, the comparing of expression level at each step a), b) and c) is performed with at least 5, preferably 10, preferably all of said genes or nucleic acid sequences of each respective group.

[0055] In various embodiments, said methods may comprise the first step of quantifying in a biological sample from said female mammal the expression level of said nucleic acids sequences.

[0056] In other various embodiments, these methods can comprise the step e) of comparing said score (S_C) from the biological sample with a baseline or a score (S_C) from a control sample.

[0057] In other various embodiments, said biological sample is a breast tumor sample. By "sample" is meant a cell or a tissue.

[0058] In other various embodiments, said methods further comprise a step of taking at least one biological sample from said female mammal.

[0059] In another embodiment, said methods comprise a step of administrating a pharmaceutical treatment, preferably a chemotherapy treatment to a female mammal, for optimizing the clinical outcome of said female mammal in response to said treatment. The pharmaceutical treatment may comprise the use of one or more taxane compounds, e.g., docetaxel or paclitaxel. This treatment may be administered if the female mammal has not

responded to a previous anti-cancer treatment, e.g., a treatment comprising the use of one or more anthracyclin compound, e.g., epirubicin, doxorubicin, pirarubicin, idarubicin, zorubicin or aclarubicin, preferably epirubicin.

[0060] In a further aspect, the methods according to the invention may be used for identifying a female mammal that has not responded to a previous anti-cancer treatment, e.g., a treatment comprising the use of one or more anthracyclin compound, e.g., epirubicin, doxorubicin, pirarubicin, idarubicin, zorubicin or aclarubicin, preferably epirubicin.

[0061] In other various embodiments, a comparison of or analysis of data may involve a statistical computer mediated analysis. Also, said methods may optionally further involve generating a printed report.

[0062] The invention further relates to a computer program comprising instructions for performing said methods.

[0063] Finally, the invention relates to a recording medium for recording said computer program.

DETAILED DESCRIPTION

[0064] Unless otherwise noted, technical terms are used according to conventional usage.

[0065] In order to facilitate review of the various embodiment of the invention, the following explanation of specific terms is provided:

[0066] Mammals corresponds to animals such as humans, mice, rats, guinea pigs, monkeys, cats, dogs, pigs, horses, or cows, preferably to humans, and most preferably to women;

[0067] Biological sample: any biological material, such as a cell, a tissue sample, or a biopsy from breast cancer.

[0068] A "Metagene" as used herein corresponds to a group of genes for which expression variation (but not necessarily expression level) across tumors is correlated. A metagene can be simply calculated by one of skill in the art according to the method as described in the examples.

[0069] A "Control" as used herein corresponds to one or more biological samples from a cell, a tissue sample or a biopsy from breast. Said control may be obtained from the same female mammal than the one to be tested or from another female mammal, preferably from the same specie, or from a population of females mammal, preferably from the same specie, that may be the same or different from the test female mammal or subject. Said control may correspond to a biological sample from a cell, a cell line, a tissue sample or a biopsy from breast cancer. Preferably, the expression of EGFR, RE, PR and/or KI-67 has been established for this biological sample, by IHC (ImmunoHistoChemistry) FISH (Fluorescence In Situ Hybridization) or Quantitative PCR, for example.

[0070] In silico research: Literally referring to "in computer" systems, in silico research involves methods to test biological models, drugs, and other interventions using computer models rather than laboratory (*in vitro*) and animal (*in vivo*) experiments. In silico methods can involve analyzing an existing database, for instance a database that includes one or more records that include quantitative analysis of nucleic acid sequence expression. Analysis of such databases may include mining, parsing, selecting, identifying, sorting, or filtering of the data in the database. Data in the database can also be subjected to a clustering algorithm, discrimination algorithm, difference test, correlation, regression algorithm or other statistical modeling algorithm.

[0071] Using in silico research, drug treatment can be selected, tested and validated, and experimental strategies can be assessed. In silico systems complement laboratory-based research, yet increase productivity and efficiency by minimizing the need for in vitro and in vivo laboratory experiments.

[0072] In certain embodiments provided herein, in silico systems are used. In particular, this disclosure provides in silico methods for assessing a condition related to the clinical outcome of a female mammal suffering from breast cancer. Such methods involve assessing data in a database. The data in the database usually includes a quantity of nucleic acids from a biological sample from one or more individuals.

[0073] Quantitative data as discussed herein include molar quantitative data or relative data (variation of expression compared to control) for individual nucleic acid sequences, or subsets of nucleic acid sequences. Quantitative aspects of nucleic acids samples may be provided and/or improved by including one or more quantitative internal standards during the analysis, for instance one control nucleic acid sequence. Internal standards described herein enable true quantification of each nucleic acid sequence expression.

[0074] Truly quantitative data can be integrated from multiple sources (whether it is work from different labs, samples from different subjects, or merely samples processed on different days) into a single seamless database, regardless of the number of nucleic acid sequences measured in each discrete, individual analysis.

[0075] In any of the provided methods, a comparison of or an analysis involves a statistical or computer-mediated analysis.

[0076] The mathematical model (or method) for establishing a relation between the combined metagene adjusted values is realized on a population of mammal females showing the same ethnic and the same breast cancer characteristics than the female mammal to be tested.

[0077] The metagene coefficients (a, b, c) in the formulas used to calculate the scores (S_C) may vary according to the used tumor samples database consisting of mammal females showing the same ethnic and the same characteristics. A skilled person may calculate these coefficients by using a so-called Cox regression as described in Wright et al. (Proc. Natl. Acad. Sci. USA, vol.100 (17), p. 9991-9996, 2003)

[0078] Optionally, in some of the provided embodiments, the methods further involve comparing the score (S_C) from the female mammal to the score (S_C) from another female mammal, preferably from the same specie, or a compiled score (S_C) from a population of females mammal, preferably from the same specie, that may be the same or different from the test female mammal or subject.

[0079] In specific examples of such methods, the control is a baseline corresponding to a score (S_C) established from a population of females mammal.

[0080] The baseline is simply determined by one of skill in the art in view of the protocol described in the examples. An optimal baseline is obtained by using score distribution separating tumors into two groups of most significant different outcome.

[0081] As an example (described below), the inventors have established that a woman having a score (S_C) of more than 0.136 have at least a double propensity of poor clinical outcome than a woman with a score (S_C) of less than 0.0393.

[0082] Any of the provided method can further involve generating a printed report, for instance a report of some or all the data, of some or all the conclusions drawn from the data, or of a score or comparison between the results of a subject or individual and other individuals or a control or baseline.

[0083] There are many ways to collect quantitative or relative data on nucleic acids sequences, and the analytical methodology does not affect the utility of nucleic acids sequences expression in assessing the clinical outcome of a female mammal suffering from breast cancer. Methods for determining quantities of nucleic acids expression in a biological sample are well known from one of skill in the art. As an example of such methods, one can cite northern blot, cDNA array, oligo arrays or quantitative Reverse Transcription-PCR.

[0084] Preferably said methodology is cDNA arrays or oligo arrays, which allows the quantitative study of numerous candidate genes mRNA expression levels.

[0085] DNA arrays consist of large numbers of DNA molecules spotted in a systematic order on a solid support or substrate such as a nylon membrane, glass slide, glass beads or a silicon chip. Depending on the size of each DNA spot on the array, DNA arrays can be categorized as microarrays (each DNA spot has a diameter less than 250 microns) and macroarrays (spot diameter is grater than 300 microns). When the solid substrate used is small in size, arrays are also referred to as DNA chips. Depending on the spotting technique used, the number of spots on a glass microarray can range from hundreds to thousands.

[0086] Typically, a method of monitoring gene expression by DNA array involves the following steps:

[0087] a) obtaining a polynucleotide sample from a subject; and

[0088] b) reacting the sample polynucleotide obtained in step (a) with a probe immobilized on a solid support wherein said probe consist of polynucleotides having the nucleic acids sequence as previously described, fragments, derivative or complementary sequence thereof.

[0089] c) detecting the reaction product of step (b).

[0090] In the present invention, the term "polynucleotide" refers to a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. A polynucleotide in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

[0091] In the present invention, the term "fragment" refers to a sequence of nucleic acids that allows a specific hybridization under stringent conditions, as an example more than 10 nucleotides, preferably more than 15 nucleotides,

and most preferably more than 25 nucleotides, as an example more than 50 nucleotides or more than 100 nucleotides.

[0092] In the present invention, the term "derivative" refers to a sequence having more than 80% identity with an identified nucleic acid sequence, preferably more than 90% identity, as an example more than 95% identity, and most particularly more than 99% identity.

[0093] In the present invention, the term "immobilized on a support" means bound directly or indirectly thereto including attachment by covalent binding, hydrogen bonding, ionic interaction, hydrophobic interaction or otherwise.

[0094] The polynucleotide sample isolated from the subject and obtained at step (a) is RNA, preferably mRNA. Said polynucleotide sample isolated from the patient can also correspond to cDNA obtained by reverse transcription of the mRNA, or a product of ligation after specific hybridization of specific probes to mRNA or cDNA.

[0095] Preferably, the polynucleotide sample obtained at step (a) is labeled before its reaction at step (b) with the probe immobilized on a solid support. Such labeling is well known from one of skill in the art and includes, but is not limited to, radioactive, colorimetric, enzymatic, molecular amplification, bioluminescent, electrochemical or fluorescent labeling.

[0096] Advantageously, the reaction product of step (c) is quantified by further comparison of said reaction product to a control sample.

[0097] Detection preferably involves calculating/quantifying a relative expression (transcription) level for each nucleic acids sequence.

[0098] Then, the determination of the relative expression level for each nucleic acid sequences previously described enables to assess the clinical outcome of the subject –i.e. female mammal- suffering from breast cancer by the method of the invention.

[0099] The method of assessing the clinical outcome of a female mammal suffering from breast cancer can further involve a step of taking a biological sample, preferably breast cancer tissue or cells from a female mammal. Such methods of sampling are well known of one of skill in the art, and as an example, one can cite surgery.

[00100] The provided method may also correspond to an *in vitro* method, which does not include such a step of sampling.

[00101] Also provided are methods to determine if a pharmaceutical treatment, especially chemotherapy treatment, influences the clinical outcome of a female mammal suffering from breast cancer, which methods involve quantifying said nucleic acids sequences expression in a biological sample from a female mammal and determining the score (S_C) for said female mammal.

[00102] Further embodiments are methods to assess or identify a therapeutic or pharmaceutical agent for its potential effectiveness, efficacy or side effects relating to the clinical outcome, which methods involve quantifying said nucleic acids sequences in a biological sample from a female mammal suffering from breast cancer and determining the score (S_C) for said female mammal.

[00103] Also provided herein are methods of assessing a change in the propensity of clinical outcome from a female mammal suffering from breast cancer, wherein the methods involve taking at least two biological samples from the female mammal, one of which is taken before and one after an event. In various specific embodiments, the event involves passage of time (e.g., minutes, hours, days, weeks, months, or years), treatment with a therapeutic agent (or putative or potential therapeutic agent), treatment with a pharmaceutical agent (or putative or potential pharmaceutical agent).

[00104] One specific provided embodiment is a method of determining whether or to what extent a condition influences the clinical outcome of a female mammal suffering from breast cancer. This method involves subjecting

a subject to the condition, taking a biological sample from the subject, analyzing the biological sample to produce a score (S_C) for said subject, and comparing said score (S_C) for the subject with a control. From this comparison, conclusions are drawn about whether or to what extent the condition influences the clinical outcome of female mammal suffering from breast cancer based on differences or similarities between the test score (S_C) and the control. As contemplated for this embodiment, a condition to which the subject is subjected can include but is not limited to application of a pharmaceutical or therapeutic agent or candidate agent.

[00105] Subject: a female mammal.

[00106] In specific examples of such methods, the nucleic acids sequences expression profile is a pre-condition score (S_C) from the subject or a compiled score (S_C) assembled from a plurality of individual score (S_C). In other examples, the control score (S_C) is a control or a baseline established from previously described control score (S_C).

[00107] Pharmaceutical treatment: any agent treatment, regimen, or dosage, such the administration of a protein, a peptide (e.g., hormone), other organic molecule or inorganic molecule or compound, or combination thereof, that has or should have beneficial effects on clinical outcome when properly administrated to a subject, preferably said agents are used in chemotherapy.

[00108] Unless otherwise explained, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including explanations of terms, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

[00109] In various embodiments, the provided methods further comprise the step of selecting the pharmaceutical treatment that improves the clinical outcome of a female mammal suffering from breast cancer.

[00110] The present invention will be understood more clearly on reading the description of the experimental studies performed in the context of the research carried out by the applicant, which should not be interpreted as being limiting in nature.

EXAMPLE 1: Identification of significant metagenes combination:

1) Goals:

[00111] While it is now possible to assess patients' responses to drugs with respect to their genomic profile, the standard adjuvant chemotherapy (anthracyclines and taxanes) for non metastatic breast cancer may not be systematically appropriate: according to their genomic profile, women may rather benefit from a treatment based on anthracyclines alone without taxane.

[00112] The primary objective was to identify a gene set, which discriminate two groups of patients with different clinical outcome based on gene expression. This goal was reached by: defining the gene expression profiles, using 9.000-genes microarrays, of 323 tumours obtained from patients treated with adjuvant anthracycline-based CT without taxanes (identification set), grouping individual genes in metagenes and identifying metagenes closely correlated with the biological status of ER, PR, HER2/Neu, MIB/KI67, EGFR status of the sample as determined by the mean of independent methods such as Immunohistochemistry or FISH. Then we combined these metagenes using a Cox proportional hazard ratio analysis to separate patients according to clinical outcome. This latter step providing a model consisting of a score expressed as a linear combination such as Score = $\sum \beta_i.x_i$ where β_i.is a fixed parameter and x_i is the value of the metagene.

[00113] The secondary objective was to prospectively validate the Cox model and its metagene component for predicting clinical outcome in an independent cohort of patients (validation set). This goal was reached by

defining the gene expression profiles of 164 tumours, using the same technology, obtained from patients treated with adjuvant anthracycline-based CT without taxanes in the context of a multicentric clinical trial.

2) Patients :

[00114] We profiled a multicentric and retrospective series of 504 early breast cancers (Institut Paoli Calmettes, Centre Léon Bérard, Institut Bergonié and tumours from clinicals trials PACS01 and PEGASE01) treated with adjuvant anthracycline-based and non taxane-based CT. Clinical and pathological criteria for each patient are summarized in the following table and correspond to the identification and the validation sets.

[00115] Global population demography:

Age	median (min - max)	50 (11 - 90)
menopausal	Y	210 (41.8 %)
	N	292 (58.2 %)
Tumour size	pT1	105 (21 %)
	pT2	317 (63.5 %)
	pT3	77 (15.4 %)
N (Node)	N-	67 (13.3 %)
	N+	437 (86.7 %)
Node category	N1 (N=0)	67 (13.3 %)
	N2 (N= 1 to 3)	248 (49.2 %)
	N3 (N>3)	189 (37.5 %)
grade SBR	I	66 (13.4 %)
	II	221 (45 %)
	III	204 (41.5 %)
RE (10%)	RE-	150 (31 %)
(Estrogen Receptor)	RE+	334 (69 %)
RP (10%)	RP-	199 (41.5 %)
(Progesterone Receptor)	RP+	280 (58.5 %)
RH (10%)	RH-	115 (23.8 %)

(Hormone Receptor; RE and/or RP)	RH+	369 (76.2 %)
Her2/neu	0-1-2 3	308 (85.1 %) 54 (14.9 %)
Hormonotherapy	N Y	212 (45.3 %) 256 (54.7 %)
Follow-up	median [IC95]	71 mois [68-73]
Metastasis	N Y	364 (72.2 %) 140 (27.8 %)
5 years MFS (Metastasis Free Survival)	MFS [IC95]	73.52 [69.55 – 77.72]
Deaths from breast cancer	N Y	412 (81.7 %) 92 (18.3 %)
Specific Survival at 5 years	SS [IC95]	84.87 [81.56 – 88.31]

[00116] Identification Set demography (IPC, Lyon, Total):

Age	median	52	48	51
menopausal	Y	110 (52%)	67 (61%)	177 (55%)
	N	103 (48%)	40 (36%)	143 (45%)
Tumour size	pT1	57 (27%)	17 (16%)	74 (23%)
	pT2	115 (54%)	73 (66%)	188 (58%)
	pT3	41 (19%)	20 (18%)	61 (19%)
N	N-	56 (26%)	11 (10%)	67 (21%)
	N+	157 (74%)	99 (90%)	256 (79%)
N.cat	N1	43 (20%)	12 (11%)	55 (17%)
	N2	72 (34%)	60 (55%)	132 (41%)
	N3	98 (46%)	41 (34%)	139 (43%)
grade SBR	I	29 (14%)	16 (15%)	45 (14%)
	II	99 (46%)	55 (50%)	154 (48%)
	III	82 (38%)	39 (35%)	121 (38%)

RE (10%)	RE- RE+	80 (43%) 104 (57%)	32 (31%) 76 (69%)	112 (38%) 180 (62%)
RP (10%)	RP- RP+	62 (38%) 100 (62%)	30 (29%) 78 (71%)	92 (34%) 178 (56%)
RH (10%)	RH- RH+	46 (24%) 143 (76%)	23 (21%) 86 (79%)	69 (23%) 229 (77%)
Her2/neu	0-1-2 3 NA	174 (82%) 35 (16%) 4 (2%)	19 (17%) 1 (<1%) 90 (78%)	193 (60%) 36 (11%) 94 (29%)
Hormonotherapy	N Y	77 (36%) 136 (64%)	38 (35%) 72 (65%)	105 (33%) 208 (67%)
Follow-up	median	61	84	70
Metastasis	N Y	163 (77%) 50 (23%)	73 (66%) 37 (34%)	236 (70%) 87 (30%)
5 year MFS	MFS	77.5%	71.8%	75.6%
Deaths from breast cancer	N Y	172 (81%) 41 (19%)	85 (77%) 25 (23%)	257 (80%) 66 (20%)
Specific Survival at 5 years	SS	81.7%	82.7%	82%

[00117] Validation Set demography (PACS01, Bordeaux, total):

Age	median (min - max)	50	44	49
menopausal	Y N	60 (37%) 104 (63%)	1 (6%) 15 (88%)	61 (34%) 119 (66%)
Tumour size	pT1 pT2 pT3	27 (16%) 116 (71%) 16 (10%)	9 (53%) 8 (47%) 0 (0%)	37 (20%) 125 (69%) 16 (9%)
N	N- N+	0 (0%) 164 (100%)	0 (0%) 17 (100%)	0 (0%) 181 (0%)
	N1	0 (0%)	0 (0%)	0 (0%)

N.cat	N2	80 (49%)	14 (82%)	94 (52%)
	N3	84 (51%)	3 (18%)	87 (48%)
grade SBR	I	24 (15%)	2 (12%)	26 (14%)
	II	60 (37%)	7 (41%)	67 (37%)
	III	75 (46%)	8 (47%)	83 (46%)
RE (10%)	RE-	121 (74%)	5 (29%)	126 (70%)
	RE+	31 (19%)	12 (71%)	43 (24%)
RP (10%)	RP-	54 (33%)	4 (24%)	58 (32%)
	RP+	110 (67%)	13 (76%)	123 (68%)
RH (10%)	RH-	33 (20%)	3 (18%)	37 (20%)
	RH+	131 (80%)	14 (82%)	145 (80%)
Her2/neu	0-1-2	113 (69%)	13 (76%)	126 (70%)
	3	15 (9%)	4 (24%)	19 (10%)
	NA	36 (22%)	0 (0%)	36 (20%)
Hormonotherapy	N	53 (32%)	14 (82%)	67 (37%)
	Y	76 (46%)	3 (18%)	79 (44%)
	NA	36 (22%)	0 (0%)	36 (19%)
Follow-up	median	59	123	59
Metastasis	N	127 (77%)	12 (71%)	139 (77%)
	Y	37 (23%)	5 (29%)	42 (23%)
5y MFS	MFS	78.6%	70.6%	77.8%
Deaths from breast cancer	N	140 (85%)	12 (71%)	152 (84%)
	Y	24 (15%)	5 (29%)	29 (16%)
Specific Survival at 5 years	SS	85.4%	70.6%	84%

3) Method for gene profiling:

[00118] Radio-labeled [$A^{33}P$]-dCTP cDNA probes are obtained by reverse transcription from 3µg of total RNA. Probes are then hybridised on IPSOGEN's 10K Discovery*Chip*™, consisting of nylon membranes containing 9600 spotted cDNA (Discovery™ platform).

[00119] Following hybridization, membranes are washed and exposed to phosphor-imaging plates, then scanned with a Fuji-BAS 5000 machine. Signal intensities are quantified using the Fuji ArrayGauge v1.2 program, and the resulting raw data are analysed.

4) Analysis :

4-1 : Normalisation and filtering :

[00120] Raw data are exported from Ipsogen database. Spots for which spotted DNA amount is too low are invalidated from further analysis. Data are then normalized as compared to a reference sample using a non-linear rank based method (Sabatti et al., 2002). Normalized data are then filtered to eliminate low intensity genes, for which expression level is comparable to non-specific signal and the measure highly uncertain.

[00121] Data quality controls are performed based on hierarchical clustering grouping samples and genes according to their profile similarity. Biological pertinence of samples and genes clusters insures good quality data and allow for further analysis.

[00122] Since we analysed several samples series we performed a supplementary data normalization to insure inter-series comparability. Comparability was checked by hierarchical clustering.

4-2: Phenotypic signatures identification:

[00123] We performed supervised analysis using MaxT method available on Bioconductor (Ge, Dudoit & Speed, 2003) for several phenotypic markers: ER, PR, HER2/Neu, MIB/KI67, EGFR. The five markers were all measured by standard immunohistochemistry (IHC).

[00124] Supervised analyses were performed on a 159 samples identification set for ER, PR, HER2/Neu and EGFR markers, and on a 114 samples identification set for MIB/KI67. Each identified signature was then validated on one to four independent datasets.

[00125] Validation consisted in status prediction for independent samples using the LPS method (Linear Predictor Score) (Wright et al., PNAS, 2003, vol. 100, no. 17, 9991-9996). Prediction of all independent samples allowed for sensitivity and specificity evaluation for each identified signature.

4-3 : Metagenes calculation :

[00126] We considered as a metagene a group of genes for which expression variation (but not necessarily expression level) across tumors is correlated. The assumption is that the error made on the measurement of expression level from a single gene is highly reduced when considering several genes. So even in the case that an individual gene is poorly measured, its contribution in the metagene value is weighted by the number of genes considered and the final value for the metagene is lowly affected.

[00127] Metagenes were calculated from both supervised and unsupervised data.

[00128] Metagenes from phenotypic signatures: Phenotypic signatures correspond to genes correlated with a given phenotypic marker assessed by current standards such as immunohistochemistry (IHC) or FISH. A gene is considered correlated by a modified t test (MaxT method) which tests the significance of differential expression with a 5% risk. Each phenotypic signature is composed of two gene subsets, which expression levels are anti-correlated. One group of gene is overexpressed in a group of tumours (for exemple ER+ tumours) while the other group is underexpressed in the same group of tumours. Although expression variation is correlated across samples, expression levels may vary between genes, then leading to non robust average expression. It is assumed that even if expression levels vary, differential expression according to a reference sample belongs to the same dynamic range for all genes, allowing average calculation. For each tumour, each gene measure is divided by the expression level of the gene in a reference sample (log ratio) and the corresponding metagene is the average of those log ratios.

[00129] Each signature allowed the calculation of two anti-correlated metagenes. For instance, ER signature gives 2 metagenes, underER (genes under expressed in ER+ tumours) and overER (genes over expressed in ER+ tumours).

[00130] Metagenes from unsupervised analyses: we also defined metagenes as groups of genes with correlated expression variation across samples based on hierarchical clustering on a 468 samples set. A group of genes was retained if it contained at least 5 genes and had a node correlation coefficient higher than 0.5. Groups of genes that corresponded to previously identified metagenes by supervised analysis were not further considered. Metagenes were obtained as the mean of the log ratios of the genes contained in a given group.

4-4 : Biostatistics

[00131] Since we failed to identify any robust gene signature based on classical supervised analysis for the metastasis, it seems that obviously a single set of correlated genes is not able to predict metastasis.

[00132] The biostatistic approach was then based on survival analysis, and the objective was, instead of separating metastasis from non metastasis patients, to identify two groups of patients with significantly different outcome. The event considered is the metastasis without considering any previous event such as local relapse.

[00133] Model calculation: We used the Cox regression to identify a combination of metagenes able to add prognostic information to already existing prognostic factors, such as SBR grade, tumour size, or lymph node involvement. Cox proportional hazard ratio analysis consists in the calculation of a likelihood function, which gives for a patient the probability to observe the event at a given time (death, metastasis), knowing that he survived until this time. The likelihood function is independent of time, and takes into account a

"baseline" risk which is common to every patient, and the risk which is associated to different explanatory variables (which values differ between patients). The baseline risk function is unknown and eliminated as far as ratios between patients are considered. Then, the log-likelihood is defined as a linear function of explanatory variables, each one being appropriately weighted by a given coefficient. The coefficients are estimated by the algorithm to maximize the log-likelihood function.

[00134] For this, we use a forward stepwise approach to select the most significant metagenes, the threshold p-value being fixed to 10%. To obtain a model dependant on metagenes information and not influenced by already known clinical parameters, the analysis was stratified on the clinical parameters SBR grade, tumour size and lymph node synthesized in a single parameter, the NPI (Nottingham Prognostic Index). Moreover, since the identification set was composed of patients originating from different anti-cancer centers, we also stratified the analysis on the center of origin.

[00135] Once a combination of metagenes was obtained we calculated for each patient a score based on the linear combination of the metagenes values weighted by the coefficient calculated by the algorithm for each metagene. The exponential value of the coefficient corresponds to the hazard ratio associated to the metagene. For each parameter estimation, the algorithm gives the 95% confidence interval. Hence any combination of values comprised in the confidence intervals can be used to separate patients into significantly different prognostic groups.

[00136] Prognostic groups determination : The distribution of the scores in the identification set was used to determine the most significant cut-off to separate patients into two groups of different outcome. We tested three thresholds, 1st, 2nd, and 3rd quartile, and performed in each case the logrank test to compare the two groups of patients. We used a step by step approach to define the optimized threshold, testing all score values as a potential threshold.

[00137] The cut-off was the one for which the p value associated to the log rank test was the most significant.

[00138] Validation on an independent validation set: for each patient of the validation set, we calculated the score and separated the patients into two prognostic groups using the coefficients and the threshold determined on the identification set. The score was calculated without considering the outcome (DFS-Disease Free Survival) of individual patients.

[00139] The validation was appreciated by the p value of the log rank test, which has to be <5% to consider the model validated.

[00140] We verified that the identified model effectively added relevant information as compared to standard parameters by performing multivariate Cox analyses which integrate clinical parameters and the model.

[00141] Sample prediction: For any new sample to be predicted raw data are normalized according to the reference sample previously defined and metagenes are calculated. The formula calculated on the identification set is then applied to the new sample, allowing the attribution of a specific score to each sample. The score is compared to the threshold optimized from the identification procedure and the patient is declared to belong to the good prognosis group if its score is lower or equal to the threshold and to the poor prognosis group if its score is higher than the threshold.

5) Results :

5-1 : Metagene selection

[00142] We started from 9 metagenes calculated from supervised analyses, and 17 metagenes from unsupervised analysis.

[00143] A first analysis based on the correlation between metagenes and robustness reduced the potential candidates to 19 metagenes, 7 from supervised analysis and 12 from unsupervised analysis.

5-2 : Univariate analysis

[00144] Each metagene was first tested in a univariate Cox analysis, and none of them could be found significant alone as shown in the following table.

Variable	Parameter Estimate	Hazard Ratio	p value
underER	0.468	1.597	0.59
underPR	-0.474	0.622	0.32
underEGFR	-1.132	0.322	0.06
overEGFR	-0.261	0.771	0.59
underMIB	-0.951	0.387	0.18
overMIB	0.927	2.528	0.37
overERBB2	0.089	1.094	0.88
MG48	-0.398	0.672	0.46
MG187	-0.453	0.636	0.38
MG66	-0.423	0.655	0.40
MG27	0.193	1.21	0.65
MG51	-0.182	0.834	0.75
MG141	-0.076	0.927	0.90
MG144	-0.256	0.774	0.70
MG171	0.131	1.14	0.82
MG240	-0.304	0.738	0.55
MG310	0.271	1.31	0.61
MG448	-1.03	0.358	0.10
MG1001	-0.34	0.712	0.31

5

5-3 : Description of selected metagenes and combination thereof

[00145] Multivariate Cox analyses allowed identification of significant metagenes and combinations thereof associated with prognosis. The

constituents of the selected metagenes and these combinations are described hereafter.

EXAMPLE 2: Identification of a first metagene combination:

[00146] The Cox analysis using forward stepwise procedure identified the three following significant metagenes (underER, underPR and underEGFR) associated with good or poor prognosis.

[00147] Table I (Metagene UnderER)

Gene	Unigene Cluster	Regulation	P value	Ref. Seq	Reduced metagene 27	Reduced metagene 20
ITGB3 integrin, beta 3 (platelet glycoprotein iiia, antigen cd61)	ughs.218040:186	-	0,00001	SEQ ID No:374 (nm_000212)	+	+
PADI2 peptidyl arginine deiminase, type ii	ughs.33455:186	-	0,00001	SEQ ID No:1027 (nm_007365)	+	+
SOD2 superoxide dismutase 2, mitochondrial	ughs.487046:186	-	0,00001	SEQ ID No:598 (nm_000636)	+	+
FLJ13154 hypothetical protein flj13154	ughs.408702:186	-	0,00003	SEQ ID No:717 (nm_024598)	-	-
HDAC2 histone deacetylase 2	ughs.3352:186	-	0,00004	SEQ ID No:573 (nm_001527)	+	+
SLAC2-B	N_A	-	0,00006	SEQ ID No:83 (nm_015065)	+	+
S100A8 s100 calcium binding protein a8 (calgranulin a)	ughs.416073:186	-	0,00006	SEQ ID No:12 (nm_002964)	+	+

GSTP1 glutathione s-transferase pi	ughs.523836:186	-	0,00006	SEQ ID No:405 (nm_000852)	+	+
LCN2 lipocalin 2 (oncogene 24p3)	ughs.204238:186	-	0,00012	SEQ ID No:856 (nm_005564)	+	+
MYBL2 v-myb myeloblastosis viral oncogene homolog (avian)-like 2	ughs.179718:186	-	0,00013	SEQ ID No:384 (nm_002466)	+	-
PFKP phosphofructokinase, platelet	ughs.26010:186	-	0,00081	SEQ ID No:167 (nm_002627)	+	+
STK6 serine/threonine kinase 6	ughs.250822:186	-	0,00134	SEQ ID No:51 (nm_198433)	+	+
GPR125 g protein-coupled receptor 125	ughs.99195:186	-	0,00153	SEQ ID No:999 (nm_145290)	-	-
DSCR1 down syndrome critical region gene 1	ughs.282326:186	-	0,00206	SEQ ID No:979 (nm_004414)	-	-
FAT fat tumor suppressor homolog 1 (drosophila)	ughs.481371:186	-	0,0023	SEQ ID No:2 (nm_005245)	+	-
VGLL1 vestigial like 1 (drosophila)	N_A	-	0,00247	SEQ ID No:98 (nm_016267)	+	+
MMP7 matrix metalloproteinase 7 (matrilysin, uterine)	ughs.2256:186	-	0,00264	SEQ ID No:751 (nm_002423)	+	+
ENO1 enolase 1, (alpha)	ughs.517145:186	-	0,00348	SEQ ID No:696 (nm_001428)	+	+
cdna clone image:4831215	ughs.175285:186	-	0,00429	SEQ ID No:1050 (BC034638)	+	-

SCP2 sterol carrier protein 2	ughs.476365:186	-	0,00469	SEQ ID No:488 (nm_002979)	-	-
CEBPB ccaat/enhancer binding protein (c/ebp), beta	ughs.517106:186	-	0,00507	SEQ ID No:262 (nm_005194)	+	+
TGM1 transglutaminase 1 (k polypeptide epidermal type i, protein-glutamine-gamma-glutamyltransferase)	ughs.508950:186	-	0,00695	SEQ ID No:1020 (nm_000359)	+	+
	N_A	-	0,00764	SEQ ID No:1106 (BC015969)	-	-
GGH gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	ughs.78619:186	-	0,00881	SEQ ID No:952 (nm_003878)	+	-
GSTA4 glutathione s-transferase a4	ughs.485557:186	-	0,00995	SEQ ID No:675 (nm_001512)	-	-
FN5 b-cell cll/lymphoma 7b	ughs.438064:186	-	0,0109	SEQ ID No:289 (nm_020179)	-	-
CCNB2 glutamate decarboxylase 1 (gad 1)	ughs.194698:186	-	0,01221	SEQ ID No:553 (nm_004701)	-	-
CTSC cathepsin c	ughs.128065:186	-	0,01501	SEQ ID No:579 (nm_001814)	+	+
PBEF1 pre-b-cell colony enhancing factor 1	ughs.489615:186	-	0,01621	SEQ ID No:760 (nm_005746)	+	+

S100A6 s100 calcium binding protein a6 (calcyclin)	ughs.275243:186	-	0,01719	SEQ ID No:805 (nm_014624)	+	+
RDX radixin	ughs.263671:186	-	0,01753	SEQ ID No:361 (nm_002906)	+	-
GPR126 g protein-coupled receptor 126	ughs.318894:186	-	0,01886	SEQ ID No:448 (nm_198569)	-	-
MMP15 matrix metalloproteinase 15 (membrane-inserted)	ughs.80343:186	-	0,0274	SEQ ID No:170 (nm_002428)	-	-
KLK6 kallikrein 6 (neurosin, zyme)	ughs.79361:186	-	0,02892	SEQ ID No:878 (nm_002774)	+	+
	N_A	-	0,0351	SEQ ID No:1117	-	-
BOK bcl2-related ovarian killer	ughs.293753:186	-	0,03747	SEQ ID No:612 (nm_032515)	+	+
CDKL5 cyclin-dependent kinase-like 5	ughs.435570:186	-	0,03754	SEQ ID No:540 (nm_003159)	-	-
CSTB cystatin b (stefin b)	ughs.695:186	-	0,0382	SEQ ID No:823 (nm_000100)	-	-
LOC151194 similar to hepatocellular carcinoma-associated antigen hca557b	ughs.552610:186	-	0,03884	SEQ ID No:131 (nm_145280)	-	-
NFIB nuclear factor i/b	ughs.370359:186	-	0,03949	SEQ ID No:705 (nm_005596)	-	-
LAD1 ladinin 1	ughs.519035:186	-	0,04184	SEQ ID No:31 (nm_005558)	+	-

MGC11271 hypothetical protein mgc11271	ughs.143288:186	-	0,04312	SEQ ID No:199 (nm_024323)	+	-

Table II (Metagene Under PR)

Gene	Unigene Cluster	Regulation	P value	Ref. Seq	Reduced Metagene 35	Reduced Metagene 6
SOD2 superoxide dismutase 2, mitochondrial	ughs.487046:186	-	0,00001	SEQ ID No:598 (nm_000636)	-	-
IGHG1 immunoglobulin heavy constant gamma 1 (g1m marker)	ughs.510635:186	-	0,00001	SEQ ID No:1122	-	-
KDR kinase insert domain receptor (a type iii receptor tyrosine kinase)	ughs.479756:186	-	0,00011	SEQ ID No:364 (nm_002253)	+	+
KLF1 kruppel-like factor 1 (erythroid)	ughs.37860:186	-	0,00014	SEQ ID No:387 (nm_006563)	+	-
CASP9 caspase 9, apoptosis-related cysteine protease	ughs.329502:186	-	0,00016	SEQ ID No:34 (nm_001229)	+	+
BCL2 b-cell cll/lymphoma 2	ughs.150749:186	-	0,00018	SEQ ID No:657 (nm_000633)	+	+
MYBL2 v-myb myeloblastosis viral oncogene homolog (avian)-like 2	ughs.179718:186	-	0,00025	SEQ ID No:384 (nm_002466)	-	-
ADAM10 a disintegrin and metalloproteinase domain 10	ughs.172028:186	-	0,00031	SEQ ID No:451 (nm_001110)	-	-
GPR125 g protein-coupled receptor 125	ughs.99195:186	-	0,00032	SEQ ID No:999 (nm_145290)	-	-

	ughs.26192:186	-	0,00049	SEQ ID No:1056 (AK126297)	+	-
TGFBR3 transforming growth factor, beta receptor iii (betaglycan, 300kda)	ughs.482390:186	-	0,00061	SEQ ID No:15 (nm_003243)	+	-
LOC91316 similar to bk246h3.1 (immunoglobulin lambda-like polypeptide 1, pre-b-cell specific)	ughs.407693:186 ; ughs.148656:186	-	0,00072	SEQ ID No:1090 (AK125808)	-	-
	ughs.416139:186	-	0,00074	SEQ ID No:1120	+	-
S100A8 s100 calcium binding protein a8 (calgranulin a)	ughs.416073:186	-	0,00079	SEQ ID No:12 (nm_002964)	-	-
PIM2 *pim-2 oncogene*	ughs.496096:186	-	0,00088	SEQ ID No:743 (nm_006875)	-	-
TP53 tumor protein p53 (li-fraumeni syndrome)	ughs.408312:186	-	0,00104	SEQ ID No:414 (nm_000546)	+	-
ITGB3 integrin, beta 3 (platelet glycoprotein iiia, antigen cd61)	ughs.218040:186	-	0,00118	SEQ ID No:374 (nm_000212)	+	-
LAMB1 laminin, beta 1	ughs.489646:186	-	0,00118	SEQ ID No:711 (nm_002291)	+	-
SILV silver homolog (mouse)	*ughs.95972:186*	-	0,00118	SEQ ID No:663 (nm_006928)	+	-
cdna flj42596 fis, clone brace3010283	ughs.113271:186	-	0,00121	SEQ ID No:1102 (AK124587)	-	-
PIGR polymeric immunoglobulin receptor	ughs.497589:186	-	0,00123	SEQ ID No:237 (nm_002644)	+	-
CSH1 chorionic somatomammotropin hormone 1 (placental lactogen)	ughs.347963:186	-	0,00161	SEQ ID No:60 (nm_022640)	+	-

RDX radixin	ughs.263671:186	-	0,00176	SEQ ID No:361 (nm_002906)	-	-
ETF1/FLT1 eukaryotic translation termination factor 1/fms-related tyrosine kinase 1	ughs.483494:186 ; ughs.507621:186	-	0,0019	SEQ ID No:119 (nm_004730) or SEQ ID No:1109 (NM_002019)	+	-
PFKP phosphofructokinase, platelet	ughs.26010:186	-	0,00193	SEQ ID No:167 (nm_002627)	-	-
CXORF38 chromosome x open reading frame 38	ughs.495961:186	-	0,002	SEQ ID No:339 (nm_144970)	+	+
MGC15606 family with sequence similarity 55, member c	ughs.130195:186	-	0,00207	SEQ ID No:333 (nm_145037)	-	-
SLAC2-B slac2-b	N_A	-	0,00236	SEQ ID No:83 (nm_015065)	-	-
FLJ10986 hypothetical protein flj10986	ughs.444301:186 ; ughs.439112:186	-	0,00261	SEQ ID No:330 (nm_018291)	+	-
SERPINB1 serine (or cysteine) proteinase inhibitor, clade b (ovalbumin), member 1	ughs.381167:186	-	0,00368	SEQ ID No:1024 (nm_030666)	+	-
RPS6KA3 ribosomal protein s6 kinase, 90kda, polypeptide 3	ughs.445387:186	-	0,00482	SEQ ID No:229 (nm_004586)	+	+
GATA6 gata binding protein 6	ughs.514746:186	-	0,00491	SEQ ID No:925 (nm_005257)	+	-
MTIF2 mitochondrial translational initiation factor 2	ughs.149894:186	-	0,00535	SEQ ID No:788 (nm_001005369)	-	-
	N_A	-	0,00572	SEQ ID No:1104 (AK128524)	+	-

	N_A	-	0,00635	SEQ ID No:1103 (BX108410)	+	-
IFNGR1 interferon gamma receptor 1	ughs.520414:186	-	0,00656	SEQ ID No:66 (nm_000416)	+	-
EBF early b-cell factor	ughs.308048:186	-	0,00665	SEQ ID No:1030 (nm_024007)	-	-
	N_A	-	0,00729	SEQ ID No:1119	+	+
p66alpha GATA zinc finger domain containing 2A (p66alpha)	ughs.551742:186	-	0,00741	SEQ ID No:1068 (AK024670)	+	-
FKBP1A fk506 binding protein 1a, 12kda	ughs.471933:186	-	0,00885	SEQ ID No:241 (nm_000801)	-	-
SNAPC3 small nuclear rna activating complex, polypeptide 3, 50kda	ughs.546299:186	-	0,00887	SEQ ID No:398 (nm_003084)	-	-
IL2RB interleukin 2 receptor, beta	ughs.474787:186 ; ughs.555488:186	-	0,0097	SEQ ID No:74 (nm_000878)	+	-
Homo sapiens mRNA for FLJ00204 protein	ughs.535157:186	-	0,00973	SEQ ID No:1087 (AK074131)	-	-
ETV4 ets variant gene 4 (e1a enhancer binding protein, e1af)	ughs.434059:186	-	0,01003	SEQ ID No:955 (nm_001986)	-	-
IL1R2 interleukin 1 receptor, type ii	ughs.25333:186	-	0,01009	SEQ ID No:71 (nm_004633)	-	-
IGHG1 immunoglobulin heavy constant gamma 1 (g1m marker)	ughs.510635:186	-	0,01039	SEQ ID No:1105 (BC072392)	-	-
LCN2 lipocalin 2 (oncogene 24p3)	ughs.204238:186	-	0,01068	SEQ ID No:856 (nm_005564)	-	-
CMRF35 cd300c antigen	ughs.2605:186	-	0,01119	SEQ ID No:231 (nm_006678)	+	-

CXCL1 chemokine (c-x-c motif) ligand 1 (melanoma growth stimulating activity, alpha)	ughs.789:186	-	0,01174	SEQ ID No:593 (nm_001511)	+	-
MYBL2 v-myb myeloblastosis viral oncogene homolog (avian)-like 2	ughs.179718:186	-	0,0122	SEQ ID No:384 (nm_002466)	+	-
SLAMF8 slam family member 8	ughs.438683:186	-	0,01309	SEQ ID No:519 (nm_020125)	-	-
CTSC cathepsin c	ughs.128065:186	-	0,016	SEQ ID No:579 (nm_001814)	-	-
ENPP2 ectonucleotide pyrophosphatase/phosphodiesterase 2 (autotaxin)	ughs.190977:186	-	0,0205	SEQ ID No:1039 (nm_006209)	+	-
LAD1 ladinin 1	ughs.519035:186	-	0,02102	SEQ ID No:31 (nm_005558)	-	-
RABL3 rab, member of ras oncogene family-like 3	ughs.444360:186 ; ughs.548087:186	-	0,02205	SEQ ID No:327 (nm_173825)	+	-
HDAC2 histone deacetylase 2	ughs.3352:186	-	0,02428	SEQ ID No:573 (nm_001527)	-	-
VGLL1 vestigial like 1 (drosophila)	N_A	-	0,02447	SEQ ID No:98 (nm_016267)	-	-
npc-a-5 nasopharyngeal carcinoma-associated antigen npc-a-5	ughs.510543:186	-	0,02592	SEQ ID No:1059 (AK091113)	-	-
CDK4 cyclin-dependent kinase 4	ughs.95577:186	-	0,02615	SEQ ID No:886 (nm_000075)	+	-
ABCC5 atp-binding cassette, sub-family c (cftr/mrp), member 5	ughs.368563:186	-	0,02624	SEQ ID No:1032 (nm_005688)	+	-

MGC9913 hypothetical protein mgc9913	ughs.23133:186	-	0,02709	SEQ ID No:1091 (XM_378178)	-	-
FUT8 fucosyltransferase 8 (alpha (1,6) fucosyltransferase)	ughs.118722:186	-	0,02833	SEQ ID No:233 (nm_178155)	-	-
SFRP1 secreted frizzled-related protein 1	ughs.213424:186	-	0,03011	SEQ ID No:938 (nm_003012)	-	-
ARPC2 actin related protein 2/3 complex, subunit 2, 34kda	ughs.529303:186	-	0,03237	SEQ ID No:264 (nm_152862)	+	-
LILRB2 leukocyte immunoglobulin-like receptor, subfamily b (with tm and itim domains), member 2	ughs.534386:186	-	0,03294	SEQ ID No:546 (nm_005874)	-	-
IGKC immunoglobulin kappa constant	ughs.449621:186 ; ughs.546620:186	-	0,03458	SEQ ID No:1099 (BC066343)	-	-
SN sialoadhesin	ughs.31869:186	-	0,03771	SEQ ID No:1037 (nm_023068)	+	-
C1ORF38 chromosome 1 open reading frame 38	ughs.10649:186	-	0,03783	SEQ ID No:550 (nm_004848)	-	-
PADI2 peptidyl arginine deiminase, type ii	ughs.33455:186	-	0,0418	SEQ ID No:1027 (nm_007365)	-	-
MONDOA mlx interactor	ughs.437153:186	-	0,04548	SEQ ID No:1005 (nm_014938)	+	-
TAP1 transporter 1, atp-binding cassette, sub-family b (mdr/tap)	ughs.352018:186 ; ughs.552165:186	-	0,04583	SEQ ID No:820 (nm_000593)	-	-
CYP2D6 cytochrome p450, family 2, subfamily d, polypeptide 6	ughs.534311:186	-	0,04704	SEQ ID No:370 (nm_000106)	-	-

Table III (Metagene UnderEGFR)

Gene	Unigene Cluster	Regulation	P value	Ref. Seq	Reduced Metagene 34	Reduced Metagene 22
LOC255743: Nephronectin	N_A	-	0,00001	SEQ ID No:1071 (NM_001033047)	+	+
LU lutheran blood group (auberger b antigen included)	ughs.155048:186	-	0,00001	SEQ ID No:254 (nm_005581)	+	+
TFF1 trefoil factor 1 (breast cancer, estrogen-inducible sequence expressed in)	ughs.162807:186	-	0,00001	SEQ ID No:6 (nm_003225)	+	+
ESR1 estrogen receptor 1	ughs.208124:186	-	0,00001	SEQ ID No:883 (nm_000125)	+	+
XBP1 x-box binding protein 1	ughs.437638:186	-	0,00001	SEQ ID No:543 (nm_005080)	+	+
SCUBE2 signal peptide, cub domain, egf-like 2	ughs.523468:186	-	0,00001	SEQ ID No:681 (nm_020974)	+	+
GATA3 gata binding protein 3	ughs.524134:186	-	0,00001	SEQ ID No:63 (nm_001002295)	+	+
EIF2C3 eukaryotic translation initiation factor 2c, 3	ughs.530333:186	-	0,00001	SEQ ID No:212 (nm_024852)	+	+
C4A complement component 4b, telomeric	ughs.534847:186	-	0,00001	SEQ ID No:635 (nm_001002029)	+	+

TFF3 trefoil factor 3 (intestinal)	ughs.82961:186	-	0,00001	SEQ ID No:535 (nm_003226)	+	+
	N_A	-	0,00003	SEQ ID No:1125	+	+
NAT1 n-acetyltransferase 1 (arylamine n-acetyltransferase)	ughs.155956:186	-	0,00003	SEQ ID No:109 (nm_000662)	+	-
COL4A2 collagen, type iv, alpha 2	ughs.508716:186	-	0,00003	SEQ ID No:342 (nm_001846)	+	-
RABEP1 rabaptin, rab gtpase binding effector protein 1	ughs.551518:186	-	0,00003	SEQ ID No:927 (nm_004703)	+	-
	N_A	-	0,00005	SEQ ID No:1124	+	+
RHOBTB3 rho-related btb domain containing 3	ughs.445030:186	-	0,00006	SEQ ID No:124 (nm_014899)	-	-
CASKIN1/flj12650 cask interacting protein 1	ughs.530863:186 ; ughs.470259:186	-	0,00006	SEQ ID No:280 (nm_020764) or SEQ ID No:1110 (NM_024522)	+	-
CXXC5 cxxc finger 5	ughs.189119:186	-	0,00009	SEQ ID No:297 (nm_016463)	+	+
MAPT microtubule-associated protein tau	ughs.101174:186	-	0,0001	SEQ ID No:791 (nm_016835)	+	+
MGC24047 chromosome 1 open reading frame 64	ughs.29190:186	-	0,0001	SEQ ID No:210 (nm_178840)	+	-
MGC45441 hypothetical protein	ughs.488337:186	-	0,00026	SEQ ID No:827	+	+

mgc45441				(nm_152499)		
CYP2B6 Cytochrome P450, family 2, subfamily B, polypeptide 6	N_A	-	0,00065	SEQ ID No:1064 (NM_000767)	-	-
CROCC ciliary rootlet coiled-coil, rootletin	ughs.309403:186 ; ughs.135718:186	-	0,00072	SEQ ID No:147 (nm_014675)	-	-
USP21 ubiquitin specific protease 21	ughs.8015:186	-	0,00075	SEQ ID No:323 (nm_001014443)	-	-
TRAF5 tnf receptor-associated factor 5	ughs.523930:186	-	0,0011	SEQ ID No:106 (nm_004619)	-	-
GSTM2 glutathione s-transferase m2 (muscle)	ughs.279837:186	-	0,00127	SEQ ID No:181 (nm_000848)	+	-
DUSP4 dual specificity phosphatase 4	ughs.417962:186	-	0,0015	SEQ ID No:376 (nm_057158)	-	-
ASF1A asf1 anti-silencing function 1 homolog a (s. cerevisiae)	ughs.292316:186	-	0,00177	SEQ ID No:116 (nm_014034)	+	-
CSF2 colony stimulating factor 2 (granulocyte-macrophage)	ughs.1349:186	-	0,0024	SEQ ID No:252 (nm_000758)	-	-
CLSTN2 calsyntenin 2	ughs.158529:186	-	0,00247	SEQ ID No:797 (nm_022131)	-	-

GLI3 gli-kruppel family member gli3 (greig cephalopolysyndactyly syndrome)	ughs.199338:186	-	0,00282	SEQ ID No:911 (nm_000168)	-	-
REPS2 ralbp1 associated eps domain containing 2	ughs.186810:186 ; ughs.131188:186	-	0,00307	SEQ ID No:720 (nm_004726)	-	-
GSTM1 glutathione s-transferase m1	ughs.301961:186	-	0,00307	SEQ ID No:889 (nm_000561)	-	-
PLAT plasminogen activator, tissue	ughs.491582:186	-	0,00335	SEQ ID No:250 (nm_000930)	+	-
DLG5 discs, large homolog 5 (drosophila)	ughs.500245:186	-	0,00393	SEQ ID No:179 (nm_004747)	-	-
FLJ00012 flj00012 protein	ughs.21051:186	-	0,00396	SEQ ID No:786 (nm_033388)	-	-
SIDT2 sid1 transmembrane family, member 2	ughs.410977:186	-	0,00409	SEQ ID No:177 (nm_015996)	+	-
	N_A	-	0,00434	SEQ ID No:1047 (BC012900)	-	-
BCL9 b-cell cll/lymphoma 9	ughs.415209:186	-	0,00434	SEQ ID No:301 (nm_004326)	-	-
USP13 ubiquitin specific protease 13 (isopeptidase t-3)	ughs.175322:186	-	0,00516	SEQ ID No:207 (nm_003940)	+	+
DNALI1 dynein, axonemal, light intermediate polypeptide 1	ughs.406050:186	-	0,00606	SEQ ID No:936 (nm_003462)	-	-

FOXC1/RHOB forkhead box c1/ras homolog gene	ughs.348883:186 ; ughs.502876:186	-	0,00652	SEQ ID No:916 (nm_001453) or SEQ ID No:1116 (NM_004040)	+	+
	N_A	-	0,00699	SEQ ID No:1052 (BX096026)	+	+
KRT18 keratin 18	ughs.406013:186	-	0,00879	SEQ ID No:159 (nm_000224)	+	+
	ughs.548040:186	-	0,00889	SEQ ID No:1096 (AK127274)	-	-
DNAJC12 dnaj (hsp40) homolog, subfamily c, member 12	ughs.260720:186	-	0,0094	SEQ ID No:28 (nm_021800)	-	-
cdna flj41270 fis, clone bramy2036387	ughs.445414:186	-	0,00963	SEQ ID No:1054 (AK123264)	-	-
SPDEF/c8orf13 sam pointed domain containing ets transcription factor / chromosome 8 open reading frame 13	ughs.124299:186 ; ughs.485158:186	-	0,00981	SEQ ID No:25 (nm_012391) or SEQ ID No:1108 (NM_053279)	+	+
C20ORF23 chromosome 20 open reading frame 23	ughs.101774:186	-	0,01019	SEQ ID No:825 (nm_024704)	+	-
FLJ20366 hypothetical protein flj20366	ughs.390738:186	-	0,01278	SEQ ID No:145 (nm_017786)	+	-
COX6C cytochrome c oxidase subunit vic	ughs.351875:186	-	0,01401	SEQ ID No:491 (nm_004374)	-	-

RGS11 regulator of g-protein signalling 11	ughs.65756:186	-	0,01422	SEQ ID No:485 (nm_003834)	-	-
Hypothetical protein LOC153561	ughs.508559:186	-	0,01475	SEQ ID No:1072 (AY007114)	-	-
SEMA6B sema domain, transmembrane domain (tm), and cytoplasmic domain, (semaphorin) 6b	ughs.465642:186	-	0,01572	SEQ ID No:274 (nm_032108)	-	-
AP1G2 adaptor-related protein complex 1, gamma 2 subunit	ughs.343244:186	-	0,01707	SEQ ID No:258 (nm_080545)	-	-
AKAP8L a kinase (prka) anchor protein 8-like	ughs.399800:186	-	0,01817	SEQ ID No:292 (nm_014371)	-	-
PRKCBP1 protein kinase c binding protein 1	ughs.446240:186	-	0,01835	SEQ ID No:803 (nm_183047)	-	-
CENTG3 centaurin, gamma 3	ughs.195048:186	-	0,02053	SEQ ID No:349 (nm_031946)	-	-
genomic region on chromosome 1	ughs.159853:186	-	0,02456	SEQ ID No:1123	-	-
SLC40A1 solute carrier family 40 (iron-regulated transporter), member 1	ughs.529285:186	-	0,02463	SEQ ID No:763 (nm_014585)	-	-
CCND2 cyclin d2	ughs.376071:186	-	0,02723	SEQ ID No:438 (nm_001759)	-	-
KLHDC2 kelch domain containing 2	N_A	-	0,02795	SEQ ID No:94 (nm_014315)	-	-

ABCA3 atp-binding cassette, sub-family a (abc1), member 3	ughs.26630:186	-	0,03438	SEQ ID No:845 (nm_001089)	+	+
LOC143381 hypothetical protein loc143381	ughs.388347:186 ; ughs.557061:186	-	0,03705	SEQ ID No:1084 (BX648964)	-	-
FLJ21439 hypothetical protein flj21439	ughs.550536:186	-	0,03746	SEQ ID No:734 (nm_025137)	-	-
HOXA4 homeo box a4	ughs.77637:186	-	0,03897	SEQ ID No:943 (nm_002141)	-	-
CACNA1D/KIF5C Calcium channel, voltage-dependent, L type, alpha 1D subunit/kinesin family member 5c	ughs.476358:186 ; ughs.435557:186	-	0,03958	SEQ ID No:1085 (NM_000720)	+	+
GNG3 guanine nucleotide binding protein (g protein), gamma 3	ughs.179915:186	-	0,04937	SEQ ID No:276 (nm_012202)	+	-

[00148] Multivariate Cox analysis allowed estimation of parameters corresponding to each of the selected metagenes:

Metagene	Parameter estimation	P value (Chi square)	Hazard Ratio
UnderER	-2.90279	0.0906	0.055
UnderPR	-1.47423	0.0143	0.229
UnderEGFR	-4.17198	0.0012	0.015

[00149] On the basis of these parameters, the score for prognosis has been established as follows:

[00150] Score = -2.90279*underER - 1.47423*underPR - 4.17198*underEGFR

[00151] Threshold optimization: we tested all the possible thresholds. As an example 1st, 2nd and 3rd quartile of the score distribution of the training set and found 0.502, 0.0057 and <0.0001 respectively for the p value associated to the log rank test.

[00152] The 3rd quartile (cut-off=0.087646) was then defined as the optimal cut-off to separate patients into two groups with the highest significance.

[00153] The error on the score was integrated by calculating a confidence interval around the threshold, within which sample classification was considered non robust. Considering the score distribution Gaussian, we estimated the confidence interval around the threshold using standard deviation calculation method (estimated standard deviation of the population / $\sqrt{n}$).

[00154] The inventors have established that a woman having a score (S_C) of more than 0.136 have at least a double propensity of poor clinical outcome than a woman with a score (S_C) of less than 0.0393.

[00155] Model validation: the score was calculated for each of the 164 patients from the validation set and we separated the patients into two groups according the cut-off determined on the identification set. On the 164 patients, the model was well validated ($p=4.7\ 10^{-02}$, log rank test) and separated the patients into a good-prognosis group with 80% 5-year MFS (84% of patients) and a poor-prognosis group with 63% 5-year MFS (13% of patients), 3% of patients being not interpretable. On a subset of the validation set, constituted of the clinical trial PACS01 (N=128), we obtained similar validation ($p= 3.9\ 10^{-03}$, logrank test) with 88% of 5-year MFS in the good-prognosis group (80% of patients) and 65% of 5-year MFS in the poor-prognosis group (16% of patients, 4% of patients not interpretable).

[00156] Model performances: we performed multivariate analysis to determine the importance of the model as compared to standard clinical parameters. Even when considering grade, lymph node, ER status, age..., the model was still significant in the multivariate analysis, suggesting that it provides an independent, complementary and significant prognostic information.

[00157] Multivariate analysis on the global population (N=347)

		Hazard ratio	CI95 upper	CI95 lower	p
Age	<35 y	1			
	>=35y	0.66	0.16	2.79	p=0.57
Menopausal	N	1			
	Y	1.24	0.82	1.9	p=0.31
Tumour size	pT1	1			
	pT2-pT3	1.61	0.95	2.74	p=0.078
N	N-	1			
	N+	1.45	0.76	2.76	p=0.26
SBR grade	I	1			
	II-III	2.27	0.96	5.35	p=0.062
HR (10%)	HR-	1			
	HR+	0.86	0.53	1.4	p=0.54
Erbb2	0-1-2	1			
	3	1.17	0.66	2.07	p=0.58
Model	Good	1			
	Poor	2.61	1.65	4.11	$P=3.8\ 10^{-5}$

[00158] Multivariate analysis on the identification set (N=222)

	Hazard ratio	CI95 upper	CI95 lower	p

Age	<35 y >=35y	1 1.43	0.19	10.8	p=0.73
Menopausal	N Y	1 1.03	0.64	1.68	p=0.89
Tumour size	pT1 pT2-pT3	1 1.12	0.63	1.97	p=0.7
N	N- N+	1 1.73	0.9	3.34	p=0.1
SBR grade	I II-III	1 2.1	0.87	5.08	p=0.098
HR (10%)	HR- HR+	1 0.93	0.53	1.62	p=0.79
Erbb2	0-1-2 3	1 0.93	0.44	1.93	p=0.84
Model	Good Poor	1 2.18	1.3	3.64	P=0.003

[00159] Multivariate analysis on the PACS01 clinical trial (N=108)

		Hazard ratio	CI95 upper	CI95 lower	p
Age	<35 y >=35y	1 518527625.4	0	Inf	p=1
Menopausal	N Y	1 1.33	0.51	3.47	p=0.56
Tumour size	pT1 pT2-pT3	1 324628156.2	0	Inf	p=1
N	N- N+	1			p=NA
SBR grade	I II-III	1 287987535.5	0	Inf	p=1
HR (10%)	HR-	1			

	HR+	1.42	0.36	5.61	p=0.62
Erbb2	0-1-2	1			
	3	2.13	0.68	6.66	p=0.19
Model	Good	1			
	Poor	5.3	1.58	17.74	P=0.0068

[00160] Metagenes reduction:

[00161] In this model with underER, underPR and underEGFR, we defined the number of genes according to their significance in the metagene identification with the MaxT method. Even if the genes are well correlated between each other, some of them may be removed from further analysis, in order to reduce the number of genes to analyze and simplify the analysis process.

[00162] We calculated the correlation between each gene composing the metagene and the metagene, sorted the genes according to their increasing correlation to the metagene and progressively eliminated the genes the least correlated to the metagene, starting from 1 removed gene to all except one removed genes.

[00163] For each of these new sets of genes, we calculated a new metagene and its correlation with the original metagene. We selected given correlation cut-offs varying from 0.91 to 0.99 and integrated the corresponding new metagene in the model. This allowed us to generate a new score and prognostic group for each patient and to compare the attribution of a given prognostic group between the original model and the model with the optimized metagene. The criterion was equivalence between the 2 patients classification (with the original model and the optimized one) within the 2 prognostic groups.

[00164] As an example, we can reduce the number of genes from the metagene underER from 42 to 27 (Table I), while keeping 97% of equivalence (meaning that only 3% of patients are predicted in the opposite prognostic

group when optimizing the metagene) for patient classification in the two prognostic groups on the validation set. With 20 genes (Table I), the concordancy is still of 95%.

[00165] In the same way, the metagene underPR may be reduced from 73 to 35 (Table II) and 6 genes (Table II) with 96% and 94% equivalence respectively for patient classification in the validation set.

[00166] The metagene underEGFR may be reduced from 71 to 34 (Table III) and 22 genes (Table III) with 95% and 91% concordancy respectively for patient classification in the validation set.

[00167] Considering optimization of the 3 metagenes, we reached on the validation set a concordancy of 91% and 90% with 102 and 50 genes respectively instead of the 186 genes used in the original model.

EXAMPLE 3: Identification of a second significant metagene combination:

[00168] Since ER and EGFR markers are correlated, with the majority of EGFR+ being ER-, we found another combination that could replace the metagenes underER and underPR by a single metagene overEGFR.

[00169] Table IV (Metagene OverEGFR)

Gene	Unigene Cluster	Regulation	P value	Ref. Seq	Reduced Metagene 12	Reduced Metagene 5
GSTP1 glutathione s-transferase pi	ughs.523836:186	+	0,00005	SEQ ID No:405 (nm_000852)	-	-
ITGB3 integrin, beta 3 (platelet glycoprotein iiia, antigen cd61)	ughs.218040:186	+	0,00008	SEQ ID No:374 (nm_000212)	-	-

IGHG1 i mmunoglobulin heavy constant gamma 1 (g1m marker)	ughs.510635:186	+	0,00011	SEQ ID No:1122	+	+
SOD2 superoxide dismutase 2, mitochondrial	ughs.487046:186	+	0,00072	SEQ ID No:598 (nm_000636)	+	+
CEBPB ccaat/enhancer binding protein (c/ebp), beta	ughs.517106:186	+	0,00089	SEQ ID No:262 (nm_005194)	+	-
IGKC immunoglobulin kappa constant	ughs.449621:186; ughs.546620:186	+	0,00177	SEQ ID No:1099 (BC066343)	+	-
ENO1 enolase 1, (alpha)	ughs.517145:186	+	0,00201	SEQ ID No:696 (nm_001428)	+	+
npc-a-5 nasopharyngeal carcinoma-associated antigen npc-a-5	ughs.510543:186	+	0,00352	SEQ ID No:1059 (AK091113)	+	+
MMP7 matrix metalloproteinase 7 (matrilysin, uterine)	ughs.2256:186	+	0,00698	SEQ ID No:751 (nm_002423)	+	-
	N_A	+	0,01196	SEQ ID No:1121	+	-
MKI67 antigen identified by monoclonal antibody ki-67	ughs.80976:186	+	0,0122	SEQ ID No:286 (nm_002417)	+	-
ARHGEF1 rho guanine nucleotide exchange factor	ughs.278186:186	+	0,01427	SEQ ID No:244 (nm_199002)	-	-

(gef) 1						
ATF2 activating transcription factor 2	ughs.425104:186	+	0,0148	SEQ ID No:18 (nm_001880)	-	-
TFCP2L1 transcription factor cp2-like 1	ughs.156471:186	+	0,0259	SEQ ID No:121 (nm_014553)	+	+
IGKC Immunoglobulin kappa variable 1-5 (IGKC)	N_A	+	0,02767	SEQ ID No:1107 (BC073775)	-	-
PRSS12 protease, serine, 12 (neurotrypsin, motopsin)	ughs.445857:186	+	0,03118	SEQ ID No:103 (nm_003619)	+	-
IGLC2 immunoglobulin lambda joining 3	ughs.449585:186	+	0,04077	SEQ ID No:1118	+	-
CSF1 colony stimulating factor 1 (macrophage)	ughs.173894:186	+	0,0412	SEQ ID No:42 (nm_000757)	-	-
LOC114659 SH3-domain GRB2-like pseudogene 1	ughs.406166:186; ughs.438861:186	+	0,04453	SEQ ID No:1067 (AK123784)	-	-

[00170] Multivariate Cox analysis allowed estimation of parameters corresponding to each of the selected metagenes:

Metagene	Parameter estimation	P value (Chi square)	Hazard Ratio
OverEGFR	-1.33	0.022	0.26
UnderEGFR	-2.28	0.0048	0.10

[00171] On the basis of these parameters, the score for prognosis has been established as follows:

[00172] Score = -1.33*overEGFR – 2.28*underEGFR

[00173] Threshold optimization: the 3rd quartile was selected (cut-off = 0.14) associated with a [0.103 – 0.177] confidence interval, separating patients into two groups of 79% 5years MFS in the good prognosis group and 60% of 5years MFS in the poor prognosis group (p=0.041, logrank test).

[00174] Model validation: we calculated the score for the 164 patients of the validation set with the formula identified on the training set, and separated the patients according to the defined threshold. The model was well validated ($p=1.1\ 10^{-03}$, log rank test), with 82% MFS at 5 years in the good prognosis group (76% of patients), and 54% MFS in the poor prognosis group (20% of patients, 5% of patients not interpretable). On a subset of the validation set, constituted of the clinical trial PACS01 (N=128), we obtained similar validation ($p= 2.9\ 10^{-03}$, logrank test) with 87% of 5-year MFS in the good-prognosis group (75% of patients) and 60% of 5-year MFS in the poor-prognosis group (19% of patients, 6% of patients not interpretable).

[00175] Model performances: we performed multivariate analysis to determine the importance of the model as previously.

[00176] Multivariate analysis on the global population (N=347)

		Hazard ratio	CI95 upper	CI95 lower	p
Age	<35 y	1			
	>=35y	0.73	0.17	3.09	p=0.67
Menopausal	N	1			
	Y	1.27	0.83	1.93	p=0.27
Tumour size	pT1	1			
	pT2-pT3	1.65	0.97	2.79	p=0.065
N	N-	1			
	N+	1.39	0.73	2.64	p=0.32
SBR grade	I	1			
	II-III	2.46	1.05	5.78	p=0.039
HR (10%)	HR-	1			
	HR+	0.8	0.48	1.33	p=0.4
Erbb2	0-1-2	1			
	3	1.05	0.59	1.85	p=0.88
Model	Good	1			
	Poor	1.79	1.09	2.94	P=0.021

[00177] Multivariate analysis on the training set (N=222)

		Hazard ratio	CI95 upper	CI95 lower	p
Age	<35 y	1			
	>=35y	1.67	0.22	12.59	p=0.62
Menopausal	N	1			
	Y	1.01	0.63	1.65	p=0.95
Tumour size	pT1	1			
	pT2-pT3	1.11	0.63	1.97	p=0.71
N	N-	1			
	N+	1.67	0.86	3.21	p=0.13
SBR grade	I	1			

	II-III	2.27	0.95	5.46	p=0.067
HR (10%)	HR- HR+	1 0.87	0.48	1.57	p=0.65
Erbb2	0-1-2 3	1 0.85	0.4	1.77	p=0.66
Model	Good Poor	1 1.32	0.73	2.38	P=0.35

[00178] Multivariate analysis on the PACS01 clinical trial (N=108)

		Hazard ratio	CI95 upper	CI95 lower	p
Age	<35 y >=35y	1 440091063.3	0	Inf	p=1
Menopausal	N Y	1 1.59	0.62	4.12	p=0.34
Tumour size	pT1 pT2-pT3	1 267234385.6	0	Inf	p=1
N	N- N+	1			p=NA
SBR grade	I II-III	1 182875754.1	0	Inf	p=1
HR (10%)	HR- HR+	1 0.95	0.26	3.5	p=0.94
Erbb2	0-1-2 3	1 1.99	0.65	6.1	p=0.23
Model	Good Poor	1 3.09	0.96	10.02	P=0.059

[00179] Metagenes reduction:

[00180] We optimized the number of genes to analyse in underEGFR and overEGFR signature as described previously for the other metagenes.

[00181] The metagene overEGFR could be reduced from 19 to 12 (Table IV) or 5 genes (Table IV) with a concordancy of 96% and 94% respectively on the validation set.

[00182] Taken with the optimized underEGFR metagene, we obtained a concordancy of 95 and 91% considering 37 (Table III) and 24 genes (Table III) respectively instead of 92.

[00183] Some metagenes could be reduced at the level of a single gene still having a significant prognostic value.

[00184] An example of such a gene-based model contains SCUBE2 (SEQ ID NO: 681) and IGKC (SEQ ID NO: 1107 or 1099). SCUBE2 is an element of underEGFR metagene, while IGKC is part of overEGFR metagene.

Metagene	Parameter estimation	P value (Chi square)	Hazard Ratio
SCUBE2	-0.746	0.0016	0.474
IGKC	-0.463	0.037	0.629

[00185] Threshold optimization: the 3rd quartile (cut-off=0.095), confidence interval [0.0513 – 0.1387]) was the most significant ($p=9.1\ 10^{-04}$, logrank test) and separated the identification set in a good-prognosis group (77% MFS at 5 years) and a poor-prognosis group (51% MFS at 5 years).

[00186] Model Validation: we used the coefficients and the threshold previously calculated to separate the 164 patients from the validation set into two groups that had statistically significant outcome ($p=4\ 10^{-04}$, logrank test). The good prognosis group had a 5y MFS of 83% (69% of the patients) while the poor prognosis group had a 5y MFS of 55% (24% of the patients, 7% of patients not interpretable). On a subset of the validation set, constituted of the clinical trial PACS01 (N=128), we obtained similar validation ($p= 1.3\ 10^{-03}$, logrank test) with 90% of 5-year MFS in the good-prognosis group (69% of patients) and 61% of 5-year MFS in the poor-prognosis group (23% of patients, 7% of patients not interpretable).

[00187] Model performances: we performed multivariate analysis to determine the importance of this simplified model as described previously.

[00188] Multivariate analysis on the global population (N=330)

		Hazard ratio	CI95 upper	CI95 lower	p
Age	<35 y	1			
	>=35y	0.91	0.22	3.77	p=0.89
Menopausal	N	1			
	Y	1.2	0.78	1.85	p=0.4
Tumour size	pT1	1			
	pT2-pT3	1.61	0.95	2.74	p=0.079
N	N-	1			
	N+	1.38	0.72	2.64	p=0.33
SBR grade	I	1			
	II-III	2.36	1	5.56	p=0.051
HR (10%)	HR-	1			
	HR+	0.71	0.44	1.13	p=0.15
Erbb2	0-1-2	1			
	3	1.09	0.62	1.92	p=0.76
Model	Good	1			
	Poor	1.82	1.17	2.82	P=0.0077

[00189] Multivariate analysis on the training set (N=222)

		Hazard ratio	CI95 upper	CI95 lower	p
Age	<35 y	1			
	>=35y	1.59	0.21	11.94	p=0.65
Menopausal	N	1			
	Y	0.99	0.61	1.6	p=0.97

Tumour size	pT1 pT2-pT3	1 1.1	0.62	1.94	p=0.75
N	N- N+	1 1.64	0.85	3.17	p=0.14
SBR grade	I II-III	1 2.11	0.87	5.12	p=0.098
HR (10%)	HR- HR+	1 0.91	0.52	1.59	p=0.74
Erbb2	0-1-2 3	1 0.91	0.44	1.89	p=0.8
Model	Good Poor	1 1.74	1.02	2.99	P=0.043

[00190] Multivariate analysis on the PACS01 clinical trial (N=108)

		Hazard ratio	CI95 upper	CI95 lower	p
Age	<35 y >=35y	1 0.73	0.09	6.22	p=0.77
Menopausal	N Y	1 1.01	0.37	2.73	p=0.99
Tumour size	pT1 pT2-pT3	1 6.19	0.79	48.7	p=0.083
N	N- N+	1			p=NA
SBR grade	I II-III	1 634794463.76	0	Inf	p=1
HR (10%)	HR- HR+	1 0.6	0.23	1.58	p=0.3
Erbb2	0-1-2 3	1 2.32	0.8	6.74	p=0.12

Model	Good Poor	1 2.47	1.01	6.06	P=0.049

[00191] Different nucleic acids array platforms may be used to work the present invention including, but not limited to, cDNA platforms (Image or "Ipso" clones described below), Affymetrix® platforms (GeneChip® probe sets) and others.

EXAMPLE 4: Use of Metagenes combinations according to the invention on a cDNA platform

[00192] The following tables are examples of metagenes of the invention that may be used on a cDNA platform according to the above described methods. For example, the following underER, underPR and underEGFR metagenes may be used in the above described method using a Cox regression analysis and the score S_C = - 2.90279 x underER – 1.47423 x underPR – 4.17198 x under EGFR, with the intervals mentioned previously in the description for "a", "b" and "c" (and similarly for the above described combination involving underEGFR and over EGFR, as well as the IGKC + SCUBE2 combination). The Seq3' and Seq5' in the tables below columns provide the sequences identifying the respective Image or Ipso clones.

[00193] Table V : Metagene UnderER

Set No.	Gene symbol	Clone ID	Gene name	Unigene Cluster	Regulation	P value	Seq 3'	Seq 5'	Ref. Seq
402	ITGB3	ipso:0000143	integrin, beta 3 (platelet glycoprotein iiia, antigen cd61)	ughs.218040:186	-	0,00001		SEQ ID No:992	SEQ ID No: 374 (nm_000212)
423	PADI2	ipso:0000610	peptidyl arginine deiminase, type ii	ughs.33455:186	-	0,00001		SEQ ID No:1026	SEQ ID No:1027 (nm_007365)
246	SOD2	image:324014	superoxide dismutase 2, mitochondrial	ughs.487046:186	-	0,00001	SEQ ID No:597		SEQ ID No:598 (nm_000636)
290	FLJ13154	image:43457	hypothetical protein flj13154	ughs.408702:186	-	0,00003	SEQ ID No:715	SEQ ID No:716	SEQ ID No:717 (nm_024598)

237	HDAC2	image:309924	histone deacetylase 2	ughs.3352:1 86	-	0,00004	SEQ ID No:571	SEQ ID No:572	SEQ ID No:573 (nm_001527)
34	SLAC2-B	image:142546	slac2-b	N_A	-	0,00006		SEQ ID No:82	SEQ ID No:83 (nm_015065)
6	S100A8	image:1089513	s100 calcium binding protein a8 (calgranulin a)	ughs.41607 3:186	-	0,00006	SEQ ID No:11		SEQ ID No:12 (nm_002964)
171	GSTP1	image:231424	glutathione s-transferase pi	ughs.52383 6:186	-	0,00006	SEQ ID No:403	SEQ ID No:404	SEQ ID No:405 (nm_000852)
343	LCN2	image:544683	lipocalin 2 (oncogene 24p3)	ughs.20423 8:186	-	0,00012	SEQ ID No:854	SEQ ID No:855	SEQ ID No:856 (nm_005564)
163	MYBL2	image:207378	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	ughs.17971 8:186	-	0,00013	SEQ ID No:382	SEQ ID No:383	SEQ ID No:384 (nm_002466)

69	PFKP	image:152714	phosphofructokinase, platelet	ughs.26010:186	-	0,00081	SEQ ID No:165	SEQ ID No:166	SEQ ID No:167 (nm_002627)
152	STK6	image:1912132	serine/threonine kinase 6	ughs.250822:186	-	0,00134	SEQ ID No:358		SEQ ID No:51 (nm_198433)
408	GPR125	ipso:0000267	g protein-coupled receptor 125	ughs.99195:186	-	0,00153		SEQ ID No:1001	SEQ ID No:999 (nm_145290)
393	DSCR1	ipso:0000077	down syndrome critical region gene 1	ughs.282326:186	-	0,00206		SEQ ID No:978	SEQ ID No:979 (nm_004414)
1	FAT	image:1028762	fat tumor suppressor homolog 1 (drosophila)	ughs.481371:186	-	0,0023	SEQ ID No:1		SEQ ID No:2 (nm_005245)
40	VGLL1	image:143622	vestigial like 1 (drosophila)	N_A	-	0,00247	SEQ ID No:96	SEQ ID No:97	SEQ ID No:98 (nm_016267)
302	MMP7	image:471134	matrix metalloproteinase 7 (matrilysin, uterine)	ughs.2256:186	-	0,00264	SEQ ID No:749	SEQ ID No:750	SEQ ID No:751 (nm_002423)

282	ENO1	image:392678	enolase 1, (alpha)	ughs.51714 5:186	-	0,00348	SEQ ID No:695	SEQ ID No:696 (nm_001428)
59		image:1493187	cdna clone image:4831215	ughs.17528 5:186	-	0,00429	SEQ ID No:142	SEQ ID No:1050 (BC034638)
203	SCP2	image:278490	sterol carrier protein 2	ughs.47636 5:186	-	0,00469	SEQ ID No:486	SEQ ID No:488 (nm_002979)
111	CEBPB	image:161993	ccaat/enhancer binding protein (c/ebp), beta	ughs.51710 6:186	-	0,00507	SEQ ID No:261	SEQ ID No:262 (nm_005194)
419	TGM1	ipso:0000488	transglutaminase 1 (k polypeptide epidermal type i, protein-glutamine-gamma-glutamyltransferase)	ughs.50895 0:186	-	0,00695	SEQ ID No:1019	SEQ ID No:1020 (nm_000359)
418		ipso:0000487		N_A	-	0,00764	SEQ ID No:1018	SEQ ID No:1106 (BC015969)

380	GGH	image:809588	gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	ughs.78619:186	-	0,00881	SEQ ID No:950	SEQ ID No:951	SEQ ID No:952 (nm_003878)
273	GSTA4	image:345309	glutathione s-transferase a4	ughs.48555 7:186	-	0,00995	SEQ ID No:673	SEQ ID No:674	SEQ ID No:675 (nm_001512)
123	FN5	image:171580	b-cell cll/lymphoma 7b	ughs.43806 4:186	-	0,0109	SEQ ID No:287	SEQ ID No:288	SEQ ID No:289 (nm_020179)
387	CCNB2	image:845594	glutamate decarboxylase 1 (gad 1)	ughs.19469 8:186	-	0,01221	SEQ ID No:969		SEQ ID No:553 (nm_004701)
239	CTSC	image:320656	cathepsin c	ughs.12806 5:186	-	0,01501	SEQ ID No:577	SEQ ID No:578	SEQ ID No:579 (nm_001814)
305	PBEF1	image:488548	pre-b-cell colony enhancing factor 1	ughs.48961 5:186	-	0,01621	SEQ ID No:758	SEQ ID No:759	SEQ ID No:760 (nm_005746)

323	S100A6	image:512420	s100 calcium binding protein a6 (calcyclin)	ughs.27524 3:186	-	0,01719	SEQ ID No:804		SEQ ID No:805 (nm_014624)
153	RDX	image:193081	radixin	ughs.26367 1:186	-	0,01753	SEQ ID No:359	SEQ ID No:360	SEQ ID No:361 (nm_002906)
187	GPR126	image:259884	g protein-coupled receptor 126	ughs.31889 4:186	-	0,01886	SEQ ID No:446	SEQ ID No:447	SEQ ID No:448 (nm_198569)
70	MMP15	image:152744	matrix metalloproteinase 15 (membrane-inserted)	ughs.80343: 186	-	0,0274	SEQ ID No:168	SEQ ID No:169	SEQ ID No:170 (nm_002428)
352	KLK6	image:724109	kallikrein 6 (neurosin, zyme)	ughs.79361: 186	-	0,02892	SEQ ID No:876	SEQ ID No:877	SEQ ID No:878 (nm_002774)
79		image:153978		N_A	-	0,0351	SEQ ID No:189	SEQ ID No:190	SEQ ID No:1117
251	BOK	image:325789	bcl2-related ovarian killer	ughs.29375 3:186	-	0,03747	SEQ ID No:611		SEQ ID No:612 (nm_032515)

225	CDKL5	image:301018	cyclin-dependent kinase-like 5	ughs.43557 0:186	-	0,03754	SEQ ID No:538	SEQ ID No:539	SEQ ID No:540 (nm_003159)
330	CSTB	image:51814	cystatin b (stefin b)	ughs.695:18 6	-	0,0382	SEQ ID No:821	SEQ ID No:822	SEQ ID No:823 (nm_000100)
54	LOC15 1194	image:147707	similar to hepatocellular carcinoma-associated antigen hca557b	ughs.55261 0:186	-	0,03884		SEQ ID No:130	SEQ ID No:131 (nm_145280)
285	NFIB	image:416959	nuclear factor i/b	ughs.37035 9:186	-	0,03949	SEQ ID No:703	SEQ ID No:704	SEQ ID No:705 (nm_005596)
14	LAD1	image:121551	ladinin 1	ughs.51903 5:186	-	0,04184	SEQ ID No:29	SEQ ID No:30	SEQ ID No:31 (nm_005558)
83	MGC11 271	image:154651	hypothetical protein mgc11271	ughs.14328 8:186	-	0,04312		SEQ ID No:198	SEQ ID No:199 (nm_024323)

Table VI : Metagene underPR

Set No.	Gene symbol	Clone ID	Gene name	Unigene Cluster	Regulation	P value	Seq3'	Seq5'	Ref. Seq
246	SOD2	image:324014	superoxide dismutase 2, mitochondrial	ughs.48704 6:186	-	1E-05	SEQ ID No:597		SEQ ID No:598 (nm_000636)
217	IGHG1	image:289337	immunoglobulin heavy constant gamma 1 (g1m marker)	ughs.51063 5:186	-	1E-05	SEQ ID No:520	SEQ ID No:521	SEQ ID No:1122
154	KDR	image:193857	kinase insert domain receptor (a type iii receptor tyrosine kinase)	ughs.47975 6:186	-	0,0001	SEQ ID No:362	SEQ ID No:363	SEQ ID No:364 (nm_002253)
164	KLF1	image:208991	kruppel-like factor 1 (erythroid)	ughs.37860: 186	-	0,0001	SEQ ID No:385	SEQ ID No:386	SEQ ID No:387 (nm_006563)
15	CASP9	image:121693	caspase 9, apoptosis-related cysteine protease	ughs.32950 2:186	-	0,0002	SEQ ID No:32	SEQ ID No:33	SEQ ID No:34 (nm_001229)

267	BCL2	image:342181	b-cell cll/lymphoma 2	ughs.150749:186	-	0,0002	SEQ ID No:655	SEQ ID No:656	SEQ ID No:657 (nm_000633)
163	MYBL2	image:207378	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	ughs.179718:186	-	0,0003	SEQ ID No:382	SEQ ID No:383	SEQ ID No:384 (nm_002466)
188	ADAM10	image:261401	a disintegrin and metalloproteinase domain 10	ughs.172028:186	-	0,0003	SEQ ID No:449	SEQ ID No:450	SEQ ID No:451 (nm_001110)
406	GPR125	ipso:0000252	g protein-coupled receptor 125	ughs.99195:186	-	0,0003		SEQ ID No:998	SEQ ID No:999 (nm_145290)
81		image:154483		ughs.26192:186	-	0,0005	SEQ ID No:194	SEQ ID No:1055	SEQ ID No:1056 (AK126297)
7	TGFBR3	image:110287	transforming growth factor, beta receptor iii (betaglycan, 300kda)	ughs.482390:186	-	0,0006	SEQ ID No:13	SEQ ID No:14	SEQ ID No:15 (nm_003243)

318	LOC91316	image:50877	similar to bk246h3.1 (immunoglobulin lambda-like polypeptide 1, pre-b-cell specific)	ughs.40769 3:186;ughs. 148656:186	-	0,0007	SEQ ID No:792	SEQ ID No:1088 or SEQ ID No:1089	SEQ ID No:1090 (AK125808)
95		image:156715		ughs.41613 9:186	-	0,0007	SEQ ID No:227	SEQ ID No:1060	SEQ ID No:1120
6	S100A8	image:1089513	s100 calcium binding protein a8 (calgranulin a)	ughs.41607 3:186	-	0,0008	SEQ ID No:11		SEQ ID No:12 (nm_002964)
299	PIM2	image:46959	pim-2 oncogene	ughs.49609 6:186	-	0,0009	SEQ ID No:741	SEQ ID No:742	SEQ ID No:743 (nm_006875)
175	TP53	image:236338	tumor protein p53 (li-fraumeni syndrome)	ughs.40831 2:186	-	0,001	SEQ ID No:412	SEQ ID No:413	SEQ ID No:414 (nm_000546)
404	ITGB3	ipso:0000152	integrin, beta 3 (platelet glycoprotein iiia, antigen cd61)	ughs.21804 0:186	-	0,0012		SEQ ID No:995	SEQ ID No:374 (nm_000212)

287	LAMB1	image:428443	laminin, beta 1	ughs.48964 6:186	-	0,0012	SEQ ID No:709	SEQ ID No:710	SEQ ID No:711 (nm_002291)
269	SILV	image:342383	silver homolog (mouse)	ughs.95972: 186	-	0,0012	SEQ ID No:661	SEQ ID No:662	SEQ ID No:663 (nm_006928)
392		ipso:0000040	cdna fij42596 fis, clone brace3010283	ughs.11327 1:186	-	0,0012		SEQ ID No:977	SEQ ID No:1102 (AK124587)
100	PIGR	image:159410	polymeric immunoglobulin receptor	ughs.49758 9:186	-	0,0012	SEQ ID No:236		SEQ ID No:237 (nm_002644)
25	CSH1	image:133891	chorionic somatomammotropin hormone 1 (placental lactogen)	ughs.34796 3:186	-	0,0016		SEQ ID No:59	SEQ ID No:60 (nm_022640)
153	RDX	image:193081	radixin	ughs.26367 1:186	-	0,0018	SEQ ID No:359	SEQ ID No:360	SEQ ID No:361 (nm_002906)

49	ETF1/FLT1	image:146976	eukaryotic translation termination factor 1/fms-related tyrosine kinase 1	ughs.48349 4:186;ughs. 507621:186	-	0,0019	SEQ ID No:118	SEQ ID No:119 (nm_004730) or SEQ ID No:1109 (NM_002019)
69	PFKP	image:152714	phosphofructokinase, platelet	ughs.26010: 186	-	0,0019	SEQ ID No:165	SEQ ID No:166	SEQ ID No:167 (nm_002627)
143	CXORF38	image:188005	chromosome x open reading frame 38	ughs.49596 1:186	-	0,002	SEQ ID No:337	SEQ ID No:338	SEQ ID No:339 (nm_144970)
140	MGC15606	image:187120	family with sequence similarity 55, member c	ughs.13019 5:186	-	0,0021	SEQ ID No:331	SEQ ID No:332	SEQ ID No:333 (nm_145037)
402	ITGB3	ipso:0000143	integrin, beta 3 (platelet glycoprotein iiia, antigen cd61)	ughs.21804 0:186	-	0,0022		SEQ ID No:992	SEQ ID No:374 (nm_000212)
34	SLAC2-B	image:142546	slac2-b	N_A	-	0,0024		SEQ ID No:82	SEQ ID No:83 (nm_015065)

139	FLJ10986	image:187119	hypothetical protein flj10986	ughs.44430 1:186;ughs. 439112:186	-	0,0026	SEQ ID No:328	SEQ ID No:329	SEQ ID No:330 (nm_018291)
421	SERPINB1	ipso:0000605	serine (or cysteine) proteinase inhibitor, clade b (ovalbumin), member 1	ughs.38116 7:186	-	0,0037		SEQ ID No:1023	SEQ ID No:1024 (nm_030666)
96	RPS6KA3	image:156808	ribosomal protein s6 kinase, 90kda, polypeptide 3	ughs.44538 7:186	-	0,0048		SEQ ID No:228	SEQ ID No:229 (nm_004586)
370	GATA6	image:771332	gata binding protein 6	ughs.51474 6:186	-	0,0049	SEQ ID No:923	SEQ ID No:924	SEQ ID No:925 (nm_005257)
316	MTIF2	image:50754	mitochondrial translational initiation factor 2	ughs.14989 4:186	-	0,0054	SEQ ID No:787		SEQ ID No:788 (nm_001005369)
413		ipso:0000376		N_A	-	0,0057		SEQ ID No:1010	SEQ ID No:1104 (AK128524)
397		ipso:0000119		N_A	-	0,0064		SEQ ID No:985	SEQ ID No:1103 (BX108410)

27	IFNGR 1	image:136478	interferon gamma receptor 1	ughs.52041 4:186	-	0,0066	SEQ ID No:64	SEQ ID No:65	SEQ ID No:66 (nm_000416)
425	EBF	ipso:0000617	early b-cell factor	ughs.30804 8:186	-	0,0067		SEQ ID No:1029	SEQ ID No:1030 (nm_024007)
92		image:156283		N_A	-	0,0073	SEQ ID No:221	SEQ ID No:222	SEQ ID No:1119.
148	p66alp ha	image:188422	GATA zinc finger domain containing 2A (p66alpha)	ughs.55174 2:186	-	0,0074	SEQ ID No:350		SEQ ID No:1068 (AK024670)
102	FKBP1 A	image:159521	fk506 binding protein 1a, 12kda	ughs.47193 3:186	-	0,0089	SEQ ID No:239	SEQ ID No:240	SEQ ID No:241 (nm_000801)
168	SNAPC 3	image:219829	small nuclear rna activating complex, polypeptide 3, 50kda	ughs.54629 9:186	-	0,0089	SEQ ID No:396	SEQ ID No:397	SEQ ID No:398 (nm_003084)
159	ITGB3	image:200209	integrin, beta 3 (platelet glycoprotein iiia, antigen cd61)	ughs.21804 0:186	-	0,0097	SEQ ID No:373		SEQ ID No:374 (nm_000212)

30	IL2RB	image:139073	interleukin 2 receptor, beta	ughs.47478 7:186;ughs. 555488:186	-	0,0097	SEQ ID No:72	SEQ ID No:73	SEQ ID No:74 (nm_000878)
313		image:50541	Homo sapiens mRNA for FLJ00204 protein	ughs.53515 7:186	-	0,0097	SEQ ID No:779	SEQ ID No:780	SEQ ID No:1087 (AK074131)
381	ETV4	image:809959	ets variant gene 4 (e1a enhancer binding protein, e1af)	ughs.43405 9:186	-	0,01	SEQ ID No:953	SEQ ID No:954	SEQ ID No:955 (nm_001986)
29	IL1R2	image:137575	interleukin 1 receptor, type ii	ughs.25333: 186	-	0,0101	SEQ ID No:69	SEQ ID No:70	SEQ ID No:71 (nm_004633)
416	IGHG1	ipso:0000434	immunoglobulin heavy constant gamma 1 (g1m marker)	ughs.51063 5:186	-	0,0104		SEQ ID No:1015	SEQ ID No:1105 (BC072392)
343	LCN2	image:544683	lipocalin 2 (oncogene 24p3)	ughs.20423 8:186	-	0,0107	SEQ ID No:854	SEQ ID No:855	SEQ ID No:856 (nm_005564)
97	CMRF3 5	image:156937	cd300c antigen	ughs.2605:1 86	-	0,0112		SEQ ID No:230	SEQ ID No:231 (nm_006678)

244	CXCL1	image:323238	chemokine (c-x-c motif) ligand 1 (melanoma growth stimulating activity, alpha)	ughs.789:186	-	0,0117	SEQ ID No:591	SEQ ID No:592	SEQ ID No:593 (nm_001511)
353	MYBL2	image:724259	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	ughs.17971 8:186	-	0,0122	SEQ ID No:879	SEQ ID No:880	SEQ ID No:384 (nm_002466)
216	SLAMF 8	image:288807	slam family member 8	ughs.43868 3:186	-	0,0131	SEQ ID No:517	SEQ ID No:518	SEQ ID No:519 (nm_020125)
239	CTSC	image:320656	cathepsin c	ughs.12806 5:186	-	0,016	SEQ ID No:577	SEQ ID No:578	SEQ ID No:579 (nm_001814)
430	ENPP2	ipso:0000727	ectonucleotide pyrophosphatase/phosphodiesterase 2 (autotaxin)	ughs.19097 7:186	-	0,0205		SEQ ID No:1038	SEQ ID No:1039 (nm_006209)
14	LAD1	image:121551	ladinin 1	ughs.51903 5:186	-	0,021	SEQ ID No:29	SEQ ID No:30	SEQ ID No:31 (nm_005558)

138	RABL3	image:186926	rab, member of ras oncogene family-like 3	ughs.44436 0:186;ughs. 548087:186	-	0,0221	SEQ ID No:326	SEQ ID No:327 (nm_173825)
237	HDAC2	image:309924	histone deacetylase 2	ughs.3352:1 86	-	0,0243	SEQ ID No:571	SEQ ID No:572	SEQ ID No:573 (nm_001527)
40	VGLL1	image:143622	vestigial like 1 (drosophila)	N_A	-	0,0245	SEQ ID No:96	SEQ ID No:97	SEQ ID No:98 (nm_016267)
94	npc-a-5	image:156691	nasopharyngeal carcinoma-associated antigen npc-a-5	ughs.51054 3:186	-	0,0259	SEQ ID No:226	SEQ ID No:1058	SEQ ID No:1059 (AK091113)
355	CDK4	image:725349	cyclin-dependent kinase 4	ughs.95577: 186	-	0,0262	SEQ ID No:884	SEQ ID No:885	SEQ ID No:886 (nm_000075)
426	ABCC5	ipso:0000654	atp-binding cassette, sub-family c (cftr/mrp), member 5	ughs.36856 3:186	-	0,0262		SEQ ID No:1031	SEQ ID No:1032 (nm_005688)
319	MGC99 13	image:50892	hypothetical protein mgc9913	ughs.23133: 186	-	0,0271	SEQ ID No:793	SEQ ID No:794	SEQ ID No:1091 (XM_378178)

98	FUT8	image:156966	fucosyltransferase 8 (alpha (1,6) fucosyltransferase)	ughs.11872 2:186	-	0,0283	SEQ ID No:232	SEQ ID No:233 (nm_178155)
375	SFRP1	image:783700	secreted frizzled-related protein 1	ughs.21342 4:186	-	0,0301	SEQ ID No:937	SEQ ID No:938 (nm_003012)
112	ARPC2	image:162208	actin related protein 2/3 complex, subunit 2, 34kda	ughs.52930 3:186	-	0,0324	SEQ ID No:263	SEQ ID No:264 (nm_152862)
227	LILRB2	image:30470	leukocyte immunoglobulin-like receptor, subfamily b (with tm and itim domains), member 2	ughs.53438 6:186	-	0,0329	SEQ ID No:544	SEQ ID No:545	SEQ ID No:546 (nm_005874)
350	IGKC	image:713852	immunoglobulin kappa constant	ughs.44962 1:186;ughs. 546620:186	-	0,0346	SEQ ID No:872	SEQ ID No:1097 or SEQ ID No:1098	SEQ ID No:1099 (BC066343)
429	SN	ipso:0000704	sialoadhesin	ughs.31869: 186	-	0,0377		SEQ ID No:1036	SEQ ID No:1037 (nm_023068)

229	C1ORF 38	image:307255	chromosome 1 open reading frame 38	ughs.10649: 186	-	0,0378	SEQ ID No:548	SEQ ID No:549	SEQ ID No:550 (nm_004848)
423	PADI2	ipso:0000610	peptidyl arginine deiminase, type ii	ughs.33455: 186	-	0,0418		SEQ ID No:1026	SEQ ID No:1027 (nm_007365)
410	MOND OA	ipso:0000314	mlx interactor	ughs.43715 3:186	-	0,0455		SEQ ID No:1004	SEQ ID No:1005 (nm_014938)
329	TAP1	image:51782	transporter 1, atp-binding cassette, sub-family b (mdr/tap)	ughs.35201 8:186;ughs. 552165:186	-	0,0458	SEQ ID No:818	SEQ ID No:819	SEQ ID No:820 (nm_000593)
157	CYP2D 6	image:199680	cytochrome p450, family 2, subfamily d, polypeptide 6	ughs.53431 1:186	-	0,047		SEQ ID No:369	SEQ ID No:370 (nm_000106)

Table VII : Metagene underEGFR

Set No.	Gene symbol	Clone ID	Gene name	Unigene Cluster	Regulation	P value	Seq3'	Seq5'	Ref. Seq
197		image:266500	LOC255743: Nephronectin	N_A	-	1E-05	SEQ ID No:472		SEQ ID No:1071 (NM_001033047)
107	LU	image:160656	lutheran blood group (auberger b antigen included)	ughs.15504 8:186	-	1E-05	SEQ ID No:253		SEQ ID No:254 (nm_005581)
3	TFF1	image:1075949	trefoil factor 1 (breast cancer, estrogen-inducible sequence expressed in)	ughs.16280 7:186	-	1E-05	SEQ ID No:5		SEQ ID No:6 (nm_003225)
354	ESR1	image:725321	estrogen receptor 1	ughs.20812 4:186	-	1E-05	SEQ ID No:881	SEQ ID No:882	SEQ ID No:883 (nm_000125)

226	XBP1	image:301950	x-box binding protein 1	ughs.43763 8:186	-	1E-05	SEQ ID No:541	SEQ ID No:542	SEQ ID No:543 (nm_005080)
275	SCUBE 2	image:346321	signal peptide, cub domain, egf-like 2	ughs.52346 8:186	-	1E-05	SEQ ID No:679	SEQ ID No:680	SEQ ID No:681 (nm_020974)
26	GATA3	image:135118	gata binding protein 3	ughs.52413 4:186	-	1E-05	SEQ ID No:61	SEQ ID No:62	SEQ ID No:63 (nm_001002295)
31	GATA3	image:139076	gata binding protein 3	ughs.52413 4:186	-	1E-05	SEQ ID No:75	SEQ ID No:76	SEQ ID No:63 (nm_001002295)
88	EIF2C3	image:155341	eukaryotic translation initiation factor 2c, 3	ughs.53033 3:186	-	1E-05		SEQ ID No:211	SEQ ID No:212 (nm_024852)
309	C4A	image:491004	complement component 4b, telomeric	ughs.53484 7:186	-	1E-05	SEQ ID No:770	SEQ ID No:771	SEQ ID No:635 (nm_001002029)
223	TFF3	image:298417	trefoil factor 3 (intestinal)	ughs.82961: 186	-	1E-05	SEQ ID No:534		SEQ ID No:535 (nm_003226)
424		ipso:0000614		N_A	-	3E-05		SEQ ID No:1028	SEQ ID No:1125

44	NAT1	image:145894	n-acetyltransferase 1 (arylamine n-acetyltransferase)	ughs.15595 6:186	-	3E-05	SEQ ID No:107	SEQ ID No:108	SEQ ID No:109 (nm_000662)
144	COL4A 2	image:188193	collagen, type iv, alpha 2	ughs.50871 6:186	-	3E-05	SEQ ID No:340	SEQ ID No:341	SEQ ID No:342 (nm_001846)
259	C4A	image:340753	complement component 4b, telomeric	ughs.53484 7:186	-	3E-05	SEQ ID No:633	SEQ ID No:634	SEQ ID No:635 (nm_001002029)
371	RABEP 1	image:772890	rabaptin, rab gtpase binding effector protein 1	ughs.55151 8:186	-	3E-05	SEQ ID No:926		SEQ ID No:927 (nm_004703)
398		ipso:0000125		N_A	-	5E-05		SEQ ID No:986	SEQ ID No:1124
51	RHOBT B3	image:147138	rho-related btb domain containing 3	ughs.44503 0:186	-	6E-05	SEQ ID No:122	SEQ ID No:123	SEQ ID No:124 (nm_014899)

119	CASKIN1/flj12650	image:166862	cask interacting protein 1	ughs.53086 3:186;ughs.470259:186	-	6E-05	SEQ ID No:279	SEQ ID No:280 (nm_020764) or SEQ ID No:1110 (NM_024522)
126	CXXC5	image:173797	cxxc finger 5	ughs.18911 9:186	-	9E-05	SEQ ID No:295	SEQ ID No:296	SEQ ID No:297 (nm_016463)
317	MAPT	image:50764	microtubule-associated protein tau	ughs.10117 4:186	-	0,0001	SEQ ID No:789	SEQ ID No:790	SEQ ID No:791 (nm_016835)
87	MGC24047	image:155072	chromosome 1 open reading frame 64	ughs.29190: 186	-	0,0001	SEQ ID No:208	SEQ ID No:209	SEQ ID No:210 (nm_178840)
332	MGC45441	image:52118	hypothetical protein mgc45441	ughs.48833 7:186	-	0,0003	SEQ ID No:826		SEQ ID No:827 (nm_152499)
135	CYP2B6	image:182295	Cytochrome P450, family 2, subfamily B, polypeptide 6	N_A	-	0,0007	SEQ ID No:319	SEQ ID No:320	SEQ ID No:1064 (NM_000767)

61	CROCC	image:149567	ciliary rootlet coiled-coil, rootletin	ughs.30940 3:186;ughs. 135718:186	-	0,0007		SEQ ID No:146	SEQ ID No:147 (nm_014675)
136	USP21	image:183062	ubiquitin specific protease 21	ughs.8015:1 86	-	0,0008	SEQ ID No:321	SEQ ID No:322	SEQ ID No:323 (nm_001014443)
43	TRAF5	image:145410	tnf receptor-associated factor 5	ughs.52393 0:186	-	0,0011	SEQ ID No:104	SEQ ID No:105	SEQ ID No:106 (nm_004619)
75	GSTM2	image:153444	glutathione s-transferase m2 (muscle)	ughs.27983 7:186	-	0,0013		SEQ ID No:180	SEQ ID No:181 (nm_000848)
160	DUSP4	image:2044325	dual specificity phosphatase 4	ughs.41796 2:186	-	0,0015	SEQ ID No:375		SEQ ID No:376 (nm_057158)
47	ASF1A	image:146634	asf1 anti-silencing function 1 homolog a (s. cerevisiae)	ughs.29231 6:186	-	0,0018	SEQ ID No:115		SEQ ID No:116 (nm_014034)

106	CSF2	image:1601601	colony stimulating factor 2 (granulocyte-macrophage)	ughs.1349:186	-	0,0024	SEQ ID No:251	SEQ ID No:252 (nm_000758)
320	CLSTN2	image:50970	calsyntenin 2	ughs.15852 9:186	-	0,0025	SEQ ID No:795	SEQ ID No:796	SEQ ID No:797 (nm_022131)
365	GLI3	image:767495	gli-kruppel family member gli3 (greig cephalopolysyndactyly syndrome)	ughs.19933 8:186	-	0,0028	SEQ ID No:910		SEQ ID No:911 (nm_000168)
291	REPS2	image:43488	ralbp1 associated eps domain containing 2	ughs.18681 0:186;ughs. 131188:186	-	0,0031	SEQ ID No:718	SEQ ID No:719	SEQ ID No:720 (nm_004726)
356	GSTM1	image:73778	glutathione s-transferase m1	ughs.30196 1:186	-	0,0031	SEQ ID No:887	SEQ ID No:888	SEQ ID No:889 (nm_000561)
407	PLAT	ipso:0000253	plasminogen activator, tissue	ughs.49158 2:186	-	0,0034		SEQ ID No:1000	SEQ ID No:250 (nm_000930)

74	DLG5	image:153368	discs, large homolog 5 (drosophila)	ughs.500245:186	-	0,0039	SEQ ID No:178	SEQ ID No:179 (nm_004747)
315	FLJ00012	image:50602	fij00012 protein	ughs.21051:186	-	0,004	SEQ ID No:784	SEQ ID No:786 (nm_033388)
73	SIDT2	image:153205	sid1 transmembrane family, member 2	ughs.410977:186	-	0,0041	SEQ ID No:176	SEQ ID No:177 (nm_015996)
39		image:143169		N_A	-	0,0043	SEQ ID No:95	SEQ ID No:1047 (BC012900)
128	BCL9	image:1756392	b-cell cll/lymphoma 9	ughs.415209:186	-	0,0043	SEQ ID No:300	SEQ ID No:301 (nm_004326)
86	USP13	image:155064	ubiquitin specific protease 13 (isopeptidase t-3)	ughs.175322:186	-	0,0052	SEQ ID No:205	SEQ ID No:207 (nm_003940)
374	DNALI1	image:782688	dynein, axonemal, light intermediate polypeptide 1	ughs.406050:186	-	0,0061	SEQ ID No:934	SEQ ID No:936 (nm_003462)

367	FOXC1 /RHOB	image:768370	forkhead box c1/ras homolog gene	ughs.34888 3:186;ughs. 502876:186	-	0,0065		SEQ ID No:915	SEQ ID No:916 (nm_001453) or SEQ ID No:1116 (NM_004040)
62		image:149760		N_A	-	0,007	SEQ ID No:148	SEQ ID No:149	SEQ ID No:1052 (BX096026)
345	GSTM2	image:664233	glutathione s-transferase m2 (muscle)	ughs.27983 7:186	-	0,0079	SEQ ID No:859	SEQ ID No:860	SEQ ID No:181 (nm_000848)
66	KRT18	image:151663	keratin 18	ughs.40601 3:186	-	0,0088	SEQ ID No:157	SEQ ID No:158	SEQ ID No:159 (nm_000224)
340		image:52898		ughs.54804 0:186	-	0,0089	SEQ ID No:847	SEQ ID No:1094 or SEQ ID No:1095	SEQ ID No:1096 (AK127274)
13	DNAJC 12	image:120138	dnaj (hsp40) homolog, subfamily c, member 12	ughs.26072 0:186	-	0,0094	SEQ ID No:26	SEQ ID No:27	SEQ ID No:28 (nm_021800)

77		image:153617	cdna flj41270 fis, clone bramy2036387	ughs.44541 4:186	-	0,0096	SEQ ID No:184	SEQ ID No:185	SEQ ID No:1054 (AK123264)
12	SPDEF /c8orf1 3	image:1188588	sam pointed domain containing ets transcription factor / chromosome 8 open reading frame 13	ughs.12429 9:186;ughs. 485158:186	-	0,0098	SEQ ID No:23	SEQ ID No:24	SEQ ID No:25 (nm_012391) or SEQ ID No:1108 (NM_053279)
331	C20OR F23	image:52103	chromosome 20 open reading frame 23	ughs.10177 4:186	-	0,0102	SEQ ID No:824		SEQ ID No:825 (nm_024704)
60	FLJ203 66	image:149549	hypothetical protein flj20366	ughs.39073 8:186	-	0,0128	SEQ ID No:143	SEQ ID No:144	SEQ ID No:145 (nm_017786)
204	COX6C	image:278531	cytochrome c oxidase subunit vic	ughs.35187 5:186	-	0,014	SEQ ID No:489	SEQ ID No:490	SEQ ID No:491 (nm_004374)

202	RGS11	image:277917	regulator of g-protein signalling 11	ughs.65756:186	-	0,0142		SEQ ID No:484	SEQ ID No:485 (nm_003834)
206		image:280743	Hypothetical protein LOC153561	ughs.50855 9:186	-	0,0148	SEQ ID No:495	SEQ ID No:496	SEQ ID No:1072 (AY007114)
116	SEMA6 B	image:166010	sema domain, transmembrane domain (tm), and cytoplasmic domain, (semaphorin) 6b	ughs.46564 2:186	-	0,0157	SEQ ID No:272	SEQ ID No:273	SEQ ID No:274 (nm_032108)
109	AP1G2	image:161763	adaptor-related protein complex 1, gamma 2 subunit	ughs.34324 4:186	-	0,0171	SEQ ID No:256	SEQ ID No:257	SEQ ID No:258 (nm_080545)
124	AKAP8 L	image:171679	a kinase (prka) anchor protein 8-like	ughs.39980 0:186	-	0,0182	SEQ ID No:290	SEQ ID No:291	SEQ ID No:292 (nm_014371)

322	PRKCB P1	image:511899	protein kinase c binding protein 1	ughs.446240:186	-	0,0184	SEQ ID No:801	SEQ ID No:802	SEQ ID No:803 (nm_183047)
120	GSTM2	image:166910	glutathione s-transferase m2 (muscle)	ughs.279837:186	-	0,0186	SEQ ID No:281	SEQ ID No:282	SEQ ID No:181 (nm_000848)
105	PLAT	image:160149	plasminogen activator, tissue	ughs.491582:186	-	0,0196	SEQ ID No:248	SEQ ID No:249	SEQ ID No:250 (nm_000930)
147	CENTG 3	image:188414	centaurin, gamma 3	ughs.195048:186	-	0,0205	SEQ ID No:347	SEQ ID No:348	SEQ ID No:349 (nm_031946)
339		image:52870	genomic region on chromosome 1	ughs.159853:186	-	0,0246	SEQ ID No:1093	SEQ ID No:846 or SEQ ID No:1092	SEQ ID No:1123
306	SLC40 A1	image:489218	solute carrier family 40 (iron-regulated transporter), member 1	ughs.529285:186	-	0,0246	SEQ ID No:761	SEQ ID No:762	SEQ ID No:763 (nm_014585)

183	CCND2	image:249688	cyclin d2	ughs.37607 1:186	-	0,0272	SEQ ID No:436	SEQ ID No:437	SEQ ID No:438 (nm_001759)
38	KLHDC 2	image:143060	kelch domain containing 2	N_A	-	0,028	SEQ ID No:92	SEQ ID No:93	SEQ ID No:94 (nm_014315)
338	ABCA3	image:52741	atp-binding cassette, sub-family a (abc1), member 3	ughs.26630: 186	-	0,0344	SEQ ID No:843	SEQ ID No:844	SEQ ID No:845 (nm_001089)
293	LOC14 3381	image:44338	hypothetical protein loc143381	ughs.38834 7:186;ughs. 557061:186	-	0,0371	SEQ ID No:724	SEQ ID No:725	SEQ ID No:1084 (BX648964)
296	FLJ214 39	image:45814	hypothetical protein flj21439	ughs.55053 6:186	-	0,0375	SEQ ID No:732	SEQ ID No:733	SEQ ID No:734 (nm_025137)
377	HOXA4	image:785930	homeo box a4	ughs.77637: 186	-	0,039	SEQ ID No:941	SEQ ID No:942	SEQ ID No:943 (nm_002141)

311	CACNA 1D/KIF 5C	image:49630	Calcium channel, voltage-dependent, L type, alpha 1D subunit/kinesin family member 5c	ughs.47635 8:186;ughs. 435557:186	-	0,0396	SEQ ID No:775	SEQ ID No:1086	SEQ ID No:1085 (NM_000720)
117	GNG3	image:166254	guanine nucleotide binding protein (g protein), gamma 3	ughs.17991 5:186	-	0,0494		SEQ ID No:275	SEQ ID No:276 (nm_012202)

Table VIII : Metagene overEGFR

Set No.	Gene symbol	Clone ID	Gene name	Unigene Cluster	Regulation	P value	Seq3'	Seq5'	Ref. Seq
171	GSTP1	image:231424	glutathione s-transferase pi	ughs.52383 6:186	+	0,00005	SEQ ID No:403	SEQ ID No:404	SEQ ID No:405 (nm_000852)

402	ITGB3	ipso:0000143	integrin, beta 3 (platelet glycoprotein iiia, antigen cd61)	ughs.218040:186	+	0,00008		SEQ ID No:374 (nm_000212)
217	IGHG1	image:289337	immunoglobulin heavy constant gamma 1 (g1m marker)	ughs.510635:186	+	0,00011	SEQ ID No:520	SEQ ID No:1122
246	SOD2	image:324014	superoxide dismutase 2, mitochondrial	ughs.487046:186	+	0,00072	SEQ ID No:597	SEQ ID No:598 (nm_000636)
111	CEBPB	image:161993	ccaat/enhancer binding protein (c/ebp), beta	ughs.517106:186	+	0,00089	SEQ ID No:261	SEQ ID No:262 (nm_005194)
350	IGKC	image:713852	immunoglobulin kappa constant	ughs.449621:186;ughs.546620:186	+	0,00177	SEQ ID No:872	SEQ ID No:1097 or SEQ ID No:1098 SEQ ID No:1099 (BC066343)
282	ENO1	image:392678	enolase 1, (alpha)	ughs.517145:186	+	0,00201	SEQ ID No:695	SEQ ID No:696 (nm_001428)

94	npc-a-5	image:156691	nasopharyngeal carcinoma-associated antigen npc-a-5	ughs.510543:186	+	0,00352	SEQ ID No:226	SEQ ID No:1058	SEQ ID No:1059 (AK091113)
302	MMP7	image:471134	matrix metalloproteinase 7 (matrilysin, uterine)	ughs.2256:186	+	0,00698	SEQ ID No:749	SEQ ID No:750	SEQ ID No:751 (nm_002423)
142		image:187744		N_A	+	0,01196		SEQ ID No:336	SEQ ID No:1121
122	MKI67	image:1693709	antigen identified by monoclonal antibody ki-67	ughs.80976:186	+	0,0122	SEQ ID No:285		SEQ ID No:286 (nm_002417)
103	ARHGEF1	image:159568	rho guanine nucleotide exchange factor (gef) 1	ughs.278186:186	+	0,01427	SEQ ID No:242	SEQ ID No:243	SEQ ID No:244 (nm_199002)
8	ATF2	image:110999	activating transcription factor 2	ughs.425104:186	+	0,0148	SEQ ID No:16	SEQ ID No:17	SEQ ID No:18 (nm_001880)
50	TFCP2L1	image:1470131	transcription factor cp2-like 1	ughs.156471:186	+	0,0259	SEQ ID No:120		SEQ ID No:121 (nm_014553)
427	IGKC	ipso:0000658	Immunoglobulin kappa variable 1-5	N_A	+	0,02767		SEQ ID No:1033	SEQ ID No:1107 (BC073775)

			(IGKC)						
42	PRSS12	image:145310	protease, serine, 12 (neurotrypsin, motopsin)	ughs.44585 7:186	+	0,03118	SEQ ID No:101	SEQ ID No:102	SEQ ID No:103 (nm_003619)
84	IGLC2	image:154809	immunoglobulin lambda joining 3	ughs.44958 5:186	+	0,04077	SEQ ID No:200	SEQ ID No:201	SEQ ID No:1118
18	CSF1	image:124554	colony stimulating factor 1 (macrophage)	ughs.17389 4:186	+	0,0412	SEQ ID No:41		SEQ ID No:42 (nm_000757)
145	LOC114659	image:188196	SH3-domain GRB2-like pseudogene 1	ughs.40616 6:186;ughs. 438861:186	+	0,04453	SEQ ID No:343 (=SEQ ID No:1066)	SEQ ID No:1065	SEQ ID No:1067 (AK123784)

EXAMPLE 5: Use of Metagenes according to the invention on an Affymetrix® platform (GeneChip® Human Genome U133 plus 2.0 array)

[00194] We profiled 113 samples from the validation set on the Affymetrix® platform to evaluate agreement between the 2 platforms.

[00195] A mapping was performed to find the Affymetrix® probesets corresponding to the sequences comprised into the 3 metagenes, using standard sequence alignment (blast) algorithms.

[00196] For a given gene, several Image clones may exist, each of them covering a particular region of the gene, more commonly in the 3' region. Affymetrix® probesets are also designed to target a specific region of a gene, of around 1000 nucleotides. Clone inserts and Affymetrix® targets do not necessarily overlap, even if the same gene is considered.

[00197] Given this information, there were two possibilities to find a correspondence between Discovery™ and Affymetrix® plateform:

[00198] i) sequence alignment of clone inserts and probesets against a Reference Sequence (ReSeq), which represents a specific gene, and selection of pairs (Clone, Probeset) with homologies to the same Refseq, even if the these sequences do not overlap;

[00199] ii) consider only pairs which overlap, assuming that signal may differ according to the region we focus on. This second approach was chosen to select Affymetrix® probe sets corresponding to the Discovery clones.

[00200] Raw data from Affymetrix® platform were first normalized using the RMA (Robust Multichip Average) method available in Bioconductor (Irizarry et al. 2...) (Affymetrix® package), then corrected to take into account the inter-platform effect and calculate the score for each sample. The data processing applied was the same as previously described on the Discovery™ platform for normalization and Metagenes calculation.

[00201] As an example, comparing sample classification into good or poor prognosis group on Discovery™ and Affymetrix® platform, we obtained 95% when using appropriate confidence interval around the threshold.

[00202] The following tables (IX to XIV) are examples of metagenes of the invention that may be used with an Affymetrix® platform according to the above described methods. For each metagene (IX to XIV), at least two, preferably five, most preferably ten or all of the markers listed, e.g., genes, or marker-derived polynucleotides, e.g., Affymetrix® Probe Sets, may be used to perform these methods. The sequences of the listed Affymetrix® Probe Sets are provided in the enclosed sequence listing and are also publicly available from internet, e.g., www.affymetrix.com. For example, these underER, underPR and underEGFR metagenes may be used in the above described method using a Cox regression analysis and the score S_C = a x underER + b x underPR + c x under EGFR, wherein "a" is comprised in the interval [-6,26 ; +0,49] , "b" is comprised in the interval [-2,65; +0,29] and "c" is comprised in the interval [-6,69; + 1,65]. For example the formula is: S_C = - 2.90279 x underER – 1.47423 x underPR – 4.17198 x under EGFR. Preferably, metagenes of tables IX to XI are used together one the one hand, and metagenes of tables XII to XIV are used together on the other hand.

[00203] The error on the score was integrated by calculating a confidence interval around the threshold, within which sample classification was considered non robust. Considering the score distribution Gaussian, we estimated the confidence interval around the threshold using standard deviation calculation method (estimated standard deviation of the population / $\sqrt{n}$).

[00204] The inventors have established that a woman having a score (S_C) of more than 0.16 have at least a double propensity of poor clinical outcome than a woman with a score (S_C) of less than 0.015.

[00205] Table IX : Metagene underER

Affymetrix ® Probe Set	Clone	Gene	symbol	Unigene Reference	Reference sequence (refseq)	genbank
213094_at	image:259884	G protein-coupled receptor 126	GPR126	hs.318894	nm_001032394, nm_001032395, nm_020455, nm_198569	al033377
204259_at	image:471134	matrix metallopeptidase 7 _matrilysin, uterine_	MMP7	hs.2256	nm_002423	nm_002423
204733_at	image:724109	kallikrein 6 _neurosin, zyme_	KLK6	hs.79361	nm_001012964, nm_001012965, nm_001012966, nm_002774	nm_002774
203560_at	image:809588	gamma-glutamyl hydrolase _conjugase, folylpolygammaglutamyl hydrolase_	GGH	hs.78619	nm_003878	nm_003878
202705_at	image:845594	cyclin B2	CCNB2	hs.194698	nm_004701	nm_004701
227004_at	image:301	Cyclin-dependent kinase-like 5	CDKL5	hs.435570	nm_003159	ai611074

	018					
202967_at	image:345309	glutathione S-transferase A4	GSTA4	hs.485557	nm_001512	nm_001512
218060_s_at	image:43457	hypothetical protein FLJ13154	FLJ13154	hs.408702	nm_024598	nm_024598
201579_at	image:1028762	FAT tumor suppressor homolog 1_Drosophila_	FAT	hs.481371	nm_005245	nm_005245
208370_s_at	ipso:0000077	Down syndrome critical region gene 1	DSCR1	hs.282326	nm_004414, nm_203417, nm_203418	nm_004414
225565_at	image:147707	CDNA FLJ34215 fis, clone FCBBF3021985	NA	hs.516646	---	aa769455
217728_at	image:512420	S100 calcium binding protein A6 _calcyclin_	S100A6	hs.275243	nm_014624	nm_014624
236449_at	image:51814	Cystatin B _stefin B_	CSTB	hs.695	nm_000100	ai885390
212501_at	image:161993	CCAAT_enhancer binding protein _C_EBP_, beta	CEBPB	hs.517106	nm_005194	al564683
201487_at	image:320656	cathepsin C	CTSC	hs.128065	nm_001814, nm_148170	nm_001814
203287_at	image:121551	ladinin 1	LAD1	hs.519035	nm_005558	nm_005558

212531_at	image:544683	lipocalin 2 _oncogene 24p3_	LCN2	hs.204238	nm_005564	nm_005564
212397_at	image:193081	radixin	RDX	hs.263671	nm_002906	al137751
202917_s_at	image:1089513	S100 calcium binding protein A8 _calgranulin A_	S100A8	hs.416073	nm_002964	nm_002964
205487_s_at	image:143622	vestigial like 1 _Drosophila_	VGLL1	hs.496843	nm_016267	nm_016267
221477_s_at	image:324014	hypothetical protein MGC5618	MGC5618	NA	---	bf575213
201037_at	image:152714	phosphofructokinase, platelet	PFKP	hs.26010	nm_002627	nm_002627

[00206] Table X : Metagene underPR

Affymetrix ® Probe Set	Clone	Gene	symbol	Unigene reference	Reference Sequence (refseq)	genbank
201487_at	image:320656	cathepsin C	CTSC	hs.128065	nm_001814, nm_148170	nm_001814
203287_at	image:121	ladinin 1	LAD1	hs.519035	nm_005558	nm_005558

	551					
212531_at	image:544683	lipocalin 2 _oncogene 24p3_	LCN2	hs.204238	nm_005564	nm_005564
212397_at	image:193081	radixin	RDX	hs.263671	nm_002906	al137751
202917_s_at	image:1089513	S100 calcium binding protein A8 _calgranulin A_	S100A8	hs.416073	nm_002964	nm_002964
205487_s_at	image:143622	vestigial like 1 _Drosophila_	VGLL1	hs.496843	nm_016267	nm_016267
221477_s_at	image:324014	hypothetical protein MGC5618	MGC5618	NA	---	bf575213
201037_at	image:152714	phosphofructokinase, platelet	PFKP	hs.26010	nm_002627	nm_002627
202603_at	image:261401	ADAM metallopeptidase domain 10	ADAM10	hs.172028	nm_001110	n51370
210785_s_at	image:307255	chromosome 1 open reading frame 38	C1orf38	hs.10649	nm_004848	ab035482
219386_s_at	image:288807	SLAM family member 8	SLAMF8	hs.438683	nm_020125	nm_020125
203988_s_at	image:156966	fucosyltransferase 8 _alpha _1,6_ fucosyltransferase_	FUT8	hs.118722	nm_004480, nm_178154, nm_178155,	nm_004480

					nm_178156, nm_178157	
202307_s_at	image:51782	transporter 1, ATP-binding cassette, sub-family B _MDR_TAP_	TAP1	hs.352018	nm_000593	nm_000593
210465_s_at	image:219829	small nuclear RNA activating complex, polypeptide 3, 50kDa	SNAPC3	hs.546299	nm_003084	u71300
207498_s_at	image:199680	cytochrome P450, family 2, subfamily D, polypeptide 6	CYP2D6	hs.534311	nm_000106, nm_001025161	nm_000106
215370_at	ipso:0000040	NA	NA	NA	---	au145394
209212_s_at	image:208991	Kruppel-like factor 5 _intestinal_	KLF5	hs.508234	nm_001730	ab030824
219336_s_at	image:50892	activating signal cointegrator 1 complex subunit 1	ASCC1	hs.500007	nm_015947	nm_015947
200709_at	image:159521	FK506 binding protein 1A, 12kDa	FKBP1A	hs.471933	nm_000801, nm_054014	nm_000801
229659_s_at	image:159410	Polymeric immunoglobulin receptor	PIGR	hs.497589	nm_002644	be501712
213572_s_at	ipso:0000605	serpin peptidase inhibitor, clade B _ovalbumin_, member 1	SERPINB1	hs.381167	nm_030666	ai554300
203095_at	image:507	mitochondrial translational	MTIF2	hs.149894	nm_001005369,	nm_002453

	54	initiation factor 2			nm_002453	
240385_at	image:771332	GATA binding protein 6	GATA6	hs.514746	nm_005257	bf002339
243011_at	image:187120	family with sequence similarity 55, member C	FAM55C	hs.130195	nm_145037	bf317081
207004_at	image:342181	B-cell CLL_lymphoma 2	BCL2	hs.150749	nm_000633, nm_000657	nm_000657
219718_at	image:187119	hypothetical protein FLJ10986	FLJ10986	hs.444301	nm_018291	nm_018291
202122_s_at	image:188005	mannose-6-phosphate receptor binding protein 1	M6PRBP1	hs.140452	nm_005817	nm_005817
211372_s_at	image:137575	interleukin 1 receptor, type II	IL1R2	hs.25333	nm_004633, nm_173343	u64094
220529_at	image:154483	hypothetical protein FLJ11710	FLJ11710	NA	---	nm_024846
207988_s_at	image:162208	actin related protein 2_3 complex, subunit 2, 34kDa	ARPC2	hs.529303	nm_005731, nm_152862	nm_005731
211430_s_at	image:289337	immunoglobulin heavy locus ___ immunoglobulin heavy constant gamma 1 _G1m marker_ ___ immunoglobulin heavy constant gamma 2 _G2m	IGH@ ___ IGHG1 ___ IGHG2 ___ IGHG3 ___ IGHM	hs.510635	---	m87789

		marker immunoglobulin heavy constant gamma 3 _G3m marker immunoglobulin heavy constant mu				
220616_at	image:156691	NA	NA	NA	---	nm_006448
213502_x_at	image:50877	similar to bK246H3.1 _immunoglobulin lambda-like polypeptide 1, pre-B-cell specific_	LOC91316	hs.407693	xm_498877	aa398569

[00207] *Table XI : Metagene underEGFR*

Affymetrix® Probe Set	Clone	Gene	symbol	Unigene reference	Reference Sequence (refseq)	Genbank
214440_at	image:145894	N-acetyltransferase 1 _arylamine N-acetyltransferase_	NAT1	hs.155956	nm_000662	nm_000662
232889_at	image:280743	hypothetical protein LOC153561	LOC153561	NA	nm_207331	au147591
219414_at	image:50970	calsyntenin 2	CLSTN2	hs.158529	nm_022131	nm_022131
223044_at	image:489218	solute carrier family 40 _iron-regulated transporter_, member 1	SLC40A1	hs.529285	nm_014585	al136944
229381_at	image:155	chromosome 1 open reading	C1orf64	hs.29190	nm_178840	ai732488

	072	frame 64				
219197_s_at	image:346321	signal peptide, CUB domain, EGF-like 2	SCUBE2	hs.523468	nm_020974	ai424243
225379_at	image:50764	microtubule-associated protein tau	MAPT	hs.101174	nm_005910, nm_016834, nm_016835, nm_016841	aa199717
219570_at	image:52103	chromosome 20 open reading frame 23	C20orf23	hs.101774	nm_024704	nm_024704
225789_at	image:188414	centaurin, gamma 3	CENTG3	hs.195048	nm_031946	be876194
219438_at	image:166862	family with sequence similarity 77, member C	FAM77C	hs.470259	nm_024522	nm_024522
204352_at	image:145410	TNF receptor-associated factor 5	TRAF5	hs.523930	nm_001033910, nm_004619, nm_145759	nm_004619
228994_at	image:52118	coiled-coil domain containing 24	CCDC24	hs.488337	nm_152499	au153816
204550_x_at	image:73778	glutathione S-transferase M1	GSTM1	hs.301961	nm_000561, nm_146421	nm_000561
204623_at	image:298417	trefoil factor 3 _intestinal_	TFF3	hs.82961	nm_003226	nm_003226
222005_s_at	image:166254	guanine nucleotide binding protein _G protein_, gamma 3	GNG3	hs.179915	nm_012202	al538966
220192_x_at	image:1188588	SAM pointed domain containing ets transcription factor	SPDEF	hs.485158	nm_012391	nm_012391
218064_s_at	image:171679	A kinase _PRKA_ anchor protein 8-like	AKAP8L	hs.399800	nm_014371	nm_014371
40093_at	image:160656	Lutheran blood group _Auberger b antigen included_	LU	hs.155048	nm_001013257, nm_005581	x83425
203428_s_at	image:146634	ASF1 anti-silencing function 1 homolog A _S. cerevisiae_	ASF1A	hs.292316	nm_014034	ab028628
204129_at	image:175	B-cell CLL_lymphoma 9	BCL9	hs.415209	nm_004326	nm_004326

	6392					
224182_x_at	image:166010	sema domain, transmembrane domain _TM_, and cytoplasmic domain, _semaphorin_ 6B	SEMA6B	hs.465642	nm_020241, nm_032108, nm_133327	af293363
204418_x_at	image:166910	glutathione S-transferase M2 _muscle_	GSTM2	hs.279837	nm_000848	nm_000848
201681_s_at	image:153368	discs, large homolog 5 _Drosophila_	DLG5	hs.500245	nm_004747	ab011155
233955_x_at	image:173797	CXXC finger 5	CXXC5	hs.189119	nm_016463	ak001782
205225_at	image:725321	estrogen receptor 1	ESR1	hs.208124	nm_000125	nm_000125
205201_at	image:767495	GLI-Kruppel family member GLI3 _Greig cephalopolysyndactyly syndrome_	GLI3	hs.199338	nm_000168	nm_000168
209049_s_at	image:511899	protein kinase C binding protein 1	PRKCBP1	hs.446240	nm_012408, nm_183047, nm_183048	bc001004
218367_x_at	image:183062	ubiquitin specific peptidase 21	USP21	hs.8015	nm_001014443, nm_012475	nm_012475
212099_at	image:768370	ras homolog gene family, member B	RHOB	hs.502876	nm_004040	ai263909
201613_s_at	image:161763	adaptor-related protein complex 1, gamma 2 subunit	AP1G2	hs.343244	nm_003917, nm_080545	bc000519
201754_at	image:278531	cytochrome c oxidase subunit VIc	COX6C	hs.351875	nm_004374	nm_004374
222282_at	image:155064	Ubiquitin specific peptidase 13 _isopeptidase T-3_	USP13	hs.175322	nm_003940	av761453
208451_s_at	image:340753	complement component 4A ___ complement component 4B ___ complement component 4B, telomeric	C4A ___ C4B	hs.534847	nm_000592, nm_001002029, nm_007293	nm_000592

214428_x_at	image:491004	complement component 4A /// complement component 4B /// complement component 4B, telomeric	C4A /// C4B	hs.534847	nm_000592, nm_001002029, nm_007293	k02403
219426_at	image:155341	eukaryotic translation initiation factor 2C, 3	EIF2C3	hs.567761	nm_024852, nm_177422	nm_024852
209604_s_at	image:139076	GATA binding protein 3	GATA3	hs.524134	nm_001002295, nm_002051	bc003070
201596_x_at	image:151663	keratin 18	KRT18	hs.406013	nm_000224, nm_199187	nm_000224

[00208] Table XII: Metagene underER

Affymetrix® Probe Set	Clone	Gene	symbol	Unigene reference	Reference Sequence (refseq)	Genbank
200824_at	image:231424	glutathione S-transferase pi	GSTP1	Hs.523836	NM_000852	NM_000852
201037_at	image:152714	phosphofructokinase, platelet	PFKP	Hs.26010	NM_002627	NM_002627
201201_at	image:51814	cystatin B (stefin B)	CSTB	Hs.695	NM_000100	NM_000100
201231_s_at	image:392678	enolase 1, (alpha)	ENO1	Hs.517145	NM_001428	NM_001428
201487_at	image:320656	cathepsin C	CTSC	Hs.128065	NM_001814	NM_001814

201579_at	image:1028762	FAT tumor suppressor homolog 1 (Drosophila)	FAT	Hs.481371	NM_005245	NM_005245
201710_at	image:207378	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	MYBL2	Hs.179718	NM_002466	NM_002466
202705_at	image:845594	cyclin B2	CCNB2	Hs.194698	NM_004701	NM_004701
202967_at	image:345309	glutathione S-transferase A4	GSTA4	Hs.485557	NM_001512	NM_001512
203256_at	ipso:0000143	cadherin 3, type 1, P-cadherin (placental)	CDH3	Hs.461074	NM_001793	NM_001793
203287_at	image:121551	ladinin 1	LAD1	Hs.519035	NM_005558	NM_005558
203560_at	image:809588	gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	GGH	Hs.78619	NM_003878	NM_003878
204092_s_at	image:1912132	aurora kinase A	AURKA	Hs.250822	NM_003600	NM_003600

204259_at	image:471134	matrix metallopeptidase 7 (*matrilysin, uterine*)	MMP7	Hs.2256	NM_002423	NM_002423
204733_at	image:724109	kallikrein-related peptidase 6	KLK6	Hs.79361	NM_001012964	NM_002774
208370_s_at	ipso:0000077	regulator of calcineurin 1	RCAN1	Hs.282326	NM_004414	NM_004414
208456_s_at	image:278490	related RAS viral (r-ras) oncogene homolog 2	RRAS2	Hs.502004	NM_012250	NM_012250
209791_at	ipso:0000610	peptidyl arginine deiminase, type II	PADI2	Hs.33455	NM_007365	AL049569
210453_x_at	ipso:0000267	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit G	ATP5L	Hs.486360	NM_006476	AL050277
212398_at	image:193081	radixin	RDX	Hs.263671	NM_002906	AI057093
212501_at	image:161993	CCAAT/enhancer binding protein (C/EBP), beta	CEBPB	Hs.517106	NM_005194	AL564683
212531_at	image:544683	lipocalin 2 (oncogene 24p3)	LCN2	Hs.204238	NM_005564	NM_005564

213094_at	image:259884	G protein-coupled receptor 126	GPR126	Hs.318894	NM_001032394	AL033377
214370_at	image:1089513	S100 calcium binding protein A8	S100A8	Hs.416073	NM_002964	AW238654
215223_s_at	image:324014	superoxide dismutase 2, mitochondrial	SOD2	Hs.487046	NM_000636	W46388
215729_s_at	image:143622	vestigial like 1 (Drosophila)	VGLL1	Hs.496843	NM_016267	BE542323
217728_at	image:512420	S100 calcium binding protein A6	S100A6	Hs.275243	NM_014624	NM_014624
218060_s_at	image:43457	chromosome 16 open reading frame 57	C16orf57	Hs.588873	NM_024598	NM_024598
221477_s_at	ipso:0000488	hypothetical protein MGC5618	MGC5618	NA	NA	BF575213

[00209] Table XIII : Metagene underPR

Affymetrix® Probe Set	Clone	Gene	symbol	Unigene reference	Reference Sequence (refseq)	Genbank
201487_at	image:320656	cathepsin C	CTSC	Hs.128065	NM_001814	NM_001814

201505_at	image:428443	laminin, beta 1	LAMB1	Hs.650585	NM_002291	NM_002291
201710_at	image:207378	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	MYBL2	Hs.179718	NM_002466	NM_002466
201710_at	image:724259	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	MYBL2	Hs.179718	NM_002466	NM_002466
202036_s_at	image:783700	secreted frizzled-related protein 1	SFRP1	Hs.213424	NM_003012	AF017987
202246_s_at	image:725349	cyclin-dependent kinase 4	CDK4	Hs.95577	NM_000075	NM_000075
202307_s_at	image:51782	transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)	TAP1	Hs.352018	NM_000593	NM_000593
202519_at	image:159783	MLX interacting protein	MLXIP	Hs.437153	NM_014938	NM_014938
203095_at	image:50754	mitochondrial translational initiation factor 2	MTIF2	Hs.149894	NM_001005369	NM_002453
203256_at	ipso:0000143	cadherin 3, type 1, P-cadherin (placental)	CDH3	Hs.461074	NM_001793	NM_001793
203287_at	image:121	ladinin 1	LAD1	Hs.519035	NM_005558	NM_005558

	551					
203685_at	image:342181	B-cell CLL/lymphoma 2	BCL2	Hs.150749	NM_000633	NM_000633
203934_at	image:193857	kinase insert domain receptor (a type III receptor tyrosine kinase)	KDR	Hs.479756	NM_002253	NM_002253
204470_at	image:323238	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	CXCL1	Hs.789	NM_001511	NM_001511
204628_s_at	image:200209	integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)	ITGB3	Hs.218040	NM_000212	NM_000212
205890_s_at	ipso:00002 52	ubiquitin D	UBD	Hs.44532	NM_006398	NM_006398
206324_s_at	image:156808	death-associated protein kinase 2	DAPK2	Hs.237886	NM_014326	NM_014326
206792_x_at	image:219829	phosphodiesterase 4C, cAMP-specific (phosphodiesterase E1 dunce homolog, Drosophila)	PDE4C	Hs.631628	NM_000923	NM_000923
207270_x_at	image:156937	CD300c molecule	CD300C	Hs.2605	NM_006678	NM_006678

207498_s_at	image:199680	cytochrome P450, family 2, subfamily D, polypeptide 6	CYP2D6	Hs.648256	NM_000106	NM_000106
207571_x_at	image:307255	chromosome 1 open reading frame 38	C1orf38	Hs.10649	NM_001039477	NM_004848
209138_x_at	ipso:00004 34	immunoglobulin lambda locus	IGL@	Hs.449585	NA	M87790
209791_at	ipso:00006 10	peptidyl arginine deiminase, type II	PADI2	Hs.33455	NM_007365	AL049569
209848_s_at	image:342383	silver homolog (mouse)	SILV	Hs.95972	NM_006928	U01874
210002_at	image:771332	GATA binding protein 6	GATA6	Hs.514746	NM_005257	D87811
211372_s_at	image:137575	interleukin 1 receptor, type II	IL1R2	Hs.25333	NM_004633	U64094
211430_s_at	image:289337	immunoglobulin heavy constant gamma 3 (G3m marker)	IGHG3	Hs.510635	NA	M87789
212398_at	image:193081	radixin	RDX	Hs.263671	NM_002906	AI057093
212531_at	image:544683	lipocalin 2 (oncogene 24p3)	LCN2	Hs.204238	NM_005564	NM_005564
213572_s_at	ipso:00006 05	serpin peptidase inhibitor, clade B (ovalbumin), member 1	SERPINB1	Hs.381167	NM_030666	AI554300

214370_at	image:1089513	S100 calcium binding protein A8	S100A8	Hs.416073	NM_002964	AW238654
215223_s_at	image:324014	superoxide dismutase 2, mitochondrial	SOD2	Hs.487046	NM_000636	W46388
215729_s_at	image:143622	vestigial like 1 (Drosophila)	VGLL1	Hs.496843	NM_016267	BE542323
215946_x_at	image:50877	similar to omega protein	CTA-246H3.1	Hs.567636	NM_001013618	AL022324
216598_s_at	ipso:0000152	chemokine (C-C motif) ligand 2	CCL2	Hs.303649	NM_002982	S69738
217865_at	image:186926	ring finger protein 130	RNF130	Hs.484363	NM_018434	NM_018434
219386_s_at	image:288807	SLAM family member 8	SLAMF8	Hs.438683	NM_020125	NM_020125
221651_x_at	image:156691	immunoglobulin kappa constant	IGKC	Hs.449621	NA	BC005332
221671_x_at	ipso:0000376	immunoglobulin kappa constant	IGKC	Hs.449621	NA	M63438
224795_x_at	image:713852	immunoglobulin kappa constant	IGKC	Hs.449621	NA	AW575927
227262_at	image:187120	hyaluronan and proteoglycan link protein 3	HAPLN3	Hs.447530	NM_178232	BE348293

243209_at	image:156966	potassium voltage-gated channel, KQT-like subfamily, member 4	KCNQ4	Hs.473058	NM_004700	BF725804

Table XIV: Metagene underEGFR

Affymetrix® Probe Set	Clone	Gene	symbol	Unigene reference	Reference Sequence (refseq)	Genbank
200670_at	ipso:0000125	X-box binding protein 1	XBP1	Hs.437638	NM_005080	NM_001079539
200670_at	image:301950	X-box binding protein 1	XBP1	Hs.437638	NM_005080	NM_001079539
201596_x_at	image:151663	keratin 18	KRT18	Hs.406013	NM_000224	NM_000224
201613_s_at	image:161763	adaptor-related protein complex 1, gamma 2 subunit	AP1G2	Hs.343244	BC000519	NM_003917
201681_s_at	image:153368	discs, large homolog 5 (Drosophila)	DLG5	Hs.654780	AB011155	NM_004747

201754_at	image:278531	cytochrome c oxidase subunit VIc	COX6C	Hs.351875	NM_004374	NM_004374
201860_s_at	image:160149	plasminogen activator, tissue	PLAT	Hs.491582	NM_000930	NM_000930
201860_s_at	ipso:00002 53	plasminogen activator, tissue	PLAT	Hs.491582	NM_000930	NM_000930
204129_at	image:1756392	B-cell CLL/lymphoma 9	BCL9	Hs.415209	NM_004326	NM_004326
204352_at	image:145410	TNF receptor-associated factor 5	TRAF5	Hs.523930	NM_004619	NM_001033910
204418_x_at	image:166910	glutathione S-transferase M2 (muscle)	GSTM2	Hs.279837	NM_000848	NM_000848
204418_x_at	image:153444	glutathione S-transferase M2 (muscle)	GSTM2	Hs.279837	NM_000848	NM_000848
204418_x_at	image:664233	glutathione S-transferase M2 (muscle)	GSTM2	Hs.279837	NM_000848	NM_000848
204550_x_at	image:73778	glutathione S-transferase M1	GSTM1	Hs.301961	NM_000561	NM_000561
204623_at	image:298417	trefoil factor 3 (intestinal)	TFF3	Hs.82961	NM_003226	NM_003226

205009_at	image:1075949	trefoil factor 1	TFF1	Hs.162807	NM_003225	NM_003225
205186_at	image:782688	dynein, axonemal, light intermediate chain 1	DNALI1	Hs.406050	NM_003462	NM_003462
205201_at	image:767495	GLI-Kruppel family member GLI3 (Greig cephalopolysyndactyly syndrome)	GLI3	Hs.21509	NM_000168	NM_000168
205225_at	image:725321	estrogen receptor 1	ESR1	Hs.208124	NM_000125	NM_000125
206107_at	image:277917	regulator of G-protein signaling 11	RGS11	Hs.65756	NM_003834	NM_003834
206289_at	image:785930	homeobox A4	HOXA4	Hs.654466	NM_002141	NM_002141
206401_s_at	image:50764	microtubule-associated protein tau	MAPT	Hs.101174	J03778	NM_005910
208451_s_at	image:340753	complement component 4A (Rodgers blood group)	C4A	Hs.655564	NM_000592	NM_007293

208451_s_at	image:491004	complement component 4A (Rodgers blood group)	C4A	Hs.655564	NM_000592	NM_007293
209048_s_at	image:511899	zinc finger, MYND-type containing 8	ZMYND8	Hs.446240	AB032951	NM_012408
209604_s_at	image:139076	GATA binding protein 3	GATA3	Hs.524134	BC003070	NM_001002295
209604_s_at	ipso:0000286	GATA binding protein 3	GATA3	Hs.524134	BC003070	NM_001002295
210108_at	image:49630	calcium channel, voltage-dependent, L type, alpha 1D subunit	CACNA1D	Hs.476358	BE550599	NM_000720
210272_at	image:182295	cytochrome P450, family 2, subfamily B, polypeptide 7 pseudogene 1	CYP2B7P1	Hs.529117	M29873	NR_001278
211038_s_at	image:149567	ciliary rootlet coiled-coil, rootletin-like 1	CROCCL1	Hs.631865	BC006312	XM_001130627

212099_at	image:768370	ras homolog gene family, member B	RHOB	Hs.502876	AI263909	NM_004040
212099_at	image:149760	ras homolog gene family, member B	RHOB	Hs.502876	AI263909	NM_004040
214440_at	image:145894	N-acetyltransferase 1 (arylamine N-acetyltransferase)	NAT1	Hs.591847	NM_000662	NM_000662
218064_s_at	image:171679	A kinase (PRKA) anchor protein 8-like	AKAP8L	Hs.399800	NM_014371	NM_014371
218211_s_at	image:155341	melanophilin	MLPH	Hs.102406	NM_024101	NM_001042467
218692_at	image:149549	Golgi-localized protein	GOLSYN	Hs.390738	NM_017786	NM_001099743
219197_s_at	image:346321	signal peptide, CUB domain, EGF-like 2	SCUBE2	Hs.523468	AI424243	NM_020974
219438_at	image:166862	family with sequence similarity 77, member C	FAM77C	Hs.470259	NM_024522	NM_024522

219570_at	image:521 03	chromosome 20 open reading frame 23	C20orf23	Hs.101774	NM_024704	NM_024704
220192_x_at	image:118 8588	SAM pointed domain containing ets transcription factor	SPDEF	Hs.485158	NM_012391	NM_012391
220778_x_at	image:527 41	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6B	SEMA6B	Hs.465642	NM_020241	NM_020241
222005_s_at	image:166 254	guanine nucleotide binding protein (G protein), gamma 3	GNG3	Hs.179915	AL538966	NM_012202
223044_at	image:489 218	solute carrier family 40 (iron-regulated transporter), member 1	SLC40A1	Hs.643005	AL136944	NM_014585

223721_s_at	image:120138	DnaJ (Hsp40) homolog, subfamily C, member 12	DNAJC12	Hs.260720	AF176013	NM_021800
224516_s_at	image:173797	CXXC finger 5	CXXC5	Hs.189119	BC006428	NM_016463
225092_at	image:772890	nucleoporin 88kDa	NUP88	Hs.584784	AL550977	NM_002532
225883_at	image:50602	ATG16 autophagy related 16-like 2 (S. cerevisiae)	ATG16L2	Hs.653186	AK024423	NM_033388
225911_at	image:266500	nephronectin	NPNT	Hs.518921	AL138410	NM_001033047
226362_at	image:280743	small EDRK-rich factor 1A (telomeric)	SERF1A	Hs.658079	AI198515	NM_021967
226373_at	image:147138	sideroflexin 5	SFXN5	Hs.368171	AW166098	NM_144579
226506_at	image:160656	thrombospondin, type I, domain containing 4	THSD4	Hs.387057	AI742570	NM_024817
227425_at	image:43488	RALBP1 associated Eps domain containing 2	REPS2	Hs.186810	AI984607	NM_001080975

227515_at	image:188414	STAM binding protein	STAMBP	Hs.469018	AU158421	NM_006463
227550_at	image:44338	hypothetical protein LOC143381	LOC143381	Hs.388347	AW242720	NA
227811_at	image:146634	FYVE, RhoGEF and PH domain containing 3	FGD3	Hs.411081	AK000004	NM_001083536
228528_at	image:153617	NA	NA	NA	AI927692	NA
228994_at	image:52118	coiled-coil domain containing 24	CCDC24	Hs.632394	AU153816	NM_152499
229150_at	ipso:0000614	melanophilin	MLPH	Hs.102406	AI810764	NM_001042467
229381_at	image:155072	chromosome 1 open reading frame 64	C1orf64	Hs.29190	AI732488	NM_178840

[00210] The above described protocol for finding a correspondence between a cDNA platform (e.g., Discovery™) and another platform (e.g., Affymetrix®) may be similarly applied by a person skilled in the art for the other metagenes according to the present invention.

CLAIMS

1. A method of assessing the clinical outcome of a female mammal suffering from breast cancer, comprising the steps of:

a) generating a metagene adjusted value underER by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 10 nucleic acid sequences selected in the group comprising or consisting of: SEQ ID No:374 (nm_000212), SEQ ID No:1027 (nm_007365), SEQ ID No:598 (nm_000636), SEQ ID No:717 (nm_024598), SEQ ID No:573 (nm_001527), SEQ ID No:83 (nm_015065), SEQ ID No:12 (nm_002964), SEQ ID No:405 (nm_000852), SEQ ID No:856 (nm_005564), SEQ ID No:384 (nm_002466), SEQ ID No:167 (nm_002627), SEQ ID No:51 (nm_198433), SEQ ID No:999 (nm_145290), SEQ ID No:979 (nm_004414), SEQ ID No:2 (nm_005245), SEQ ID No:98 (nm_016267), SEQ ID No:751 (nm_002423), SEQ ID No:696 (nm_001428), SEQ ID No:1050 (BC034638), SEQ ID No:488 (nm_002979), SEQ ID No:262 (nm_005194), SEQ ID No:1020 (nm_000359), SEQ ID No:1106 (BC015969), SEQ ID No:952 (nm_003878), SEQ ID No:675 (nm_001512), SEQ ID No:289 (nm_020179), SEQ ID No:553 (nm_004701), SEQ ID No:579 (nm_001814), SEQ ID No:760 (nm_005746), SEQ ID No:805 (nm_014624), SEQ ID No:361 (nm_002906), SEQ ID No:448 (nm_198569), SEQ ID No:170 (nm_002428), SEQ ID No:878 (nm_002774), SEQ ID No:1117, SEQ ID No:612 (nm_032515), SEQ ID No:540 (nm_003159), SEQ ID No:823 (nm_000100), SEQ ID No:131 (nm_145280), SEQ ID No:705 (nm_005596), SEQ ID No:31 (nm_005558), and SEQ ID No:199 (nm_024323) fragments, derivatives or complementary sequences thereof;

b) generating a metagene adjusted value underPR by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 6 nucleic acid sequences selected in the group comprising or consisting of : SEQ ID No:598 (nm_000636), SEQ ID No:1122, SEQ ID No:364 (nm_002253), SEQ ID No:387 (nm_006563), SEQ ID No:34 (nm_001229), SEQ ID No:657 (nm_000633), SEQ ID No:384 (nm_002466),

SEQ ID No:451 (nm_001110), SEQ ID No:999 (nm_145290), SEQ ID No:1056 (AK126297), SEQ ID No:15 (nm_003243), SEQ ID No:1090 (AK125808), SEQ ID No:1120, SEQ ID No:12 (nm_002964), SEQ ID No:743 (nm_006875), SEQ ID No:414 (nm_000546), SEQ ID No:374 (nm_000212), SEQ ID No:711 (nm_002291), SEQ ID No:663 (nm_006928), SEQ ID No:1102 (AK124587), SEQ ID No:237 (nm_002644), SEQ ID No:60 (nm_022640), SEQ ID No:361 (nm_002906), SEQ ID No:119 (nm_004730) (or SEQ ID No:1109 (NM_002019)), SEQ ID No:167 (nm_002627), SEQ ID No:339 (nm_144970), SEQ ID No:333 (nm_145037), SEQ ID No:83 (nm_015065), SEQ ID No:330 (nm_018291), SEQ ID No:1024 (nm_030666), SEQ ID No:229 (nm_004586), SEQ ID No:925 (nm_005257), SEQ ID No:788 (nm_001005369), SEQ ID No:1104 (AK128524), SEQ ID No:1103 (BX108410), SEQ ID No:66 (nm_000416), SEQ ID No:1030 (nm_024007), SEQ ID No:1119, SEQ ID No:1068 (AK024670), SEQ ID No:241 (nm_000801), SEQ ID No:398 (nm_003084), SEQ ID No:74 (nm_000878), SEQ ID No:1087 (AK074131), SEQ ID No:955 (nm_001986), SEQ ID No:71 (nm_004633), SEQ ID No:1105 (BC072392), SEQ ID No:856 (nm_005564), SEQ ID No:231 (nm_006678), SEQ ID No:593 (nm_001511), SEQ ID No:384 (nm_002466), SEQ ID No:519 (nm_020125), SEQ ID No:579 (nm_001814), SEQ ID No:1039 (nm_006209), SEQ ID No:31 (nm_005558), SEQ ID No:327 (nm_173825), SEQ ID No:573 (nm_001527), SEQ ID No:98 (nm_016267), SEQ ID No:1059 (AK091113), SEQ ID No:886 (nm_000075), SEQ ID No:1032 (nm_005688), SEQ ID No:1091 (XM_378178), SEQ ID No:233 (nm_178155), SEQ ID No:938 (nm_003012), SEQ ID No:264 (nm_152862), SEQ ID No:546 (nm_005874), SEQ ID No:1099 (BC066343) SEQ ID No:1037 (nm_023068), SEQ ID No:550 (nm_004848), SEQ ID No:1027 (nm_007365), SEQ ID No:1005 (nm_014938), SEQ ID No:820 (nm_000593), and SEQ ID No:370 (nm_000106), fragments, derivatives or complementary sequences thereof;

c) generating a metagene adjusted value underEGFR by comparing the level, in a biological sample from said female mammal and in a control, of at least 10 nucleic acid sequences selected in the group comprising or consisting of : SEQ ID No:1071 (NM_001033047), SEQ ID No:254

(nm_005581), SEQ ID No:6 (nm_003225), SEQ ID No:883 (nm_000125), SEQ ID No:543 (nm_005080), SEQ ID No:681 (nm_020974), SEQ ID No:63 (nm_001002295), SEQ ID No:212 (nm_024852), SEQ ID No:635 (nm_001002029), SEQ ID No:535 (nm_003226), SEQ ID No:1125, SEQ ID No:109 (nm_000662), SEQ ID No:342 (nm_001846), SEQ ID No:927 (nm_004703), SEQ ID No:1124, SEQ ID No:124 (nm_014899), SEQ ID No:280 (nm_020764) (or SEQ ID No:1110 (NM_024522)), SEQ ID No:297 (nm_016463), SEQ ID No:791 (nm_016835), SEQ ID No:210 (nm_178840), SEQ ID No:827 (nm_152499), SEQ ID No:1064 (NM_000767), SEQ ID No:147 (nm_014675), SEQ ID No:323 (nm_001014443), SEQ ID No:106 (nm_004619), SEQ ID No:181 (nm_000848), SEQ ID No:376 (nm_057158), SEQ ID No:116 (nm_014034), SEQ ID No:252 (nm_000758), SEQ ID No:797 (nm_022131), SEQ ID No:911 (nm_000168), SEQ ID No:720 (nm_004726), SEQ ID No:889 (nm_000561), SEQ ID No:250 (nm_000930), SEQ ID No:179 (nm_004747), SEQ ID No:786 (nm_033388), SEQ ID No:177 (nm_015996), SEQ ID No:1047 (BC012900), SEQ ID No:301 (nm_004326), SEQ ID No:207 (nm_003940), SEQ ID No:936 (nm_003462), SEQ ID No:916 (nm_001453) (or SEQ ID No:1116 (NM_004040)), SEQ ID No:1052 (BX096026), SEQ ID No:159 (nm_000224), SEQ ID No:1096 (AK127274), SEQ ID No:28 (nm_021800), SEQ ID No:1054 (AK123264), SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (NM_053279)), SEQ ID No:825 (nm_024704), SEQ ID No:145 (nm_017786), SEQ ID No:491 (nm_004374), SEQ ID No:485 (nm_003834), SEQ ID No:1072 (AY007114), SEQ ID No:274 (nm_032108), SEQ ID No:258 (nm_080545), SEQ ID No:292 (nm_014371), SEQ ID No:803 (nm_183047), SEQ ID No:349 (nm_031946), SEQ ID No:1123, SEQ ID No:763 (nm_014585), SEQ ID No:438 (nm_001759), SEQ ID No:94 (nm_014315), SEQ ID No:845 (nm_001089), SEQ ID No:1084 (BX648964), SEQ ID No:734 (nm_025137), SEQ ID No:943 (nm_002141), SEQ ID No:1085 (NM_000720), and SEQ ID No:276 (nm_012202), fragments, derivatives or complementary sequences thereof;

d) generating a score (S_C) from said metagene adjusted values using a mathematical method establishing a relation between the combined metagene values and the clinical outcome of said female mammal.

2. The method of claim 1 wherein the mathematical method used in step d) comprises a Cox regression or CART analysis.

3. The method of claim 1 wherein the mathematical method used in step d) is a Cox regression and the score (S_C) is generated according to the following formula: S_C = a x underER + b x underPR + c x under EGFR, wherein "a" is comprised in the interval [-6,26 ; +0,49] , "b" is comprised in the interval [-2,65; +0,29] and "c" is comprised in the interval [-6,69; + 1,65].

4. The method of claim 3 wherein the score S_C is generated according to the following formula: S_C = - 2.90279 x underER – 1.47423 x underPR – 4.17198 x under EGFR.

5. The method of claim 1, wherein the step a) of generating a metagene adjusted value underER is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 20 nucleic acid sequences selected in said group.

6. The method of claim 5, wherein the step a) of generating a metagene adjusted value underER is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 25 nucleic acid sequences selected in said group.

7. The method of claim 1, wherein said metagene adjusted value underER is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 20 nucleic acid sequences selected in the group consisting of : SEQ ID No:374 (nm_000212); SEQ ID No:1027 (nm_007365); SEQ ID No:598 (nm_000636); SEQ ID No:573

(nm_001527); SEQ ID No:83 (nm_015065); SEQ ID No:12 (nm_002964); SEQ ID No:405 (nm_000852); SEQ ID No:856 (nm_005564); SEQ ID No:167 (nm_002627); SEQ ID No:51 (nm_198433); SEQ ID No:98 (nm_016267); SEQ ID No:751 (nm_002423); SEQ ID No:696 (nm_001428); SEQ ID No:262 (nm_005194); SEQ ID No:1020 (nm_000359); SEQ ID No:579 (nm_001814); SEQ ID No:760 (nm_005746); SEQ ID No:805 (nm_014624); SEQ ID No:878 (nm_002774); and SEQ ID No:612 (nm_032515), fragments, derivatives or complementary sequences thereof.

8. The method of claim 7, wherein said metagene adjusted value underER is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 27 nucleic acid sequences selected in the group consisting of : SEQ ID No:374 (nm_000212); SEQ ID No:1027 (nm_007365); SEQ ID No:598 (nm_000636); SEQ ID No:573 (nm_001527); SEQ ID No:83 (nm_015065); SEQ ID No:12 (nm_002964); SEQ ID No:405 (nm_000852); SEQ ID No:856 (nm_005564); SEQ ID No:167 (nm_002627); SEQ ID No:51 (nm_198433); SEQ ID No:98 (nm_016267); SEQ ID No:751 (nm_002423); SEQ ID No:696 (nm_001428); SEQ ID No:262 (nm_005194); SEQ ID No:1020 (nm_000359); SEQ ID No:579 (nm_001814); SEQ ID No:760 (nm_005746); SEQ ID No:805 (nm_014624); SEQ ID No:878 (nm_002774); SEQ ID No:612 (nm_032515); SEQ ID No:384 (nm_002466); SEQ ID No:2 (nm_005245); SEQ ID No:1050 (BC034638); SEQ ID No:952 (nm_003878); SEQ ID No:361 (nm_002906); SEQ ID No:31 (nm_005558); and SEQ ID No:199 (nm_024323), fragments, derivatives or complementary sequences thereof.

9. The method of claim 1, wherein the step b) of generating a metagene adjusted value underPR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 10 nucleic acid sequences selected in said group.

10. The method of claim 9, wherein the step b) of generating a metagene adjusted value underPR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 20 nucleic acid sequences selected in said group.

11. The method of claim 10, wherein the step b) of generating a metagene adjusted value underPR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 30 nucleic acid sequences selected in said group.

12. The method of claim 10, wherein the step b) of generating a metagene adjusted value underPR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 36 nucleic acid sequences selected in said group.

13. The method of claim 1, wherein said metagene adjusted value underPR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 6 nucleic acid sequences selected in the group consisting of : SEQ ID No:364 (nm_002253); SEQ ID No:34 (nm_001229); SEQ ID No:657 (nm_000633); SEQ ID No:339 (nm_144970); SEQ ID No:229 (nm_004586); SEQ ID No:1119, fragments, derivatives or complementary sequences thereof.

14. The method of claim 1, wherein said metagene adjusted value underPR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 36 nucleic acid sequences selected in the group consisting of : SEQ ID No:364 (nm_002253); SEQ ID No:34 (nm_001229); SEQ ID No:657 (nm_000633); SEQ ID No:339 (nm_144970); SEQ ID No:229 (nm_004586); SEQ ID No:1119; SEQ ID No:387 (nm_006563); SEQ ID No:1056 (AK126297); SEQ ID No:15 (nm_003243); SEQ ID No:1120; SEQ ID No:414 (nm_000546); SEQ ID No:374 (nm_000212); SEQ ID No:711 (nm_002291); SEQ ID No:663

(nm_006928); SEQ ID No:237 (nm_002644); SEQ ID No:60 (nm_022640); SEQ ID No:119 (nm_004730) ; SEQ ID No:330 (nm_018291); SEQ ID No:1024 (nm_030666); SEQ ID No:925 (nm_005257); SEQ ID No:1104 (AK128524); SEQ ID No:1103 (BX108410); SEQ ID No:66 (nm_000416); SEQ ID No:1068 (AK024670); SEQ ID No:374 (nm_000212); SEQ ID No:74 (nm_000878); SEQ ID No:231 (nm_006678); SEQ ID No:593 (nm_001511); SEQ ID No:384 (nm_002466); SEQ ID No:1039 (nm_006209); SEQ ID No:327 (nm_173825); SEQ ID No:886 (nm_000075); SEQ ID No:1032 (nm_005688); SEQ ID No:264 (nm_152862); SEQ ID No:1037 (nm_023068); and SEQ ID No:1005 (nm_014938), fragments, derivatives or complementary sequences thereof.

15. The method of claim 1, wherein the step c) of generating a metagene adjusted value underEGFR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 20 nucleic acid sequences selected in said group.

16. The method of claim 15, wherein the step c) of generating a metagene adjusted value underEGFR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 24 nucleic acid sequences selected in said group.

17. The method of claim 15, wherein the step c) of generating a metagene adjusted value underEGFR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 30 nucleic acid sequences selected in said group.

18. The method of claim 17, wherein the step c) of generating a metagene adjusted value underEGFR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 37 nucleic acid sequences selected in said group.

19. The method of claim 1, wherein said metagene adjusted value underEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 24 nucleic acid sequences selected in the group consisting of : SEQ ID No:1071 (nm_001033047); SEQ ID No:254 (nm_005581); SEQ ID No:6 (nm_003225); SEQ ID No:883 (nm_000125); SEQ ID No:543 (nm_005080); SEQ ID No:681 (nm_020974); SEQ ID No:63 (nm_001002295); SEQ ID No:212 (nm_024852); SEQ ID No:635 (nm_001002029); SEQ ID No:535 (nm_003226); SEQ ID No:1125); SEQ ID No:1124; SEQ ID No:297 (nm_016463); SEQ ID No:791 (nm_016835); SEQ ID No:827 (nm_152499); SEQ ID No:207 (nm_003940); SEQ ID No:916 (nm_001453) (or SEQ ID No:1116 (nm_004040)); SEQ ID No:1052 (BX096026) ; SEQ ID No:159 (nm_000224); SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (NM_053279)); SEQ ID No:845 (nm_001089); and SEQ ID No:1085 (NM_000720), fragments, derivatives or complementary sequences thereof.

20. The method of claim 19, wherein said metagene adjusted value underEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 37 nucleic acid sequences selected in the group consisting of : SEQ ID No:1071 (nm_001033047); SEQ ID No:254 (nm_005581); SEQ ID No:6 (nm_003225); SEQ ID No:883 (nm_000125); SEQ ID No:543 (nm_005080); SEQ ID No:681 (nm_020974); SEQ ID No:63 (nm_001002295); SEQ ID No:212 (nm_024852); SEQ ID No:635 (nm_001002029); SEQ ID No:535 (nm_003226); SEQ ID No:1125; SEQ ID No:1124; SEQ ID No:297 (nm_016463); SEQ ID No:791 (nm_016835); SEQ ID No:827 (nm_152499); SEQ ID No:207 (nm_003940); SEQ ID No:916 (nm_001453) (or SEQ ID No:1116 (nm_004040)); SEQ ID No:1052 (BX096026) ; SEQ ID No:159 (nm_000224); SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (NM_053279)); SEQ ID No:845 (nm_001089); SEQ ID No:1085 (NM_000720) ; SEQ ID No:109 (nm_000662); SEQ ID No:342 (nm_001846); SEQ ID No:927 (nm_004703); SEQ ID No:280 (nm_020764) (or SEQ ID No:1110 (NM_024522)); SEQ ID No:210

(nm_178840); SEQ ID No:181 (nm_000848); SEQ ID No:116 (nm_014034); SEQ ID No:250 (nm_000930); SEQ ID No:177 (nm_015996); SEQ ID No:825 (nm_024704); SEQ ID No:145 (nm_017786); and SEQ ID No:276 (nm_012202), fragments, derivatives or complementary sequences thereof.

21. The method of claim 1, further comprising the first step of quantifying in a biological sample from said female mammal the expression level of said nucleic acids sequences.

22. The method of claim 1, further comprising the step e) of comparing said score (S_C) from the biological sample with a baseline or a score (S_C) from a control sample.

23. The method according to claim 1, wherein said score (S_C) from the biological sample is correlated to the clinical outcome from said female mammal.

24. The method of claim 1, further comprising a step of taking at least one biological sample from said female mammal.

25. The method of claim 1, wherein said method is an *in vitro* method and does not comprise any step of taking at least one biological sample from said female mammal.

26. The method of claim 1, wherein said biological sample is a breast tumor sample.

27. The method according to claim 1, wherein said mammal is selected in the group comprising humans, mice, rats, guinea pigs, monkeys, cats, dogs, pigs, horses and cows.

28. The method according to claim 27, wherein said mammal is a human.

29. The method of claim 1, further comprising the step of administrating a pharmaceutical treatment to a female mammal, for optimizing the clinical outcome of said female mammal in response to said treatment.

30. The method of claim 29, further comprising the step of taking at least two biological samples from the female mammal, wherein said two samples are taken before and after said pharmaceutical treatment.

31. The method according to claim 29, further comprising the step of selecting the pharmaceutical treatment that improves the clinical outcome of said female mammal.

32. The method of claim 1, further comprising the step of generating a printed report.

33. A Computer program comprising instructions for performing the method according to claim 1.

34. A recording medium for recording the computer program according to claim 33.

35. A method of assessing the clinical outcome of a female mammal suffering from breast cancer, comprising the steps of:

a) generating a metagene adjusted value underEGFR by comparing the expression level, in a biological sample from said female mammal and in a control, of at least one nucleic acid sequence selected in the group consisting of: SEQ ID No:1071 (NM_001033047), SEQ ID No:254 (nm_005581), SEQ ID No:6 (nm_003225), SEQ ID No:883 (nm_000125), SEQ ID No:543 (nm_005080), SEQ ID No:681 (nm_020974), SEQ ID No:63 (nm_001002295),

SEQ ID No:212 (nm_024852), SEQ ID No:635 (nm_001002029), SEQ ID No:535 (nm_003226), SEQ ID No:1125, SEQ ID No:109 (nm_000662), SEQ ID No:342 (nm_001846), SEQ ID No:927 (nm_004703), SEQ ID No:1124, SEQ ID No:124 (nm_014899), SEQ ID No:280 (nm_020764) (or SEQ ID No:1110 (NM_024522)), SEQ ID No:297 (nm_016463), SEQ ID No:791 (nm_016835), SEQ ID No:210 (nm_178840), SEQ ID No:827 (nm_152499), SEQ ID No:1064 (NM_000767), SEQ ID No:147 (nm_014675), SEQ ID No:323 (nm_001014443), SEQ ID No:106 (nm_004619), SEQ ID No:181 (nm_000848), SEQ ID No:376 (nm_057158), SEQ ID No:116 (nm_014034), SEQ ID No:252 (nm_000758), SEQ ID No:797 (nm_022131), SEQ ID No:911 (nm_000168), SEQ ID No:720 (nm_004726), SEQ ID No:889 (nm_000561), SEQ ID No:250 (nm_000930), SEQ ID No:179 (nm_004747), SEQ ID No:786 (nm_033388), SEQ ID No:177 (nm_015996), SEQ ID No:1047 (BC012900), SEQ ID No:301 (nm_004326), SEQ ID No:207 (nm_003940), SEQ ID No:936 (nm_003462), SEQ ID No:916 (nm_001453) (or SEQ ID No:1116 (NM_004040)), SEQ ID No:1052 (BX096026), SEQ ID No:159 (nm_000224), SEQ ID No:1096 (AK127274), SEQ ID No:28 (nm_021800), SEQ ID No:1054 (AK123264), SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (NM_053279)), SEQ ID No:825 (nm_024704), SEQ ID No:145 (nm_017786), SEQ ID No:491 (nm_004374), SEQ ID No:485 (nm_003834), SEQ ID No:1072 (AY007114), SEQ ID No:274 (nm_032108), SEQ ID No:258 (nm_080545), SEQ ID No:292 (nm_014371), SEQ ID No:803 (nm_183047), SEQ ID No:349 (nm_031946), SEQ ID No:1123, SEQ ID No:763 (nm_014585), SEQ ID No:438 (nm_001759), SEQ ID No:94 (nm_014315), SEQ ID No:845 (nm_001089), SEQ ID No:1084 (BX648964), SEQ ID No:734 (nm_025137), SEQ ID No:943 (nm_002141), SEQ ID No:1085 (NM_000720), and SEQ ID No:276 (nm_012202), fragments, derivatives or complementary sequences thereof;

b) generating a metagene adjusted value overEGFR by comparing the expression level, in a biological sample from said female mammal and in a control, of at least one nucleic acid sequences selected in the group consisting of SEQ ID No:405 (nm_000852), SEQ ID No:374 (nm_000212), SEQ ID

No:1122, SEQ ID No:598 (nm_000636), SEQ ID No:262 (nm_005194), SEQ ID No:1099 (BC066343), SEQ ID No:696 (nm_001428), SEQ ID No:1059 (AK091113), SEQ ID No:751 (nm_002423), SEQ ID No:1121, SEQ ID No:286 (nm_002417), SEQ ID No:244 (nm_199002), SEQ ID No:18 (nm_001880), SEQ ID No:121 (nm_014553), SEQ ID No:1107 (BC073775), SEQ ID No:103 (nm_003619), SEQ ID No:1118, SEQ ID No:42 (nm_000757), and SEQ ID No:1067 (AK123784), fragments, derivatives or complementary sequences thereof;

c) generating a score (S_C) from said metagene adjusted values using a mathematical method establishing a relation between the combined metagene values and the clinical outcome of said female mammal.

36. The method of claim 35, wherein the mathematical method used in step c) comprises a Cox regression or CART analysis.

37. The method of claim 35, wherein the mathematical method used in step c) is a Cox regression and the score (S_C) is generated according to the following formula: S_C = a x overEGFR + b x underEGFR, wherein "a" is comprised in the interval [-1,85; 0,81] and "b" is comprised in the interval [-3,86; 0,70].

38. The method of claim 37, wherein the score S_C is generated according to the following formula: S_C = – 1.33 x over EGFR – 2.28 under EGFR.

39. The method of claim 35, wherein the step a) of generating a metagene adjusted value underEGFR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 10 nucleic acid sequences selected in said group.

40. The method of claim 35, wherein the step a) of generating a metagene adjusted value underEGFR is obtained by comparing the

expression level, in a biological sample from said female mammal and in a control, of at least 20 nucleic acid sequences selected in said group.

41. The method of claim 40, wherein the step a) of generating a metagene adjusted value underEGFR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 24 nucleic acid sequences selected in said group.

42. The method of claim 41, wherein the step a) of generating a metagene adjusted value underEGFR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 37 nucleic acid sequences selected in said group.

43. The method of claim 35, wherein said metagene adjusted value underEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the nucleic acid sequence consisting of : SEQ ID No:681 (nm_020974).

44. The method of claim 35, wherein said metagene adjusted value underEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 24 nucleic acid sequences selected in the group consisting of : SEQ ID No:1071 (nm_001033047); SEQ ID No:254 (nm_005581); SEQ ID No:6 (nm_003225); SEQ ID No:883 (nm_000125); SEQ ID No:543 (nm_005080); SEQ ID No:681 (nm_020974); SEQ ID No:63 (nm_001002295); SEQ ID No:212 (nm_024852); SEQ ID No:635 (nm_001002029); SEQ ID No:535 (nm_003226); SEQ ID No:1125); SEQ ID No:1124; SEQ ID No:297 (nm_016463); SEQ ID No:791 (nm_016835); SEQ ID No:827 (nm_152499); SEQ ID No:207 (nm_003940); SEQ ID No:916 (nm_001453) (or SEQ ID No:1116 (nm_004040)); SEQ ID No:1052 (BX096026) ; SEQ ID No:159 (nm_000224); SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (NM_053279)); SEQ ID No:845

(nm_001089); and SEQ ID No:1085 (NM_000720), fragments, derivatives or complementary sequences thereof.

45. The method of claim 44, wherein said metagene adjusted value underEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 37 nucleic acid sequences selected in the group consisting of : SEQ ID No:1071 (nm_001033047); SEQ ID No:254 (nm_005581); SEQ ID No:6 (nm_003225); SEQ ID No:883 (nm_000125); SEQ ID No:543 (nm_005080); SEQ ID No:681 (nm_020974); SEQ ID No:63 (nm_001002295); SEQ ID No:212 (nm_024852); SEQ ID No:635 (nm_001002029); SEQ ID No:535 (nm_003226); SEQ ID No:1125; SEQ ID No:1124; SEQ ID No:297 (nm_016463); SEQ ID No:791 (nm_016835); SEQ ID No:827 (nm_152499); SEQ ID No:207 (nm_003940); SEQ ID No:916 (nm_001453) (or SEQ ID No:1116 (nm_004040)); SEQ ID No:1052 (BX096026) ; SEQ ID No:159 (nm_000224); SEQ ID No:25 (nm_012391) (or SEQ ID No:1108 (NM_053279)); SEQ ID No:845 (nm_001089); SEQ ID No:1085 (NM_000720) ; SEQ ID No:109 (nm_000662); SEQ ID No:342 (nm_001846); SEQ ID No:927 (nm_004703); SEQ ID No:280 (nm_020764) (or SEQ ID No:1110 (NM_024522)); SEQ ID No:210 (nm_178840); SEQ ID No:181 (nm_000848); SEQ ID No:116 (nm_014034); SEQ ID No:250 (nm_000930); SEQ ID No:177 (nm_015996); SEQ ID No:825 (nm_024704); SEQ ID No:145 (nm_017786); and SEQ ID No:276 (nm_012202), fragments, derivatives or complementary sequences thereof.

46. The method of claim 35, wherein the step b) of generating a metagene adjusted value overEGFR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 5 nucleic acid sequences selected in said group.

47. The method of claim 46, wherein the step b) of generating a metagene adjusted value overEGFR is obtained by comparing the expression

level, in a biological sample from said female mammal and in a control, of at least 10 nucleic acid sequences selected in said group.

48. The method of claim 47, wherein the step b) of generating a metagene adjusted value overEGFR is obtained by comparing the expression level, in a biological sample from said female mammal and in a control, of at least 12 nucleic acid sequences selected in said group.

49. The method of claim 35, wherein said metagene adjusted value overEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the nucleic acid sequence consisting of : SEQ ID No: 1107 (BC073775) or SEQ ID No: 1099 (BC066343), fragments, derivatives or complementary sequences thereof.

50. The method of claim 35, wherein said metagene adjusted value overEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 5 nucleic acid sequences selected in the group consisting of : SEQ ID No:1122; SEQ ID No:598 (nm_000636); SEQ ID No:696 (nm_001428); SEQ ID No:1059 (AK091113); and SEQ ID No:121 (nm_014553), fragments, derivatives or complementary sequences thereof.

51. The method of claim 50, wherein said metagene adjusted value overEGFR is generated by comparing the expression level, in a biological sample from said female mammal and in a control, of the 12 nucleic acid sequences selected in the group consisting of : SEQ ID No:1122; SEQ ID No:598 (nm_000636); SEQ ID No:696 (nm_001428); SEQ ID No:1059 (AK091113); SEQ ID No:121 (nm_014553); SEQ ID No:262 (nm_005194); SEQ ID No:1099 (BC066343); SEQ ID No:751 (nm_002423); SEQ ID No:1121; SEQ ID No:286 (nm_002417); SEQ ID No:103 (nm_003619); and SEQ ID No:1118, fragments, derivatives or complementary sequences thereof.

52. The method of claim 35, further comprising the first step of quantifying in a biological sample from said female mammal the expression level of said nucleic acids sequences.

53. The method of claim 35, further comprising the step e) of comparing said score (S_C) from the biological sample with a baseline or a (S_C) from a control sample.

54. The method according to claim 35, wherein said score (S_C) from the biological sample is correlated to the clinical outcome of said female mammal.

55. The method of claim 35, further comprising a step of taking at least one biological sample from said female mammal.

56. The method of claim 35, wherein said method is an *in vitro* method and does not comprise any step of taking at least one biological sample from said female mammal.

57. The method of claim 35, wherein said biological sample is a breast tumor sample.

58. The method according to claim 35, wherein said mammal is selected in the group comprising humans, mice, rats, guinea pigs, monkeys, cats, dogs, pigs, horses and cows.

59. The method according to claim 58, wherein said mammal is a human.

60. The method of claim 35, further comprising the step of administrating a pharmaceutical treatment to a female mammal, for optimizing the clinical outcome of said female mammal in response to said treatment.

61. The method of claim 60, further comprising the step of taking at least two biological samples from the female mammal, wherein said two samples are taken before and after said pharmaceutical treatment.

62. The method according to claim 60, further comprising the step of selecting the pharmaceutical treatment that improves the clinical outcome of said female mammal.

63. The method of claim 35, further comprising the step of generating a printed report.

64. A Computer program comprising instructions for performing the method according to claim 35.

65. A recording medium for recording the computer program according to claim 64.

66. The method of claim 29 or 60, wherein the pharmaceutical treatment comprises the use of one or more taxane compounds.

67. The method of claim 66, wherein the female mammal has not responded to a previous anti-cancer treatment.

68. The method of claim 67, wherein said previous treatment comprise the use of one or more anthracyclin.

69. The method of claim 1 or 35 for identifying a female mammal that has not responded to a previous anti-cancer treatment.

70. The method of claim 69, wherein said previous treatment comprise the use of one or more anthracyclin.

71. A method of assessing the clinical outcome of a female mammal suffering from breast cancer, comprising the steps of:

a) generating a metagene adjusted value underER by comparing the expression level, in a biological sample from said female mammal and in a control, of at least two genes, e.g. by using nucleic acid sequences selected in the group of Affymetrix® Probe Sets, of table IX or XII, preferably table XII,

b) generating said metagene adjusted value underPR by comparing the expression level, in a biological sample from said female mammal and in a control, of at least two genes, e.g. by using nucleic acid sequences selected in the group of Affymetrix® Probe Sets, of table X or XIII, preferably table XIII,

c) generating said metagene adjusted value underEGFR by comparing the expression level, in a biological sample from said female mammal and in a control, of at least two genes, e.g. by using nucleic acid sequences selected in the group of Affymetrix® Probe Sets, of table XI or XIV preferably table XIV,

d) generating a score (S_C) from said metagene adjusted values using a mathematical method establishing a relation between the combined metagene values and the clinical outcome of said female mammal.

72. The method of claim 71, wherein the mathematical method used in step d) comprises a Cox regression or CART analysis.

73. The method of claim 71, wherein the mathematical method used in step d) is a Cox regression and the score (S_C) is generated according to the following formula: S_C = a x underER + b x underPR + c x under EGFR, wherein "a" is comprised in the interval [-6,26 ; +0,49] , "b" is comprised in the interval [-2,65; +0,29] and "c" is comprised in the interval [-6,69; + 1,65].

74. The method of claim 71, wherein the score S_C is generated according to the following formula: S_C = - 2.90279 x underER – 1.47423 x underPR – 4.17198 x under EGFR.

75. The method of claim 71, wherein the comparing of expression level at each step a), b) and c) is performed with at least 5 genes or nucleic acid sequences of each respective group.

76. The method of claim 71, wherein the comparing of expression level at each step a), b) and c) is performed with at least 10, preferably all, genes or nucleic acid sequences of each respective group.

77. The method of claim 71, further comprising the first step of quantifying in a biological sample from said female mammal the expression level of said genes or nucleic acids sequences.

78. The method of claim 71, further comprising the step e) of comparing said score (S_C) from the biological sample with a baseline or a score (S_C) from a control sample.

79. The method according to claim 71, wherein said score (S_C) from the biological sample is correlated to the clinical outcome from said female mammal.

80. The method of claim 71, further comprising a step of taking at least one biological sample from said female mammal.

81. The method of claim 71, wherein said method is an *in vitro* method and does not comprise any step of taking at least one biological sample from said female mammal.

82. The method of claim 71, wherein said biological sample is a breast tumor sample.

83. The method according to claim 71, wherein said mammal is a human.

84. The method of claim 71, further comprising the step of administrating a pharmaceutical treatment to a female mammal, for optimizing the clinical outcome of said female mammal in response to said treatment.

85. The method of claim 71, further comprising the step of taking at least two biological samples from the female mammal, wherein said two samples are taken before and after said pharmaceutical treatment.

86. The method according to claim 71, further comprising the step of selecting the pharmaceutical treatment that improves the clinical outcome of said female mammal.

87. The method of claim 71, further comprising the step of generating a printed report.

88. A Computer program comprising instructions for performing the method according to claim 71.

89. A recording medium for recording the computer program according to claim 88.

Figure 1

>201505_at

aagagtcagctgatgccagaaggaaagccgaaatgctacaaaatgaagcaaaaactcttttagctca
agcaaatagcaagctgcaactgctcaaagatttagaaagaaaatatgaagacaatcaaagatactta
gaagataaagctcaagaattagcaagactggaaggagaagtccgttcactcctaaaggatataagcc
agaaagttgctgtgtatagcacatgcttgtaacagaggag

>201710_at
cccctatgtccagtgcctggaagacggtggcctgcgggggggaccagggaccagcttttcatgcagga
gaaagcccggcagctcctgggccgcctgaagcccagccacacatctcggaccctcatcttgtcctgag
gtgttgagggtgtcacgagcccattctcatgtttacaggggttgtgggggcagagggggtctgtgaatctg
agagtcattcaggtgacctcctgcagggagccttctgccaccagcccctccccagactctcaggtgga
ggcaacagggccatgtgctgccctgttgccgagcccagctgtgggcggctcctggtgctaacaacaaa
gttccacttccaggtctgcctggttccctccccaaggccacagggagctccgtcagcttctcccaagccc
acgtcaggcctggcctcatctcagaccctgcttaggatgggggatgt

>203256_at
gtggccgtagcaacttggcggagacaggctatgagtctgacgttagagtggttgcttccttagcctttcag
gatggaggaatgtgggcagtttgacttcagcactgaaaacctctccacctgggccagggttgcctcaga
ggccaagtttccagaagcctcttacctgccgtaaaatgctcaaccctgtgtcctgggcctgggcctgctgt
gactgacctacagtggactttctctctggaatggaaccttcttaggcctcctggtgcaacttaatttttttttta
atgctatcttcaaaacgttagagaaagttcttcaaaagtgcagcccagagctgctgggcccactggccg
tcctgcatttctggtttccagaccccaatgcctcccattcggatggatctctgcgtttttatactgagtgtgcct
aggttgccccttatttttattttccctgttgcgttgctatagatg

>203685_at
gatggaaaggctcgctcaatcaagaaaattctgagactattaataaataagactgtagtgtagatactg
agtaaatccatgcacctaaacctttggaaaatctgccgtgggccctccagatagctcatttcattaagtttt
tccctccaaggtagaatttgcaagagtgacagtggattgcatttcttttggggaagctttcttttggtggttttgt
ttattataccttcttaagttttcaaccaaggtttgcttttgttttgagttactggggttatttttgttttaaataaaaat
aagtgtacaataagtgtttttgtattgaaagcttttgttatcaagattttcatacttttaccttccatggctcttttta
agattgatacttttaagaggtggctgatattctgcaacactgtacacataaaaaatacggtaaggatactt
tacatggttaaggtaaagtaagtctccagttggccaccattagctataatggcactttgtttgtgttgttggaa
aaagtcacattgccattaaactttccttgtctgtctagttaata

>203934_at
gtttcttctggaatgaataccctcatatctgtcctgatgtgatatgtctgagactgaatgcgggaggttcaat
gtgaagctgtgtgtggtgtcaaagtttcaggaaggattttacccttttgttcttccccctgtccccaacccact
ctcaccccgcaacccatcagtattttagttatttggcctctactccagtaaacctgattgggtttgttcactctc
tgaatgattattagccagacttcaaaattattttatagcccaaattataacatctattgtattatttagacttta
acatatagagctatttctactgatttttgcccttgttctgtccttttttcaaaaaagaaaatgtgtttttgtttggta

ccatagtgtgaaatgctgggaacaatgactataagacatgctatggcacatatatttatagtctgtttatgt
agaaacaaatgtaatatattaaagccttatatataatgaactttgtactattcacattttgtatcagtattatgt
agcataacaaaggtcataatgctttcagcaat

>204628_s_at
tgagccttggtgtgctgtatccatccatggggctgattgtatttaccttctacctcttggctgccttgtgaagga
attattcccatgagttggctgggaataagtgccaggatggaatgatgggtcagttgtatcagcacgtgtg
gcctgttcttctatgggttggacaacctcattttaactcagtctttaatctgagaggccacagtgcaattttattt
tatttttctcatgatgaggttttcttaacttaaaagaacatgtatataaacatgcttgcattatatttgtaaatttat
gtgtatggcaaagaaggagagcataggaaaccacacagacttgggcagggtacagacactcccac
ttggcatcattcacagcaagtcactggccagtggctggatctgtgaggggctctctcatgatagaaggct
atggggatagatgtgtggacacattggacctttcctgaggaagagggactgttcttttgtcccagaaaag
cagtggctccattggtgttgacatacatccaacattaaaagccacccc

>227262_at
ccactccaggaatcatgcttgctcccctgggccatttgcngttttgtgggcttctggagggttccccgccat
ccaggctggtctccctcccttaaggaggttggtgcccagagtgggcggtggcctgtctagaatgccgcc
gggagtccgggcatggtgggcacagttctccctgcccctcagcctgggggaagaagagggcctcgg
gggcctccggagctgggctttgggcctctcctgcccacctctacttctctgtgaagccgctgaccccagtc
tgcccactgaggggctagggctggaagccagttctaggcttccaggcgaaagctgagggaaggaag
aaactcccctccccgttccccttcccctctcggttccaaagaatctgt

>205890_s_at
gatcttaaagccacggagaagcctctcatcttatggcattgacaaagagaagaccatccaccttaccct
gaaagtggtgaagcccagtgatgaggagctgcccttgtttcttgtggagtcaggtgatgaggcaaaga
ggcacctcctccaggtgcgaaggtccagctcagtggcacaagtgaaagcaatgatcgagactaaga
cgggtataatccctgagacccagattgtgacttgcaatggaaagagactggaagatgggaagatgat
ggcagattacggcatcagaaagggcaacttactcttcctggcatcttattgtattggagggtgaccaccct
ggggatggggtgttggcaggggtcaaaaagcttatttcttttaatctcttactcaacgaacacatcttctgat
gatttcccaaaattaatgagaatgagatgagtagagtaagatttgggtgggatgggtaggatgaagtat
attgcccaactctatgtttctttga

>206324_s_at
tgggcccaaggaattcagaagagcttgcaggcaagccaggagaccctgggagctgtggctgtcttctg
tggaggaggctccagcattcccaaagctcttaattctccataaaatgggctttcctctgtctgccatcctca
gagtctggggtgggagtgtggacttaggaaaacaatataaaggacatcctcatcatcacggggtgaa
ggtcagagtaaggcagccttcttcacaggctgaggggttcagaaccagcctggccaaaaattacac
cagagagacagagtcctccccattgggaacagggtgattgaggaaagtgaaccttgggtgtgaggga
ccaatcctgtgacctcccagaaccatggaagcca

>206792_x_at
ggtgggaggactgcttaagctcaggagtttgagaccagcctgggcaacatagcaagacccccaccat
ctctgaaaaaaaaaattagccaggcatggtggtgtgcacctgagaatcccagctactcagaaggttga
gacaaaggggatcgcttgagcccaggagttggaggctgaagagagctatgactgcatcactgcactc
cagcctgggcaacacagcaagatcctgtctaaaaataaaaagaaaagagaaggaaaggaaaga
gacggggctctgaggccgagcacagtggcccatgcctataatcccagcactttgggaggctgaggca

ggtggatcacctgaggttaggagttcgagaccagcctggccaacatggtgaaaccccatctctactaa
aaatacaaaaattggctgggcatggtggcgggtgcctgtaatcccagctactggggaggctgaggca
ggagaatcacttgaattcaggaggtggaggttgcagtgagccgacatcatgccactgcactccagcct
ggggctgacagagcaagacac

>207270_x_at
tggctccgagactttcatgatcccattgtcgaggttgaggtgtccgtgttcccggccgggacgaccacag
cctccagcccccagagctccatgggcacctcaggtcctcccacgaagctgcccgtgcacacctggcc
cagcgtgaccagaaaggacagccccgaacccagcccacaccctggctccctgttcagcaatgtccg
cttcctgctcctggtcctcttggagctgcccctgctcctgagcatgctgggtgccgtcctctgggtgaacag
acctcagagaagctctagaagcaggcagaattggcccaagggtgagaaccagtagcatctgctgtcc
atcaaggccctgtgctgcaacagagcccctctggggactggaatgacctcctgaccatcaaggcctgc
aacagagcccctctgggggactggaatgacctcctgaccactccctcccgggctgctctctccaacatc
tcctggaatcctttgtgagcctc

>207571_x_at
gcttccttcaggttctagattcttgctacttagggcgggctggtttggacctaacatctcgcacgtgactccct
cagcctcagagccttgggatgcagagcagctggcagggttcctctcaatcctgcaaccccagctgtcc
caccggtggatgcagaggggaatccgaggccatcaaccttggtgacagcagcgcagtgccaatgct
gatcacactgcatgggagattttgttaacgtctgccacccccactctcacccccaagctctaagccccg
ggaggcctggactgtcttcctcatctctgtagcaccaagcctgatagatctgtatatggtaaacaggggttt
aaccacatgtggttaacatggattaatgtgggaatttggcttcaagaacacaaccttaggaccttgggcc
ccaaaagctggtggtgaaatgagaggagccaatttaagaagacccttatggagacctgaggctgc

>209138_x_at
tctctgggctccaggctgaggacgaggctgattattactgctgctcatatgcaggtagttacactgtggtttt
cggcggagggaccaaactgaccgtcctaggtcagcccaaggctgccccctcggtcactctgttcccgc
cctcctctgaggagcttcaagccaacaaggccacactggtgtgtctcataagtgacttctacccgggag
ccgtgacagtggcctggaaggcagatagcagccccgtcaaggcgggagtggagaccaccacaccc
tccaaacaaagcaacaacaagtacgcggccagcagctatctgagcctgacgcctgagcagtggaa
gtcccacagaagctacagctgccaggtcacgcatgaagggagcaccgtggagaagacagtggccc
ctacagaatgtt

>209791_at
ctccagccccagagctgaaaacaccaagtgcctatttgagggtgtctgtctggagacttagagtttgtcat
gtgtgtgtgtgtnnttggttaatgtgggtttatgggttttctttctttttttttttttttttttnnagtctacattaggggga
agtgagcgcctcccatgtgcagacagtgtgtctttatagatttttctaaggctttccccaatgatgtcggtaat
ttctgatgtttctgaagttcccaggactcacacacccgttcccatctcacttgcccacccagtgtgacaacc
ctcggtgtggatataccccgtggactcatggctcttccccacccccactttctataaatgtaggcctaga
atacgcttctctgttgcaaaactcagctaagttcctgcttccaccttgatgttgaaatatcttatgtaagaggg
caggggatgtcgtgaagatggc

>212398_at
ttaacactaattatcacgtctgacaaatgtgtatgtgtggtttcagttctgtgtacattttaaaggataatggtg
aacattttaatgggtttcccttgcccttccatatttaacctatttccacattctctctcactcacattttctcagtgt
gcccttctcttatctgccatgtccatagccataattccaccatcatacagatcaggcagtgtttaaaatgat

ggtaggtagcacagtggacagtctttgatcatcatgtagaatatggctatgaatcaggaaagagattag
aacatttaataatgtatgtacagctggtgcttagtttttttttaatctaaatttaattaccttattggatatttgatatt
tggttatttaatcacagtcatctttaacagcttacactgattggtgttttatctcctgtgatcctttgatggctttttt
gcctaccatttcacagaggtt

>215223_s_at
acctcagccctaacggtggtggagaacccaaaggggagttgctggaagccatcaaacgtgactttggt
tcctttgacaagtttaaggagaagctgacggctgcatctgttggtgtccaaggctcaggttggggttggctt
ggtttcnaataagnnaacnggggacacttacnaaattgctgcttgtccaaatcaggatccactgcaag
gaacaacaggccttattccnactgctggggattgatgtgtgggagcacgcttactaccttcagtataaaa
atgtcaggcctgattatctaaaagctatttggaatgtaatcaactgggagaatgtaactgaaagatacat
ggcttgcaaaaagtaaaccacgatcgttatgctga

>215729_s_at
gaaactatatgtatctcgtggatctgccagtaccagccttccaaatgaaactctttcagagttagagacac
ctgggaaatactcacttacaccaccaaaccactggggccacccacatcgatacctgcagcatctttagt
caagttggaggagaaagacaacacttggtctaagacacggcagcaagacatccctgcatantgttcc
agataaaaatgaaagctgntcacacccacttgcctccccaatctgttnaaacagcttcgtgtctagtatg
agctcagta

>215946_x_at
gcaacatgcaggttcctgctccagcgnggctcctggactggcnccngntgctgnccnnggggggtttca
atccaagcataattcagtgaagcatgtgtttggcagngggacccagctcacngttttaggtcagcccaa
gncnaccccntcggtcantctgttcctgccgtcctntgaggagcnccaagccaacaaggccacactg
gtgtgtctcatgaatnacttnnnncngggaatcttganggtgacctggaaggcagatggtacccncatc
acccagngcgtggagangaccacgccnctccaaacagagcaacancaagtacatggccagcagc
tacctgagcctgacgcccgagcagtggaggtcccgcagaagctacagctgccaggtcatgcatgaa
gggagcactgcagagaagacggtggcccctgcagaatgttcataggttcccagcccccaccccacc
cacaggggcctggagctgcaggatcccaggggaggcgtc

>217865_at
tcccctgcaagcatgttttccacaaatcctgcgtggatccctggcttagtgaacattgtacctgtcctatgtg
caaacttaatatattgaaggccctgggaattgtgccgaatttgccatgtactgataacgtagcattcgatat
ggaaaggctcaccagaacccaagctgttaaccgaagatcagccctcggcgacctcgccggcgaca
actcccttggccttgagccacttcgaacttcggggatctcacctcttcctcaggatggggagctcactccg
agaacaggagaaatcaacattgcagtaacaaaagaatggtttattattgccagttttggcctcctcagtg
ccctcacactctgctacatgatcatcagagccacagctag

>243209_at
ggtagtggaccaatggcaaccctctgtgggaacaaggggccgaggccacgctgcctgcatctcgtgc
tggggacctgcatgcgccagcaccagggcttggactggatcttactcagtccatggtgcccagcctctg
ccccaacatgccctntgcatgtgaccgtcatgccctggatggagccactcctggctcaccccacctgca
ctgcactgtccccagagagccacccctccacccactcagagacagctgtggagagggccaggaga
atgggattaccctatgaccaaggagacatgggaagaagccctccttccttccacgatcgaggttccgcc
atcaactcggttctcggatatgcaagtacctcactttgttaac

>221651_x_at
gttatcctgtcacttttggccaggggaccaagctggagatcaaacgaactgtggctgcaccatctgtcttc
atcttcccgccatctgatgagcagttgaaatctggaactgcctctgttgtgtgcctgctgaataacttctatc
ccagagaggccaaagtacagtggaaggtggataacgccctccaatcgggtaactcccaggagagtg
tcacagagcaggacagcaaggacagcacctacagcctcagcagcaccctgacgctgagcaaagc
agactacgagaaacacaaagtctacgcctgcgaagtcacccatcagggcctgagctcgcccgtcac
aaagagcttcaacaggggagagtgttagagggagaagtgcccccacctgctcctcagttccagcctg
acccccctcccatcctttggcctctgaccctttttccacaggggacctacccctattgcggtcctccagctcat
ctttcacctcacccccctcctcctccttggctttaattatgc

>202036_s_at
gacaaaccatttccaacagcaacacagccactaaaacacaaaaaggggattgggcggaaagtg
agagccagcagcaaaaactacattttgcaacttgttggtgtggatctattggctgatctatgcctttcaact
agaaaattctaatgattggcaagtcacgttgttttcaggtccagagtagtttctttctgtctgctttaaatggaa
acagactcataccacacttacaattaaggtcaagcccagaaagtgataagtgcagggaggaaaagt
gcaagtccattatgtaatagtgacagcaaaggcccaggggagaggcattgccttctctgcccacagtct
ttccgtgtgattgtctttgaatctgaatcagccagtctcagatgccccaaagtttcggttcctatgagcccgg
ggcatgatctgatccccaagacatg

>204470_at
tatgattaactctacctgcacactgtcctattatattcattctttttgaaatgtcaaccccaagttagttcaatct
ggattcatatttaatttgaaggtagaatgttttcaaatgttctccagtcattatgttaatatttctgaggagcctg
caacatgccagccactgtgatagaggctggcggatccaagcaaatggccaatgagatcattgtgaag
gcaggggaatgtatgtgcacatctgttttgtaactgtttagatgaatgtcagttgttatttattgaaatgatttca
cagtgtgtggtcaacatttctcatgttgaaactttaagaactaaaatgttctaaatatcccttggacattttatg
tctttcttgtaaggcatactgccttgtttaatggtagttttacagtgtttctggcttagaaca

>209848_s_at
gtatcgatatggttccttttccgtcaccctggacattgtccagggtattgaaagtgccgagatcctgcaggc
tgtgccgtccggtgaggggatgcatttgagctgactgtgtcctgccaaggcgggctgcccaaggaag
cctgcatggagatctcatcgccagggtgccagccccctgcccagcggctgtgccagcctgtgctaccc
agcccagcctgccagctggttctgcaccagatactgaagggtggctcggggacatactgcctcaatgt
gtctctggctgataccaacagcctggcagtggtcagcacccagcttatcatgcctggtcaagaagcag
ggggccttgggcaggttccgctgatcgtgggcatcttgctggtgttgatggctgtggtccttgcatctctgat
atataggcgcagacttatgaagcaagacttctccgtaccccagttgccacatagcagcagtcactggct
gcgtctaccccgcatcttctgctcttgtcccattggtgagaa

>214370_at
ggtattgatgactttattattctgcaggtacatgtccaggggcccagcctctgggcccagtaactcagcta
ctctttgtggctttcttcatggctttttttgtgggctgccacgcccatctttatcaccagaatgaggaactcctg
gaagttaactgcaccatcagtgttgatatccaactctttgaaccagacgtctgcacccttttttcctgatatac
tgaggacactcggtctctagcaatttcttcaggtcatccctgtagacggcatggaaattcccctttatcagg
gagtacttgtggtagacgtcgatgatagagttcaaggctttctccagctcggtcaacatgatgc

>216598_s_at
ctcgctcagccagatgcaatcaatgccccagtcacctgctgctataacttcaccaataggaagatctca
gtgcagaggctcgcgagctatagaagaatcaccagcagcaagtgtcccaaagaagctgtgatcttca
agaccattgtggccaaggagatctgtgctgaccccaagcagaagtgggttcaggattccatggaccac
ctggacaagcaaacccaaactccgaagacttgaacactcactccacaacccaagaatctgcagcta
acttattttcccctagctttccccagacatcctgttttatttattataatgaattttgtttgttgatgtgaaacattat
gccttaagtaatgttaattcttatttaagttattgatgttttaagtttatctttcatggtactagtgtttttagataca
gagacttggggaaattgcttttcctcttgaaccacagttctaccc

>202246_s_at
gatgtttcgtcgaaagcctctcttctgtggaaactctgaagccgaccagttgggcaaaatctttgacctgat
tgggctgcctccagaggatgactggcctcgagatgtatccctgccccgtggagcctttccccccagagg
gccccgcccagtgcagtcggtggtacctgagatggaggagtcgggagcacagctgctgctggaaatg
ctgacttttaacccacacaagcgaatctctgcctttcgagctctgcagcactcttatctacataaggatgaa
ggtaatccggagtgagcaatggagtggctgccatggaaggaagaaaagctgccatttcccttctggac
actgagagggcaatctttgcctttatctctgaggctatggagggtcctcctccatctttctacagagattactt
tgctgccttaatgacattcccctcccacctctccttttgaggcttctccttctccttcccatttctctacactaag
g

>202519_at
ctccctctgctgacttaatgtcgtgattctgtttcttcagatatttaaggctgttaggttgtgtgagccttgaagt
gtgtgtgtgtgtcccagcgactgtccactgtccaggagatgcatgtctttgtattggagatatttctgtaactc
attctcttggtgctcacgattgccatggccatagggccacagtgccgtatctgctgcagacatgattgtttct
tgttctagaggttttcttgttttcgaatcttgcctgatgaatccagccagaccaaggggcctagatttgacctc
tgtcctgggctcctgggccaggtgcaggaacatctgaggccactctgctggccacctccagtgggtgct
gaccacaggatgggctttgtttacactcattttcaccctgattcttgcccccactttcataaaaga

>221671_x_at
caacaccgtgacaattggcctccgggggccactttcggcggagggaccaaggtggagatcaaacga
actgtggctgcaccatctgtcttcatcttcccgccatctgatgagcagttgaaatctggaactgcctctgttgt
gtgcctgctgaataacttctatcccagagaggccaaagtacagtggaaggtggataacgccctccaat
cgggtaactcccaggagagtgtcacagagcaggacagcaaggacagcacctacagcctcagcag
caccctgacgctgagcaaagcagactacgagaaacacaaagtctacgcctgcgaagtcacccatca
gggcctgagctcgcccgtcacaaagagcttcaacaggggagagtgttagagggagaagtgccccca
cctgctcctcagttccagcctgacccctcccatcctttggcctctgaccctttttccacaggggacctacc
cctattgcggtcctccagctcatctttcacctcacccccctcctcctccttggctttaattatgc

>210002_at
gtgtggctgatctgaagtcagtcggaatttgtaaacagggtagcaaacaagatatttttcttccatgtatac
aataatttttttaaaaagtgcaatttgcgttgcagcaatcagtgttaaatcatttgcataagatttaacagcat
ttttttataatgaatgtaaacatttttaacttaatggtacttaaaataatttaaaagaaaaatgttaacttagaca
ttcttatgcttcttttacaactacatcccatttttatatttccaattgttaaagaaaaatatttcaagaacaaatctt
ctctcaggaaaattgcctttctctatttgttaagaatttttatacaagaacaccaatatacccccttttattttact
gtggaatatgtgctggaaaaattgcaacaacactttactacctaacggatagcatttgtaaatactctagg
tatctgtaaacactctgatgaagtctgtatagtgtgactaacccacaggcaggttggtttacattaatttttttt
ttgaatgggatgtcctatggaaacctatttcaccag

>224795_x_at
atcttcccgccatctgatgagcagttgaaatctggaactgcctctgttgtgtgcctgctgaataacttctatc
ccagagaggccaaagtacagtggaaggtnggataacgccctccaatcgggtaactncccaggaga
gtgtcacagagcaggacagcaanggacagcancctacagcctcagcagcaccctgacgctgagca
aagcagactacgagaaacacaaagtctacgcctgcgaagtcancccatcagggcctgagctcgccc
gtcacaaagagcttcaacaggggagagtgttagagggagaagtgcccccacctgctcctcagttcca
gcctgacccccctcccatcctttggcctctgacccttttccacaggggacctacccctattgcggtcctcca
gctcatctttcacctcacccccctcctcctccttggctttaattatgc

>223721_s_at
gaagattactacacattactgggatgtgatgaactatcttcggttgaacaaatcctggcagaatttaaagt
cagagctctggaatgtcacccagacaagcatcctgaaaaccccaaagctgtggagacttttcagaaa
ctgcagaaggcaaaggagattctgaccaatgaagagagtcgagcccgctatgaccactggcgaag
gagccagatgtcgatgccattccagcagtgggaagctttgaatgactcagtg

>225883_at
gagactagtgatgacctctgttctgaccccagccctgtctgtcctcagggccgatggcttcaagtgtggttc
tgactggaccaaagctgtgttcagcccggacagaagctatgcactggcaggctcctgtgatggggccc
tttacatctgggatgtggacaccgggaaactggagagcagactacagggaccccattgcgctgccgtc
aacgccgtggcctggtgctactccgggagccacatggtgagcgtggaccagggcaggaaggttgtgc
tctggcagtagggccacgacctgcctgcctgggctggagctcttgcccgaagcctgaagcttccttcgg
cgccatgcaggggttggggttgggactggagctggccttgggatttaatggggaagaaggcctggcag
gacctggcctgtttgtttaaaaatgaagtatgggttgggggattacgctagtttttctttgtat

>225911_at
agattttcatcgggtgcattctctctgcttcgtgtgtgacaagttatcttggctgctgagaaagagtgccctgc
cccacaccggcagacctttccttcacctcatcagtatgattcagtttctcttatcaattggactctcccaggtt
ccacagaacagtaatatttttttgaacaataggtacaatagaaggtcttctgtcatttaacctggtaaaggc
agggctggaggggggaaaataaatcattaagcctttgagtaacggcagaatatatggctgtagatccatt
tttaatggttcatttcctttatggtcatataactgcacagctgaagatgaaaggggaaaataaatgaaaatt
ttacttttcgatgccaatgatacattgcactaaactgatggaagaagttatccaaagtactgtataacatctt
gttta

>204418_x_at
ggatgccttcccaaacctgaaggacttcatctcccgatttgagggcttggagaagatctctgcctacatg
aagtccagccgcttcctcccaagacctgtgttcacaaagatggctgtctggggcaacaagtagggcctt
gaaggcaggaggtgggagtgaggagcccatactcagcctgctgcccaggctgtgcagcgcagctgg
actctgcatcccagcacctgcctcctcgttcctttctcctgtttattcccatctttactcccaagacttcattgtcc
ctcttcactcccctaaacccctgtcccatgcaggccctttgaagcctcagctacccactatccttcgtgaa
catcccctcccatcattacccttccctgcactaaagccagcctgaccttccttcctgttagtggttgtgtctgc
tttaaagggcctgcctggccctcgcctgtggagctcagccccgagctgtccccgtgttgcatgaaggag
cagcattgactggtttacaggccctgc

>226362_at
gaggtctcattgtatctcacagaattttgatttgcctactcacccacaggagacatttctgggacagtagca
acataaggtccatcccaaaactttggctcattatcattaatatcctgcactttgatgatgaattctgattcagg
ctccaggggctttctggtttctatgtccacagcctgagcacgaagagtgtagaaaggtttttcttctctatcta
ggctccttattgcatgaatgtcccctgtggtttcatcaatggtaaaaacggtgccagcgccatctcctgag
agggtgtatttcgcagtgccct

>226373_at
gaggcctgaggggcagattggttcatgcctggggtggaggctaagcctggacacagtcangggcgg
ggctggccagctgtgcgagaacacaagccacgcctgcgatggtgctgcacccctggtgtccctgcctg
gccctcctctggtcacttcaaacatgccctcaggcttggggatgcccttcttccattcccagacagcagtgt
gagggtgcagggaccaagatgtcaagctggccgtggagtcgagctggccgacggacccctttcacct
ggttatgagccgacttctttggtgttctggcctctggcctgggaaaagggcaggagcctagaggaggaa
aaaggctggtggacagacccagggtggtcttcaagcctggcccagtgagagtgaggcccccgcacg
caagcctcagccactcccaggggcctttgcagcgtctttttaacctcagaaaatttctcaatctatgtgattt
gtgtaatactaatgagctttgggca

>205009_at
agacagagacgtgtacagtggccccccgtgaaagacagaattgtggttttcctggtgtcacgccctccc
agtgtgcaaataagggctgctgtttcgacgacaccgttcgtggggtcccctggtgcttctatcctaatacc
atcgacgtccctccagaagaggagtgtgaattttagacacttctgcagggatctgcctgcatcctgacgg
ggtgccgtccccagcacggtgattagtcccagagctcggctgccacctccaccggacacctcagaca
cgcttctgcagctgtgcctcggctcacaacacagattgactgctctgactttgactac

>205186_at
cccctacgcaccaggatggacgtggtgcacctccaggagcagttagacttaaagctgcagcagcggc
aggccagggaaacaggcatctgccctgtccgcaggaactctactcacagtgttttgatgagttgatccg
ggagtcaccatcaactgtgcggagaggggctgctgctgctgcagtcgggacgagatccgcatgacc
atcgctgcctaccagaccctgtacgagagcagcgtggcgtttggcatgaggaaggcactgcaggctg
agcaggggaagtcagacatggagaggaaaatcgcagaattggagacggaaaagagagacctgg
agaggcaagtgaacgagcagaaggcaaaatgtgaagccactgagaagcgggagagcgagagg
cggcaggtggaggagaagaagcacaatgaggagattcagttcctgaagcgaacaaatcagcagct
gaaggcccaactggaaggcattattgcaccaaagaagtgataatttccacatgattaatttccaacaag
acacttgggagttatttactgtgttcctctggcag

>227425_at
gtggagattggatcaactgagttgtgttatttttgttttaagtcaccttgtgcgagaacctcaaccgcacctat
cttgggaaccgggtataccctttctttagcactgctatccttttttgtcttcagcacaaataagatgttcaagtg
agccagaagcaagaagagccatttagtcttcatagctattggctaagagagataatgagctgatggtct
attttaaccttgaaagtaaaaatgtatatttttttcactacaagtacattgaacagtaaaaagtgtaacagca
aagtatatatttgatgccttctgtctttttcatgataatgtgctaacaagttgtgttaatattttactcagccagagt
ctcattcatttgctaagcattgggaacattatgtatattgaccttaaacataggtgtcctatatttggattgtga
cttgtagactcaagctaaccttactgcctctttttcacacttgttgaaangtctgtgaagaacatagttaaag
atctccaactttggaaaatatacatgatgtgaaactggggtgctatgt

>227515_at
agaagctctatgtcaccttgtaggaagccatctgaatttgactcagcatcatatgctgggcctgtgctgca
cttgagctttctttccagccgccaggagggatggttgtcctggccaacagcgaatttaatgggtggtcaaa
agaagatgtgcctgcatagaaggctgaaagagggagctggaagtaggaggagacaaggagatagt
gacaccagtatcttgtcacagaaatcaaggaaggcaacagtttaaaggaggggagggagatcagtg
gtgccaactggctcagaagaggtttgctgggttaaatctcaagatgagcccatgagatctggtggttagc
agaggcctttcctgacaaccccattttaagcacctcttccccatcactctatctcatcactgttactgtctgtgt
agccctgtcatgagctaaaatcatcctgtttgtttattgctagcttatgtgtctaaagtatactccttgagaac
agcaactttgtctttttcacctttgtactgagtc

>227550_at
tgttgcatttttcactcttaacccgagggtgtgtttcagcttatgttcgttctgtttcatgcaggtttatagcacgg
tagagtagaaggcggcttctgatttttagggtatttttagaattcattcctgagtgaggggttcagacaccca
gtctcctcggaacaggggtgaggggtcgactgagctttgttgagaagcctccagttaggcttcgggcgg
gtctccatgttatattgtgtgtttactgagcttcccactggtagaagatgacacatttgtccatcgtcctgtgtat
ctgctttccagaggacaccggagcattctcctggggtcactcccacatggctgcctcacatagctgttttg
caacagcctttattgccaacacccttggaagaacatctcctgtagaagtacaagttttgcaatatggacttt
ggaagtggttttgtgatgttcgaattttctgttagcctattgctcaagcactacagaatgtacatagtccctttg
gcca

>206107_at
gcgaatacccggaagcaagatgcacaggcgggcggcttcccacacacccgtcacaagacgcgga
catgcaggtctcggcgcgagctctgccccgtccaagagcctctccgctgtcgcccagtgtgagcctgga
agaggacccaagagagtgccgtgctgaggctgcctcgaggtcactgccttccggagctgcgcctattc
ctccctcgccaaacgcgttccagaatttgtccacaggtgcgccggcacctgctttcccacctcgaggcc
gcggcctcccccccgatttatagacaactctgacattgtcaccccactgacgaggcccgattccatagg
gtggatccttgccaggc

>206289_at
ccatctggaccataatagggacaccaaaacaaacccaaattggtaaaaagaataatcaaaaagaa
gacattatccggttaagagtctgtgctggttgccacnnaagagagaacagttgtccaggatgctggctg
gtggaacaacctgctggcccgaaacaaggctgccaggtgtggaacagcccatgctgactggggaca
tacacttgcatctttgttgaaagcagaagaagacagacccttccccaccttccttacctcctcttcccccat
taaggcagctcatccaagcttgtatttaactgaataaatgagtagacattgtggacctcacaagattattta
attcttaagatgtgtagaccttgatggtaggtgtgacatgttagtttttcttacttgcatttatttaagacacact
gttacagagatactgttgtcaccttctggggcacggtctttggggagaggggagtgcatttagacttatgtg
g

>206401_s_at
aggtggcagtggtccgtactccacccaagtcgccgtcttccgccaagagccgcctgcagacagcccc
cgtgcccatgccagacctgaagaatgtcaagtccaagatcggctccactgagaacctgaagcaccag
ccgggaggcgggaaggtgcaaatagtctacaaaccagttgacctgagcaaggtgacctccaagtgt
ggctcattaggcaacatccatcataaaccaggaggtggccaggtggaagtaaaatctgagaagcttg
acttcaaggacagagtccagtcgaagattgggtccctggacaatatcacccacgtccctggcggagg
aaataaaaagattgaaacccacaagctgaccttccgcgagaacgccaaagccaagacagaccac
ggggcggagatcgtgtacaagtcgccagtggtgtctggggacacgtctccacggcatctcagcaatgt

ctcctccaccggcagcatcgacatggtagactcgccccagctcgccacgctagctgacgaggtgtctg
cctcc

>228528_at
gagggaccattttggtccacatcgacaccatcagtctnctcttagaggatgccnnngagtattcggcgtt
gattncngggcacccnaaatcagacttgccacctnggactgtcgaggtgcagaccctgggagcacc
actggcccatctcttacacaggctgaccgatttctcctggtgttcagagtctgtttttgtctagcaccatttga
aatcggttatgatgtaggggaaaagcagcagcctcgaagcctcatgccaactctgggcagcagcag
cctgtggtttcctggaagatggatgggcagagaatagggaaggaagatcatgcttttccctactaacttct
gtaactgcatgtatgatacattattgcagaggtaagagatagtttaatggatttttaaaaacaaattactata
atttatctgatgttctctagttgcattttgctgaaatgtagtgctgttctaaattctgtaaattgattgctgttgaatt
atctttctgttgagaagagtctattcatgcatcctgaccttaataaatact

>229150_at
tagaaacgacacataacaccgggaaaagccatgctgtcaagcatctcagaatctcaaccccagagg
cagtctttagggactggggtgggctcaggaagctgtattatttcaaaggacgcagtgtttccacagctca
gcccagagctgctggagttcacacggcccaccagaacctgctgacttcaaaatcatcaccaccaagg
ggccaggaggtggcgctgtggggcaacagaaaacagaaaacagctgctccttgaacttggattccca
gcagctcccccaagcgcccacaggggctgcattcctccctgctggacacctgtaactacaacattaaa
acagggcgcaagcaaggaacacaaacac

>208451_s_at
ccaccttgtgttctgctgaagtctgccagtgtgctgaggggaagtgccctcgccagcgtcgcgccctgga
gcggggtctgcaggacgaggatggctacaggatgaagtttgcctgctacaccccgtgtggagtacg
gcttccaggttaaggttctccgagaagacagcagagctgctttccgcctctttgagaccaagatcaccca
agtcctgcacttcaccaaggatgtcaaggccgctgctaatcagatgcgcaacttcctggttcgagcctcc
tgccgccttcgcttggaacctgggaaagaatatttgatcatgggtctggatggggccacctatgacctcg
agggacacccccagtacctgctggactcgaatagctggatcgaggagatgccctctgaacgcc

>209048_s_at
caaggagagtggctcgaccttgacctttctggctccagagagacgccctcctccattctcttaggctcc
aaccaaggctctnnnnnnnnnnngnaccattcccggagtaataaatccagttggagcagcagtgat
gagaagaggggatcgacacgttccgatcacaacaccagtaccagcacgaagagcctcctcccgaa
agagtctcggctggacaccttctgggactagcagtgaatcgggacacaaaccacccaccccattggg
agaaaaacccagacgccaggaaaagaagaaacaacaaaggcaggagaacagccactttcaga
cttgaaaatgacaaaaccctcagttgagcctnngnncncggcgcgggggctgctacactacaggac
acccagcatcggctttgactgcagactgttcacccacacgagccctgtgcttttggtgtaaataatgtaca
atttgtggatgtcattgaatctagaggactttccccttttt

>210108_at
agagaatatcactccgatggtcggtttctgactgtcacgctaagggcaactgtaaactggaataataatg
cactcgcaaccaggtaaacttagatacactagtttgtttaaaattatagatttactgtacatgacttgtaatat
actataatttgtatttgtaaagagatggtctatattttgtaattactgtattgtatttgaactgcagcaatatccat
gggtcctaataattgtagttccccactaaaatctagaaattattagtatttttactcgggctatccagaagta
gaagaaatagagccaattctcatttattcagcgaaaatcctctggggttaaaattttaagtttgaaagaac
ttgacactacagaaatttttctaaaatattttgagtcactataaacctatcatctttcc

>210272_at
tttcctataaacagaatcacacaatatttgatttttttttaaaactaagccttgctctgtctcccaggctggagt
gctgtggcgtgattttggttcactgcaacctccgccttccaagttcaagagattctcctgcctcagcttccaa
gtagctgggattacaggcatgtggtaccacgcctggctaattttcttgtatttttagtagggacatgttggcc
aggctggttgtgagctcctggcctcaggtgatccacacgcctcagtgtcccagagtgctgatattacagg
cgtaatatgtgatcttttgtgtctggttcctttcacgttgaacgctatttttgaggttcgtgcctgttgtagaccac
agtcacacactgctgtagtcttcccccatcctcattcccagctgcctcctcctactgtttccctctatcaaaa
agcctc

>211038_s_at
tggtgttccttgaatgcctatcttccttttgtgcctcggaacctctcacgcctgccacaagttactcttttccttg
gtagttctgaactttaaataaggtaatgcctgtaagaatgccataaatgctcaataattgtcatctgttattat
tttcatcagtaacatcatctgaatcatcagtattgtctgcttttaacagctgcatttttcattgtccaaatatagtc
acatacatttgaccattttataattattgaataataaattcgttctgctattttacaatgaaaaataatgctgca
gagagcatttttgcacatgtatcgtggcagatgtaggccagaggctcttctttatccatcctatggccaacc
tatgaatgtatacacgtttaatgagattttgccagcaatcaaagccttcagggaaaatgtccctagctcttt
actacatcagatcaaggactctggataattggcataacatcctggaatagctgaaacagagatattattc
tctgctgtcctctgttgtctttgtcttttcacgtcttaata

>212099_at
gtggggttgtttttgccataagcgaactttgtgcctgtcctagaagtgaaaattgttcagtccaagaaactg
atgttatttgatttatttaaaggctaaaatttgtttttttattctttgcacaattgtttcattgtttgacacttaatgcac
tcgtcatttgcatacgacagtagcattctgaccacacttgtacgctgtaacctcatctacttctgatgtttttaa
aaaatgacttttaacaaggagagggaaaagaaacccactaaattttgctttgtttccttgaagaatgtgg
caacactgttttgtgattttatttgtgcaggtcatgcacacagttttgataaagggcagtaacaagtattggg
gcctatttttttttttttnnccacaaggcattctctaaagctatgtgaaattttctctgcacctctgtacagagaat
acacctgcccctgtatatccttttttcccctccctcctcccagtggtacttctactaaat

>218211_s_at
gccacaccttcgcgaaacctgtggtggcccaccagtcctaacgggacaggacagagagacagagc
agccctgcactgttttccctccaccacagccatcctgtccctcattggctctgtgctttccactatacacagt
caccgtcccaatgagaaacaagaaggagcaccctccacatggactcccacctgcaagtggacagc
gacattcagtcctgcactgctcacctgggtttactgatgactcctggctgccccaccatcctctctgatctgt
gagaaacagctaagctgctgtgacttccctttaggacaatgttgtgtaaatctttgaaggacacaccgaa
gacctttatactgtgatcttttacccctttcactcttggctttcttatgttgc

>218692_at
aaaccggaggtatcttcaaaggcatggagacctggttccagtaaatgtcccaccagtggggtatagaa
agcatgctcatgaccctgccgtgtcgtctgaggtacccgttcttatcctagtggttcaggaagagaaaac
gcagtttgcactttcaagacagcttctctaaggctggcatgttatctccttgctttgctttttgccgttttaaaatg
tgtaattgttccagcattccaatggtcttgtgcatagcaggggactgtaaccaaaaataaacatgtatttgt
gtaattggtttgaagaagtcttgaatagctctttactgtcttacttggggttgataagatttgagtgtttgcaattt
tttactaaatgtagctccaagtcttaaatggctt

>220778_x_at

cccaccctctgatggacgaagcggtgccctcgctgggccatgcgccctggatcctgcggaccctgatg
aggcaccagctgactcgagtggctgtggacgtgggagccggcccctggggcaaccagaccgttgtct
tcctgggttctgaggcggggacggtcctcaagttcctcgtccggcccaatgccagcacctcagggacgt
ctgggcgtgtgtgtcaagtgggccacgcgtgcagggtgtgtgtccacgagcgacgatcgtggtggccc
cagcggcctgggcgttggctgagccgacgctggggcttccagaaggcccgggggcctccgaggtgc
cggttaggagtttgaaccccccccactctgcagagggaagcggggacaatgccggggtttcaggcag
gagacacgaggagggcctgcccggaagtcacatcggcagcagctgtctaaagggcttgggggcctg
gggggcggcgaaggtgggtggggcccctctgtaaatacggccccagggtggtgagagagtcccatg
cca

>200670_at
agcagagaagcctgtctgtacttcattcaaaaagccaaaatagagagtatacagtcctagagaatccc
tctatttgttcagatctcatagatgacccccaggtattgccttttgacatccagcagtccaaggtattgagac
atattactggaagtaagaaatattactataattgagaactacagcttttaagattgtacttttaagattgtactt
ttatcttaaaagggtggtagttttccctaaaatacttattatgtaagggtcattagacaaatgtcttgaagtag
acatggaatttatgaatggtctttatcatttctcttccccctttttggcatcctggcttgcctccagttttaggtcct
ttagtttgcttctgcaagcaacgggaacacctgctgagggggctctttccctcatgtatacttcaagtaaga
tcaagaatcttttgtgaaattatagaaatttactatgtaaatgcttgatggaatttttttcctgctagtgtagcttct
gaaaggtgctttctccatt

>201860_s_at
tcccagatacttcccattttggaagttttcaggacttggtctgatttcaggatactctgtcagatgggaagac
atgaatgcacactagcctctccaggaatgcctcctccctgggcagaagtggccatgccaccctgttttca
gctaaagcccaacctcctgacctgtcaccgtgagcagctttggaaacaggaccacaaaaatgaaag
catgtctcaatagtaaaagataacaagatctttcaggaaagacggattgcattagaaatagacagtata
tttatagtcacaagagcccagcagggctcaaagttggggcaggctggctggcccgtcatgttcctcaaa
agcaccttgacgtcaagtctccttccccctttccccactccctggctctcagaaggtattccttttgtgtacag
tgtgtaaagtgtaaatccttttttctttataaactttagagtagcatgagagaattgtatcatttgaacaactag
gcttcagcatatttatagcaatccatgttagtttttactttctgttgccacaaccctgttttatact

>224516_s_at
cctttgattctttccgaccatgaaatagtgcatagtttgcctggagaatccactcacgttcataaagagaat
gttgatggcgccgtgtagaagccgctctgtatccatccacgcgtgcagagctgccagcagggagctca
cagaaggggagggagcaccaggccagctgagctgcacccacagtcccgagactgggatccccca
ccccaacagtgattttggaaaaaaaaatgaaagttctgttcgtttatccattgcgatctggggagccccat
ctcgatatttccaatcctggctacttttcttagagaaaataagtccttttttctggccttgctaatggcaacag
aagaaagggcttctttgcgtggtcccctgctggtgggggtgggtccccaggggggccccctgcggcctg
ggcccccctgcccacggccagcttcctgctgatgaacatgctgtttgtattgttttaggaaaccaggctgttt
tgtgaataaaacgaatgcatgtttgtgtcacga

>227811_at
gtgttggaggccccatgaagagcgccctggactgctgagggtgggccaacagcccagagctcagga
canttggctttgggggggaaggaaacngaggcccagagaggggcaaccactggccaagggtcaccc
agcaagttttggctaagagcctggcctccagccccagcagtgtggcccagagcaggggccgactgcc
aaagtaaccatcatccatatgggccgtgtggtgatgctggcccggaaggcagaaagaggcagcatg
ggcactgccagggacagccacatcctgctggtctgcagcgtggtccaccccgcctctgcccagcctgt

ctacaccgtgtgagctgaatcgtgacttgcttcccacctcctttctctgtcctctcctgaggttctgcctgcag
cccccaggaggtgggcctgccccatcctagctggactcatggttcctaaataaccacgctcagaagct
ctgctaggacttaccccanccactgagtggcaggcgcatgagatttgtggctgttcctgatgctagt

>209604_s_at
ttcgacttgcatttttgcaggagcagtatcatgaagcctaaacgcgatggatatatgtttttgaaggcagaa
agcaaaattatgtttgccactttgcaaaggagctcactgtggtgtctgtgttccaaccactgaatctggacc
ccatctgtgaataagccattctgactcatatcccctatttaacagggtctctagtgctgtgaaaaaaaaaa
aatgctgaacattgcatataacttatattgtaagaaatactgtacaatgactttattgcatctgggtagctgt
aaggcatgaaggatgccaagaagtttaaggaatatgggagaaatagtgtggaaattaagaagaaac
taggtctgatattcaaatggacaaactgccagttttgtttcctttcactggccacagttgtttgatgcatta

>226506_at
taatcagtccagttccctgaggtttaagatcaaatataaattactctgcttttcgactcattcaggtagcattg
tacctgaacctgattgctacttttttcatcttaaatattatatttcctcatctaatctgccttcccctcatccacaga
catttggagaaggaaatgggagggtgtctgttatccctttctctttgctttgtccccgttgttagactggcagc
gtcagttgctcggtgggcttggttagagccgtgggtgaggcaggtggctggcggggacagggagagg
ctgagagggaagtggtggcatttactgctctgacacttccactgtccctgctggggatgctgggccaagg
cctgtggggcctgtgaactgcacagccaggagcaaggaacccactaaatactccgtcacctccatgtc
ccctctacagtgttaaatt

>225092_at
acaattctttcaactgggttttcagcataaatgggaactgatgtagaaggcaggatttagcccttctaggc
aaaagaaaagctcagttgggtttcacgagtgttcctgtgcttatattcagtctgtgcctacatgttctcatgc
atgtctaacctgatttacctcttacctgtaacctaccttatcatgtggcttttaattgncagtcactcngccattt
ctaagcagatatagtantacctttcagaactcacattggcaagtgtaaaaagatgacttaaggtgaagt
gaggacaaaatcacattctgcatactaacctatttttttctccctttaaggtgctaaacttgcacctcatgtcc
a
>200824_at

aggagggctcactcaaagcctcctgcctatacgggcagctccccaagttccaggacggagacctcac
cctgtaccagtccaataccatcctgcgtcacctgggccgcaccttgggctctatgggaaggaccagc
aggaggcagccctggtggacatggtgaatgacggcgtggaggacctccgctgcaaatacatctccct
catctacaccaactatgaggcgggcaaggatgactatgtgaaggcactgcccgggcaactgaagcct
tttgagaccctgctgtcccagaaccagggaggcaagaccttcattgtgggagaccagatctccttcgct
gactacaacctgctggacttgctgctgatccatgaggtcctagcccctggctgcctggatgcgttcccct
gctctcagcatatgtggggcgc

>201201_at
tccgtcgccgccaagatgatgtgcggggcgccctccgccacgcagccggccaccgccgagaccca
gcacatcgccgaccaggtgaggtcccagcttgaagagaaagaaaacaagaagttccctgtgtttaag
gccgtgtcattcaagagccaggtggtcgcggggacaaactacttcatcaaggtgcacgtcggcgacg
aggacttcgtacacctgcgagtgttccaatctctccctcatgaaaacaagcccttgaccttatctaactac
cagaccaacaaagccaagcatgatgagctgacctatttctgatcctgactttggacaaggcccttcagc
cagaagactgacaaagtcatcctccgtctaccagagcgtgcacttgtgat

>201231_s_at
ggcccaggccaatggttggggcgtcatggtgtctcatcgttcgggggagactgaagataccttcatcgct
gacctggttgtggggctgtgcactgggcagatcaagactggtgccccttgccgatctgagcgcttggcc
aagtacaaccagctcctcagaattgaagaggagctgggcagcaaggctaagtttgccggcaggaac
ttcagaaaccccttggccaagtaagctgtgggcaggcaagccttcggtcacctgttggctacacagacc
cctcccctcgtgtcagctcaggcagctcgaggcccccgaccaacacttgcaggggtccctgctagttag
cgccccaccgccgtggagttcgtaccgcttccttagaacttctacagaagccaagctccctggagccct
gttggcagctctagcttttgcagtcgtgtaatgggcccaagtcattgtttttctcgcctcactttccaccaagt
gtctagagtcatgtgagcct

>201710_at
cccctatgtccagtgcctggaagacggtggcctgcgggggggaccagggaccagcttttcatgcagga
gaaagcccggcagctcctgggccgcctgaagcccagccacacatctcggaccctcatcttgtcctgag
gtgttgagggtgtcacgagcccattctcatgtttacaggggttgtgggggcagagggggtctgtgaatctg
agagtcattcaggtgacctcctgcagggagccttctgccaccagcccctccccagactctcaggtgga
ggcaacagggccatgtgctgccctgttgccgagcccagctgtgggcggctcctggtgctaacaacaaa
gttccacttccaggtctgcctggttccctccccaaggccacagggagctccgtcagcttctcccaagccc
acgtcaggcctggcctcatctcagaccctgcttaggatgggggatgt

>203256_at
gtggccgtagcaacttggcggagacaggctatgagtctgacgttagagtggttgcttccttagcctttcag
gatggaggaatgtgggcagtttgacttcagcactgaaaacctctccacctgggccagggttgcctcaga
ggccaagtttccagaagcctcttacctgccgtaaaatgctcaaccctgtgtcctgggcctgggcctgctgt
gactgacctacagtggactttctctctggaatggaaccttcttaggcctcctggtgcaacttaatttttttttta
atgctatcttcaaaacgttagagaaagttcttcaaaagtgcagcccagagctgctgggcccactggccg
tcctgcatttctggtttccagaccccaatgcctcccattcggatggatctctgcgtttttatactgagtgtgcct
aggttgccccttatttttttatttccctgttgcgttgctatagatg

>204092_s_at
tgccctgaccccgatcagttaaggagctgtgcaataaccttcctagtacctgagtgagtgtgtaacttattg
ggttggcgaagcctggtaaagctgttggaatgagtatgtgattctttttaagtatgaaaataaagatatatg
tacagacttgtatttttttctctggtggcattcctttaggaatgctgtgtgtctgtccggcaccccggtaggcctg
attgggtttctagtcctccttaaccacttatctcccatatgagagtgtg

>209791_at
ctccagccccagagctgaaaacaccaagtgcctatttgagggtgtctgtctggagacttagagtttgtcat
gtgtgtgtgtgtnnttggttaatgtgggtttatgggttttctttctttttttttttttttttttnnagtctacattaggggga
agtgagcgcctcccatgtgcagacagtgtgtctttatagatttttctaaggctttccccaatgatgtcggtaat
ttctgatgtttctgaagttcccaggactcacacacccgttcccatctcacttgcccacccagtgtgacaacc
ctcggtgtggatatacccccgtggactcatggctcttccccacccccactttctataaatgtaggcctaga
atacgcttctctgttgcaaaactcagctaagttcctgcttccaccttgatgttgaaatatcttatgtaagaggg
caggggatgtcgtgaagatggc

>212398_at
ttaacactaattatcacgtctgacaaatgtgtatgtgtggtttcagttctgtgtacattttaaaggataatggtg
aacattttaatgggtttcccttgccctttccatatttaacctatttccacattctctctcactcacattttctcagtgt
gcccttctcttatctgccatgtccatagccataattccaccatcatacagatcaggcagtgtttaaaatgat
ggtaggtagcacagtggacagtctttgatcatcatgtagaatatggctatgaatcaggaaagagattag
aacatttaataatgtatgtacagctggtgcttagttttttttaatctaaatttaattaccttattggatatttgatatt
tggttatttaatcacagtcatctttaacagcttacactgattggtgttttatctcctgtgatcctttgatggctttttt
gcctaccatttcacagaggtt

>215223_s_at
acctcagccctaacggtggtggagaacccaaaggggagttgctggaagccatcaaacgtgactttggt
tcctttgacaagtttaaggagaagctgacggctgcatctgttggtgtccaaggctcaggttggggttggctt
ggtttcnaataagnnaacnggggacacttacnaaattgctgcttgtccaaatcaggatccactgcaag
gaacaacaggccttattccnactgctggggattgatgtgtgggagcacgcttactaccttcagtataaaa
atgtcaggcctgattatctaaaagctatttggaatgtaatcaactgggagaatgtaactgaaagatacat
ggcttgcaaaaagtaaaccacgatcgttatgctga

>215729_s_at
gaaactatatgtatctcgtggatctgccagtaccagccttccaaatgaaactctttcagagttagagacac
ctgggaaatactcacttacaccaccaaaccactggggccacccacatcgatacctgcagcatctttagt
caagttggaggagaaagacaacacttggtctaagacacggcagcaagacatccctgcatantgttcc
agataaaaatgaaagctgntcacacccacttgcctccccaatctgttnaaacagcttcgtgtctagtatg
agctcagta

>221477_s_at
tgttgaaatttctgtggggctctatttttgctttggctttctggtgagagagtgaggaagcattctttccttcacta
agtttgtctttcttgtcttctggatagattgnattttaagagactaagggaatttacaaactaaagattttagtc
atctggtggaaaaggagactttaagattgtttagggctgggcggggtgactcacatctgtaatcccagca
ctttgggaggccaaggcaggcagaacacttgaaggagttcaagaccagcgtggccaacgtggtgaa
accctgtctctactaaaaatacaaaaattgtttagctctgtttttcataatagaaatagaaaaggtaaaatt
gcttttcttctgaaaagaacaagtattgttcatccaagaagggtttttgtgactgaatcagcagtgcctgcc
ctagtcatagctgtgcttcaaaaacctcagcatgattagtgttggagcaaaacaaggaagcaaagcaa
atacngtttttgaaattctatctgttgcttg

>208456_s_at
gatccaaccattgaagattcttacacaaagcagtgtgtgatagatgacagagcagcccggctagatatt
ttggatacagcaggacaagaagagtttggagccatgagagaacagtatatgaggactggcgaaggc
ttcctgttggtcttttcagtcacagatagaggcagttttgaagaaatctataagtttcaaagacagattctca
gagtaaaggatcgtgatgagttcccaatgattttaattggtaataaagcagatctggatcatcaaagaca
ggtaacacaggaagaaggacaacagttagcacggcagcttaaggtaacatacatggaggcatcag
caaagattaggatgaatgtagatcaagctttccatgaacttgtccgggttatcaggaaatttcaagagca
ggaatgtcctccttcaccagaaccaacacggaaagaaaaagacaagaaaggctgccattgtgtca

>210453_x_at
ccggcgggtgacattcagccggcggttcggggcgacggactctccattccagaaccatggcccaattt
gtccgtaaccttgtggagaagaccccggcgctggtgaacgctgctgtgacttactcgaagcctcgattg

gccacattttggtactacgccaaggttgagctggttcctcccacccctgctgagatccctagagctattca
gagcctgaaaaaaatagtcaatagtgctcagactggtagcttcaaacagctcacagttaaggaagctg
tgctgaatggtttggtggccactgaggtgttgatgtggttttatgtcggagagattataggcaagcggggc
atcattggctatgatgtttgaagaccaatctttaacatctgattatatttgatttattatttgagtgttgttggacc
atgtgtgatcagactgctat

>214370_at
ggtattgatgactttattattctgcaggtacatgtccaggggcccagcctctgggcccagtaactcagcta
ctctttgtggctttcttcatggctttttttgtgggctgccacgcccatctttatcaccagaatgaggaactcctg
gaagttaactgcaccatcagtgttgatatccaactctttgaaccagacgtctgcacccttttttcctgatatac
tgaggacactcggtctctagcaatttcttcaggtcatccctgtagacggcatggaaattcccctttatcagg
gagtacttgtggtagacgtcgatgatagagttcaaggctttctccagctcggtcaacatgatgc