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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 December 2012 (06.12.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/163941 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/EP2012/060097

(22) International Filing Date:
30 May 2012 (30.05.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
00921/11 30 May 2011 (30.05.2011) CH

(71) Applicant *(for all designated States except US)*: **NEBION AG** [CH/CH]; Hohlstrasse 515, CH-8045 Zuerich (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LAULE, Oliver** [DE/DE]; Heimbachweg 5, 79713 Bad Säckingen (DE). **BLEULER, Stefan** [CH/CH]; Grepperstrasse 14, CH-6403 Küssnacht am Rigi (CH). **ZIMMERMANN, Philip** [CH/CH]; Bettlachstrasse 153a, CH-2540 Grenchen (CH).

(74) Agent: **RENTSCH PARTNER AG**; Fraumünsterstrasse 9, CH-8045 Zürich (CH).

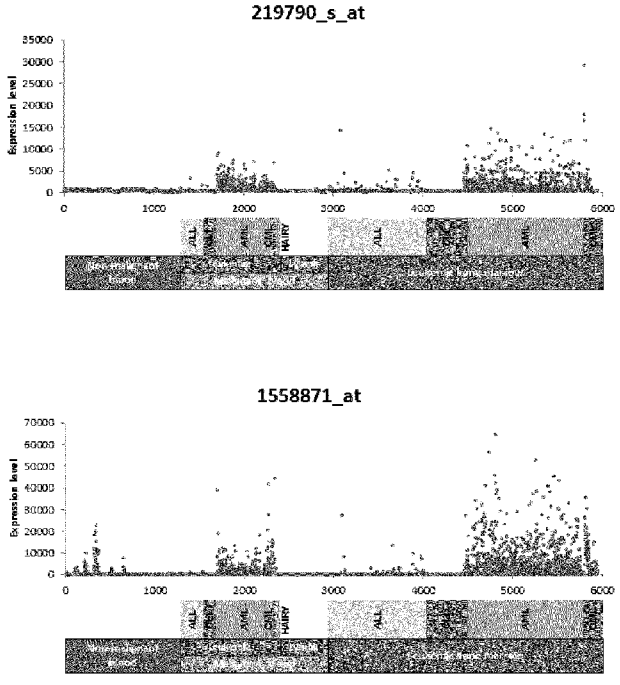
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

[Continued on next page]

(54) Title: MARKER FOR THE DETECTION AND CLASSIFICATION OF LEUKEMIA FROM BLOOD SAMPLES

Fig. 1(f)



(57) Abstract: The invention relates to a method for detecting leukemia from patient samples, preferably blood samples, and to preferably classify it into specific types of leukemia, comprising the steps of: (a) determining the expression level of at least one marker gene; and (b) comparing the expression level of the at least one marker gene to a reference to predict leukemia, preferably a specific type of leukemia; wherein the at least one marker gene corresponds to at least one of the genes listed in one or more of Tables I to VI and/or is represented by at least one probe set or at least one target sequence listed in one or more of Tables I to VI. The invention further relates to a kit to perform the method, to the use of the marker genes and to a reagent for detecting leukemia.

WO 2012/163941 A2

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

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

Marker for the detection and classification of leukemia from blood samples

TECHNICAL FIELD

The present invention is related to a method and a kit for detecting leukemia from patient samples and to preferably classify it into specific types and subtypes of leukemia. The invention further relates to the use of at least one marker gene for detecting leukemia and a reagent for detecting leukemia.

PRIOR ART

Leukemia is a disease of the bone marrow and is characterized by an abnormal proliferation of white blood cells. Leukemia belongs to the haematological neoplasms. It is subdivided into several clinically and pathologically classified subgroups. A first classification level divides it into its acute and chronic forms (acute leukemia and chronic leukemia). A second level of classification refers to the affected cell type, such as lymphoblastic or lymphocytic leukemia and myelogenous or myeloid leukemia. From these two levels, four main types of leukemia therefore have been defined: (a) Acute lymphoblastic leukemia (ALL), (b) Acute myeloid leukemia (AML), (c) Chronic lymphocytic leukemia (CLL), and (d) Chronic myeloid leukemia (CML).

Other types or subtypes of leukemia exist, such as T-cell prolymphocytic leukemia, Hairy cell leukemia, adult T-cell leukemia, and several others based on histopathological and staging characteristics.

Leukemia is usually diagnosed by physical examination, blood examination, bone marrow puncture, bone marrow biopsy and blood testing. A number of technical approaches exist to identify and classify leukemia from harvested body samples (bone marrow or blood), including complete blood count, blast cell count, hematocrit assays, microscopy examination (cell morphology), cell biochemical staining (immunophenotyping), flow cytometry, and cytogenetic tests. Once the medical diagnosis is made, the potential spread to other organs (metastasis) is usually evaluated using diagnostic imaging tools.

The diagnosis of leukemia therefore involves several steps that, in combination, provide sufficient information to correctly diagnose and classify leukemia such as to direct an appropriate therapy. While some of the methods can be performed on blood, several of them are more invasive, requiring bone marrow aspiration and/or biopsy. Samples of liquid bone marrow and bone marrow biopsies are subsequently examined by specialists, including a pathologist, a hematologist and/or an oncologist. In particular, they perform several tests involving microscopic examination, cytochemistry, flow cytometry, immunocytochemistry, cytogenetic testing, and immunophenotyping. The precise and reliable diagnosis of leukemia therefore involves several labor-intensive and expensive steps requiring the expertise from multiple persons. Additionally, they generally involve invasive steps that may cause pain to the patient, such as bone marrow biopsies. Finally, the combination of multiple diagnostic steps increases the time needed for the final diagnosis, which can require up to several days.

Recent developments in molecular measurement techniques have raised the hopes to detect diseases with minimally invasive methods, such as measuring nucleic acids, proteins or metabolites from a patient's blood samples. Ideally, a single diagnostic test measuring one or a plurality of molecular signatures from blood samples could allow identifying precisely the presence of a disease and its malignancy, progression, metastasis formation, and responsiveness to distinct treatment strategies.

EXPLANATION OF THE INVENTION

It is an objective of the invention to provide a method for detecting leukemia, preferably specific leukemia types, with a single, non-invasive test that can be carried out in a short time for immediate and reliable diagnosis.

The objective of the present invention is achieved by a method according to claim 1 for detecting leukemia from patient samples and to preferably classify it into specific types of leukemia. The method of the invention comprises the steps of: (a) determining the expression level of at least one marker gene; and (b) comparing the expression level of the at least one marker gene to a reference, wherein an up-regulated expression level of the at least one marker gene predicts leukemia, preferably a specific type or subtype of leukemia; wherein the at least one marker gene corresponds to at least one of the genes listed in one or more of Tables I to VI and/or is represented by at least one probe set or at least one target sequence (SEQ ID NOs) listed in one or more of Tables I to VI.

Within the present specifications of the invention, marker genes and contrasting marker genes (or contrasting genes) are represented by probe sets that match to a target DNA sequence (target sequence; cf. Table VIII) from said marker gene or contrasting gene. The identifiers of the probe sets of the present invention refer to the identifiers of Affymetrix Probe Sets from Affymetrix, Inc., USA, and are unique. The at least one marker gene can also be identified or is represented either by the probe set or by the target sequence.

With the present invention the expression levels of at least one marker gene in a patient's sample is determined. If the patient's sample is affected with leukemia or a specific type of leukemia the expression level of the marker gene is up-regulated compared to the sample if it would not be affected with the disease. To predict if leukemia or a specific type or subtype of leukemia is present the possibly up-regulated expression level of the at least one marker gene is compared to a reference. The reference can be a control sample, e.g. from a healthy person, where the expression level of the same marker gene is determined. In this case a significant deviation of an up-regulated expression level of the marker gene in the patient's sample to the expression level of the marker gene in a healthy sample predicts the disease.

Alternatively, the reference can be at least one internal reference gene within the patient's sample, which is invariant between leukemic and non-leukemic samples or which is down-regulated in leukemic samples, e.g. a contrasting gene of Table VII. Such a reference would avoid the need of a control sample from a healthy person. A patient sample is positive if the ratio of expression level between the at

least one marker gene and the internal reference gene or contrasting gene is beyond an empirically defined threshold, e.g. 10-fold. The threshold would have to be determined during clinical validation of a given kit for detecting leukemia from patient samples. However, there are other statistical methods that could be used to classify healthy from diseased samples.

Another alternative can be a template signature as reference, when determining the expression levels of at least two or more marker genes. The expression levels of the two or more marker genes (also called expression signature) is then compared with the typical expression signatures (template signature) of these genes in the blood of diseased and healthy patients, respectively, wherein the higher similarity of the signature of the tested sample to the typical signature of diseased patients predicts leukemia, preferably a specific type of leukemia.

The method is performed outside of the human body *in vitro*. The patient sample is preferably a blood, preferably peripheral blood mononuclear cells (PBMC), or bone marrow sample. In the case of a control sample as reference, the control is preferably a blood sample, preferably peripheral blood mononuclear cells (PBMC), or bone marrow sample of a healthy person.

In another aspect of the method of the invention ALL type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table I and/or is represented by at least one probe set or at least one target sequence listed in Table I.

In another aspect of the method of the invention AML type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table II and/or is represented by at least one probe set or at least one target sequence listed in Table II.

In another aspect of the method of the invention CLL type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table III and/or is represented by at least one probe set or at least one target sequence listed in Table III.

In another aspect of the method of the invention CML type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table IV and/or is represented by at least one probe set or at least one target sequence listed in Table IV.

In another aspect of the method of the invention hairy cell type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table V and/or is represented by at least one probe set or at least one target sequence listed in Table V.

In another aspect of the method of the invention multiple types leukemia are detected and the at least one marker gene corresponds to at least one of the genes listed in Table VI and/or is represented by at least one probe set or at least one target sequence listed in Table VI.

In another aspect of the method of the invention the expression levels of a group of at least two marker genes are determined and compared to a reference, wherein each marker gene of the group of at least two marker genes corresponds to one of the genes listed in Tables I to VI and/or is represented by one probe set or one target sequence listed in Tables I to VI. Each selected maker gene can be a marker gene specific for a different specific type of leukemia. With the method it is possible to identify the specific type of leukemia from one patient's sample.

In another aspect of the method of the invention at least one contrasting gene corresponding to at least one of the genes listed in Table VII and/or being represented by at least one probe set or at least one target sequence listed in Table VII is used to increase the discriminating power of the method for detecting leukemia from patient samples. The contrasting genes are genes that are down-regulated in the presence of leukemia or a specific type of leukemia. The contrasting genes can be used as the reference in step b or additionally to the reference in step b. The at least one contrasting gene of Table VII is preferably used to increase the discriminating power of the method for detecting a specific type of leukemia specified in Table VII.

The invention further relates to a kit for detecting leukemia from patient samples and to preferably classify it into specific types of leukemia, the kit comprising at least one reagent for determining the expression level of at least one marker gene and optionally at least one contrasting gene, wherein the at least one marker gene corresponds to at least one of the genes listed in one or more of Tables I to VI and/or is represented by at least one probe set or at least one target sequence

listed in one or more of Tables I to VI and wherein the at least one contrasting gene corresponds to at least one of the genes listed in Table VII and/or is represented by at least one probe set or at least one target sequence listed in Tables VII. The at least one reagent or the reagents can be the listed probe sets, or other probes that can be used to assess the expression level of the at least one marker gene. In particular, the reagents can be probes to measure nucleic acids, preferably gene transcripts (mRNA), as well as the products (preferably proteins) and regulators (preferably microRNAs and other small RNAs) of the at least one marker gene. The reagents are preferably oligonucleotides that match part of the sequence of the at least one marker gene and are used to determine the expression levels of the at least one marker gene, e.g. by RT-qPCR or microarrays.

Preferably, the kit comprises at least two, preferably three or more, marker genes and preferably one or more contrasting gene. With the kit the expression levels of several marker genes can be measured to determine an expression signature of the selected marker genes. By comparing the expression signature to a reference the specific type of leukemia can be identified from a patient's sample.

The invention further relates to the use of at least one marker gene and/or at least one contrasting gene for detecting leukemia from patient samples, preferably a specific type of leukemia, wherein the at least one marker gene corresponds to at least one of the genes listed in one or more of Tables I to VI and/or is represented by at least one probe set or at least one target sequence listed in one or more of Tables I to VI.

In another aspect of the inventive use, ALL type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table I and/or is represented by at least one probe set or at least one target sequence listed in Table I.

In another aspect of the inventive use, AML type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table II and/or is represented by at least one probe set or at least one target sequence listed in Table II.

In another aspect of the inventive use, CLL type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table III and/or is represented by at least one probe set or at least one target sequence listed in Table III.

In another aspect of the inventive use, CML type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table IV and/or is represented by at least one probe set or at least one target sequence listed in Table IV.

In another aspect of the inventive use, hairy cell type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table V and/or is represented by at least one probe set or at least one target sequence listed in Table V.

In another aspect of the inventive use, multiple types leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table VI and/or is represented by at least one probe set or at least one target sequence listed in Table VI.

The invention further relates to the a reagent for detecting leukemia from patient samples and to preferably classify it into specific types of leukemia, the reagent being suitable for determining the expression level of at least one marker gene, wherein the at least one marker gene corresponds to at least one of the genes listed in one or more of Tables I to VI and/or is represented by at least one probe set or at least one target sequence listed in one or more of Tables I to VI.

The reagent can be one of the listed probe sets, or other probes that can be used to assess the expression level of the at least one marker gene. In particular, the reagent can be a probe to measure nucleic acids, preferably gene transcripts (mRNA), as well as the products (preferably proteins) and regulators (preferably microRNAs and other small RNAs) of the at least one marker gene. The reagent is preferably an oligonucleotide that matches a part of the sequence of the at least one marker gene and are used to determine the expression levels of the at least one marker gene, e.g. by RT-qPCR or microarrays.

In another aspect of the invention, the reagent is for detecting ALL type leukemia and is suitable for determining the expression level of at least one marker gene, wherein the at least one marker gene corresponds to at least one of the genes

listed in Table I and/or is represented by at least one probe set or at least one target sequence listed in Table I.

In another aspect of the invention, the reagent is for detecting AML type leukemia and is suitable for determining the expression level of at least one marker gene, wherein the at least one marker gene corresponds to at least one of the genes listed in Table II and/or is represented by at least one probe set or at least one target sequence listed in Table II.

In another aspect of the invention, the reagent is for detecting CLL type leukemia and is suitable for determining the expression level of at least one marker gene, wherein the at least one marker gene corresponds to at least one of the genes listed in Table III and/or is represented by at least one probe set or at least one target sequence listed in Table III.

In another aspect of the invention, the reagent is for detecting CML type leukemia and is suitable for determining the expression level of at least one marker gene, wherein the at least one marker gene corresponds to at least one of the genes listed in Table IV and/or is represented by at least one probe set or at least one target sequence listed in Table IV.

In another aspect of the invention, the reagent is for detecting hairy cell type leukemia and is suitable for determining the expression level of at least one marker gene, wherein the at least one marker gene corresponds to at least one of the

genes listed in Table V and/or is represented by at least one probe set or at least one target sequence listed in Table V.

In another aspect of the invention, the reagent is for detecting multiple types leukemia and is suitable for determining the expression level of at least one marker gene, wherein the at least one marker gene corresponds to at least one of the genes listed in Table VI and/or is represented by at least one probe set or at least one target sequence listed in Table VI.The marker genes are represented by probe sets that match to a target DNA sequence (target sequence; cf. Table VIII) from said marker genes. The identifiers of the probe sets of the present invention refer to the identifiers of Affymetrix Probe Sets from Affymetrix, Inc., USA, and are unique.

In one aspect, the invention relates to at least one molecular signature, comprising at least one marker gene and optionally at least one contrasting gene, measured from human blood samples, to detect leukemia and to classify it into specific types of leukemia. The invention covers several areas of the method, in particular: (a) the marker gene (and optionally the used contrasting gene) is detected in blood and associated with the disease and its specific types, (b) the choice of marker genes allows to distinguish between healthy and diseased samples, including disease types and subtypes; (c) the marker genes and optionally the contrasting genes can be grouped into diagnostic signatures; (d) the use of the contrasting genes increase the power of the diagnostic test; and (e) the method can be applied in a clinical environment, in particular for the detection and classification of leukemia in blood from human patients.

Type of molecule measured in blood samples

The present invention relates to the measurement of at least one marker gene in human blood, preferably from peripheral blood mononuclear cells, to detect signatures associated with leukemia. In particular, it relates to the measurement of nucleic acids, preferably gene transcripts (mRNA), as well as the products (preferably proteins) and regulators (preferably microRNAs and other small RNAs) of these marker genes.

Molecules associated with the disease

The present invention relates to the following marker genes for the detection of leukemia:

a) ALL type leukemia. The marker genes from Table I are strongly expressed, or up-regulated, in the blood of ALL type leukemia patients and weakly expressed in the blood of healthy patients (control). These genes may, but must not, be weakly expressed in thymus and bone marrow of healthy patients. The invention relates in particular to the genes represented by the Affymetrix probe sets: 220389_at (ENSG00000149201, CCDC81, coiled-coil domain containing 81; target sequence: SEQ ID NO:1), 215117_at (ENSG00000175097, RAG2, recombination activating gene 2; target sequence: SEQ ID NO:2), 208302_at (ENSG00000158497, HMHB1, histocompatibility (minor) HB-1; target sequence: SEQ ID NO:3), 213060_s_at (ENSG00000064886, CHI3L2, chitinase 3-like 2; target sequence: SEQ ID NO:4), 240179_at (uncharacterized

LOC100505801; target sequence: SEQ ID NO:5), and 1556598_at, ENSG00000172995, ARPP21, cAMP-regulated phosphoprotein, 21kDa; target sequence: SEQ ID NO:49). The genes represented by the Affymetrix probe sets 220389_at and 240179_at are preferred as single marker genes or in combination with other marker genes from Tables I to VI. The genes represented by the probe sets 215117_at, 208302_at, 213060_s_at, and 1556598_at are preferred in combination with other marker genes indicated in Tables I to VI.

b) AML type leukemia. The marker genes from Table II are strongly expressed, or up-regulated, in the blood of AML type leukemia patients but are weakly or non-expressed in other leukemia types and in healthy blood samples (control). The invention relates in particular to the genes represented by the Affymetrix probe sets: 1553613_s_at (FOXC1 (forkhead box C1); target sequence: SEQ ID NO:6) and 240766_at (Homo sapiens cDNA FLJ43834 fis, clone TESTI4005801; target sequence: SEQ ID NO:7), 236738_at (ENSG00000180044, C3orf80, chromosome 3 open reading frame 80; target sequence: SEQ ID NO:8), 236892_s_at (ENSG00000233101, Hs.660088, HOXB-AS3, HOXB cluster antisense RNA 3; target sequence: SEQ ID NO:9). The gene represented by the Affymetrix probe set 236892_s_at is preferred as single marker genes or in combination with other marker genes from Tables I to VI. The genes represented by the probe sets 1553613_s_at, 240766_at, and 236738_at are preferred in combination with other marker genes indicated in Tables I to VI.

c) CLL type leukemia. The marker genes from Table III are strongly expressed, or up-regulated, in the blood of CLL type leukemia patients but are weakly or non-

expressed in other leukemia types and in healthy blood samples (control). The invention relates in particular to the genes represented by the Affymetrix probe sets: 1562587_at and 241483_at (ENSG00000109684, Cytokine-dependent hematopoietic cell linker; target sequence: SEQ ID NO:10 and SEQ ID NO:11), 205414_s_at (ENSG00000006740, ARHGAP44, Rho GTPase activating protein 44; target sequence: SEQ ID NO:12), 236184_at (Hs.687695, no description; target sequence: SEQ ID NO:13), 1562754_at (Hs2.436003.1, hypothetical LOC339260; target sequence: SEQ ID NO:14), 239185_at (ENSG00000154258, ABCA9, ATP-binding cassette, sub-family A, member 9; target sequence: SEQ ID NO:15), 241278_at (GenBank AI674679; target sequence: SEQ ID NO:16). The genes represented by the Affymetrix probe sets 1562587_at, 241483_at, 236184_at, 1562754_at, and 241278_at are preferred as single marker genes or in combination with other marker genes from Tables I to VI. The genes represented by the probe sets 205414_s_at and 239185_at are preferred in combination with other marker genes indicated in Tables I to VI.

d) CML type leukemia. The marker genes of Table IV are strongly expressed, or up-regulated, in the blood of CML type leukemia patients but have significantly lower expression in other leukemia types and in healthy blood samples (control/reference). The invention relates in particular to the genes represented by the Affymetrix probe sets 236031_x_at (ENSG00000164946, FRAS1 related extracellular matrix 1 (FREM1); target sequence: SEQ ID NO:17), 204624_at (ENSG00000123191, ATP7B, ATPase, Cu++ transporting, beta polypeptide; target sequence: SEQ ID NO:18), and 205984_at (ENSG00000145708, CRHBP,

corticotropin releasing hormone binding protein (CRHBP); target sequence: SEQ ID NO:29). The genes represented by the Affymetrix probe sets 236031_x_at and 204624_at are preferred as single marker genes or in combination with other marker genes from Tables I to VI. The gene represented by the probe set 205984_at is preferred in combination with other marker genes indicated in Tables I to VI.

e) Hairy cell type leukemia (HCL). The marker genes of Table V are strongly expressed, or up-regulated, in the blood of CML type leukemia patients but have significantly lower expression in other leukemia types and in healthy blood samples (control/reference). The invention relates in particular to the genes represented by the Affymetrix probe sets 238009_at (Hs.434948, no description; target sequence: SEQ ID NO:19) and 207336_at (ENSG00000134532, SOX5, SRY (sex determining region Y)-box 5; target sequence: SEQ ID NO:20). The genes represented by the Affymetrix probe sets 238009_at and 207336_at are preferred as single marker genes or in combination with other marker genes from Tables I to VI.

f) Multiple types of leukemia. The marker genes of Table VI are strongly expressed, or up-regulated, in the blood in response to two or more of the types ALL, CLL, AML, CML, and/or Hairy Cell type leukemia patients and weakly or non-expressed in the blood of healthy patients (control). The invention relates in particular to the genes represented by the Affymetrix probe sets 242520_s_at (ENSG00000198520, C1orf228, chromosome 1 open reading frame 228; target sequence: SEQ ID NO:21), 1566482_at (Hs.684006, no description; target

sequence: SEQ ID NO:22), 214575_s_at (ENSG00000172232, AZU1, azurocidin 1; target sequence: SEQ ID NO:23), 230551_at (ENSG00000171435, KSR2, kinase suppressor of ras 2; target sequence: SEQ ID NO:24), 213714_at (ENSG00000165995, CACNB2, calcium channel, voltage-dependent, beta 2 subunit; target sequence: SEQ ID NO:25), 1558871_at (clone IMAGE:4105785; target sequence: SEQ ID NO:26), 219790_s_at (ENSG00000113389, NPR3, natriuretic peptide receptor C/guanylate cyclase C; target sequence: SEQ ID NO:27), 238206_at (ENSG00000171509, RXFP1, relaxin/insulin-like family peptide receptor 1; target sequence: SEQ ID NO:30), 1552715_a_at (ENSG00000171509, RXFP1, relaxin/insulin-like family peptide receptor 1; target sequence: SEQ ID NO:31), 236632_at (ENSG00000248890, hypothetical LOC100505470; target sequence: SEQ ID NO:32), 1559315_s_at (ENSG00000246985, hypothetical protein LOC144481; target sequence: SEQ ID NO:33). The marker gene relating to probe set 230551_at is predominant in CLL and shows signals also in ALL. The marker gene relating to probe set 1558871_at is predominant in AML. The marker gene relating to probe set 1552715_a_at is particularly high at blast crisis in bone marrow. The marker genes relating to probe sets 236632_at and 1559315_s_at are predominant in ALL. The genes represented by the Affymetrix probe sets 242520_s_at, 1566482_at, 213714_at, 1558871_at, 236632_at are preferred as single marker genes or in combination with other marker genes from Tables I to VI. The genes represented by the probe sets 214575_s_at, 230551_at, 219790_s_at, 203948_s_at, 203949_at, 238206_at, 1552715_a_at, and 1559315_s_at are preferred in combination with other marker genes indicated in Tables I to VI.

Table VI also indicates the types of leukemia, in which the expression is highest, but it does not exclude expression in further types of leukemia.

Grouping of molecules into diagnostic signatures

The present invention relates to the use of these marker genes (Tables I to VI) and/or contrasting gene (Table VII), alone or in combination, to present a diagnostic prediction with increased specificity and/or sensitivity. It is statistically established that the simultaneous use of multiple markers with independent classification power increases the performance of the diagnostic test. In some cases, a single marker gene may have the highest discriminating power between healthy and diseased samples, whereas in other cases, the combination of a plurality of marker genes offers the highest discriminating power. The present invention makes use of this property to select suitable combinations of marker genes for the detection of leukemia and its specific types.

Contrasting molecules

The present invention further relates to the use of contrasting genes (also called reference genes) listed in Table VII that have opposite properties, such as non-expressed, or down-regulated, in diseased samples but expressed in healthy samples (control/reference), to increase the discriminating power of the diagnostic test. In particular, it relates to the use of such genes either for correcting the values of the marker genes, or for the generation of ratios between marker genes and contrasting genes. The contrasting genes can also directly be used as refer-

ence in step b of the method. Table VII further indicates the preferred leukemia type for which the contrasting gene is especially well suited.

Application in a clinical environment

The present invention relates to the use of at least one of the aforementioned marker genes and optionally at least one of the contrasting molecules to create a diagnostic test to detect diseases, preferably cell proliferating diseases, preferably leukemia, more preferably specific types of leukemia, in human blood samples.

BRIEF EXPLANATION OF THE FIGURES

The invention is described in greater detail below with reference to embodiments that are illustrated in the figures. The figures show:

Fig. 1 Expression evidence for the selected marker genes (Tables I-VI), based on results obtained from Genevestigator. The plots represent the expression level of individual markers across 5941 Affymetrix Human 133 Plus 2.0 arrays. Under (a) Acute Lymphocytic Leukemia (ALL); (b) Chronic Lymphocytic Leukemia (CLL); (c) Acute Myelocytic Leukemia (AML); (d) Chronic Myelocytic Leukemia (CML); (e) Hairy Cell Leukemia (HCL); and (f) Multiple leukemia types.

Fig. 2 Heat map representing the average log-expression values of all marker genes across different blood tissues, leukemia types and subtypes, and different types of lymphatic tissue neoplasms. The number of arrays taken to process each average is indicated on the right of the heat map.

DETAILED DESCRIPTION OF THE INVENTION

Materials and Methods

Identification of marker genes:

Gene expression biomarkers were identified using Genevestigator (Hruz et al., 2008). The search for molecular markers was performed by comparing the average expression level in selected "target" categories with the average expression level in selected "base" categories. "Target" categories refer to individual or multiple diseases in which marker genes are strongly expressed. "Base" categories, which include the target categories, refer to the complete set of categories the analysis is done with. Marker genes that are highly expressed in the "target" categories are expected to be weakly or non-expressed in the other "base" categories. As "base" categories, we chose either:

- Selection A: leukemic and non-leukemic blood and cells extracted from blood;

- Selection B: selection A + leukemic bone marrow samples;

- Selection C: selection B + blood samples having a variety of lymphatic system neoplasms; and

- Selection D: selection C + all cancer samples and all non-cancer tissue samples available from the Genevestigator database in May 2012.

In an initial search, we looked for genes that are highly expressed in leukemia samples but are not expressed (or minimally expressed) in other cancer types and healthy tissue types. In this case, the "base" categories comprise all categories, i.e. corresponds to selection D above.

- Target: separately ALL (B-ALL; c-ALL; T-ALL) or CLL (B-lineage) or AML or CML or Hairy Cell Leukemia

- Base: all categories of selection D above

In a second search, we identified genes that are expressed in specific types or subtypes of leukemia but are not expressed in healthy blood samples or healthy cells extracted from blood (corresponds to selection A above) and in other hematologic diseases (corresponds to selection C above). We verified expression of these genes in leukemic and healthy bone marrow samples (corresponds to selection B above) to confirm their response to the leukemia type chosen as target.

- Target: separately ALL (B-ALL; c-ALL; T-ALL) or CLL (B-lineage) or AML or CML or Hairy Cell Leukemia

- Base: blood leukemia (all available subtypes), blood, platelet, peripheral blood leukocyte, peripheral blood mononuclear cells (all), monocyte derived dendritic cell, leukocyte, B-lymphocyte, monocyte, natural killer T-cell, T-lymphocytes, macrophages, CD4 - and CD8 T-lymphocytes

In a third step, we filtered the candidate marker genes (between 10 and 100, depending on leukemia subtype) against the Human Perturbations datasets available in Genevestigator to exclude genes that are up-regulated in response to common diseases, infections, drugs or other conditions. This dataset contained at the time of analysis more than 1,000 different perturbations, including response to drugs, diseases, hormones, stress, mutations, and a group of miscellaneous conditions. This step is to maximize disease specificity and minimize the probability of false diagnosis caused by other conditions affecting the tested patient. This filtering resulted in a small number of genes that are highly specifically expressed in a chosen type or subtype of leukemia, weakly or non-expressed in any other disease, and weakly or non-responsive to common conditions such as viral or bacterial infection, to commonly used drugs or other common conditions.

The figures 1 and 2 were generated from figures or data exported from the Genevestigator analysis tool (https://www.genevestigator.com; Nebion AG, Switzerland).

Mappings of probe set identifiers to gene model identifiers were obtained using publicly available Bioconductor packages. Transcript sequences were obtained from the array manufacturer (Affymetrix, NetAffx™ Analysis Center).

Blood collection and Peripheral Blood Mononuclear Cell separation:

The collection of blood, the separation of peripheral blood mononuclear cells (PBMC) and their storage were performed according to the COLOX Clinical Study

protocol (www.diagnoplex.com). CPT VACUTAINER TUBES (Becton Dickinson) were used to collect blood in collection centers. Within 6 hours, PBMC were extracted by centrifugation and washed twice in 1x PBS. PBMC were stored in RNA later ([kit provider]) at -20 °C until shipment to the test laboratory, where they were stored at -80°C.

RNA extraction:

An equivalent of 700'000 to 2'000'000 PBMC were used for total RNA extraction using RNeasy Mini Kit and Qiacube (Qiagen) for automatization. 600 ng to 3 µg of total RNA were obtained in a final elution volume of 30 µL. A Bioanalyzer (Agilent) and a NanoDrop (Thermo Scientific) were performed for quality and quantity check. RNA was stored at -80 °C.

Reverse Transcription:

For reverse transcription, we used 200 ng of total RNA and SuperScript VILO cDNA Synthesis Kit (Invitrogen). The final volume of cDNA was 20 µL. The positive control was 200 ng of Human Universal Reference Total RNA (Clontech). The negative control was H_2O (Roche). The program in the Thermocycler was set according to the SuperScript VILO Technical Data Sheet. cDNA was diluted 1:2 in H_2O (Roche) and stored for 24 hours at -20 °C.

Real-Time PCR:

Primers (forward and reverse) and TaqMan probes (FAM/TAMRA) for the tested marker genes were designed, whereas the TaqMan assays related to the reference genes were purchased from Life Technologies. In each well, 2 µM of each primer and 2.5 µM of probe was added. TaqMan Fast Advanced Master Mix (Life Technologies) was used at 1x concentration, and the final reaction volume per well was 10 µL. Plate Loading was automated with the CAS1200 (Corbett Robotics) adapted to 384-wells plates and real-time PCR was performed using the VIIA7 instrument platform (Life Technologies). The PCR program with 45 cycles was set according to the Technical Data Sheet of the Master Mix. Ct values were obtained after threshold adjustment to 0.05 Delta Rn.

Quality Control:

The quality control of the experiment was performed by checking the positive controls for each run of RT-qPCR. The 71 subjects were divided in 54 runs of RT-qPCR. Therefore 4 independent runs of RT were performed with 1 positive and 1 negative control per run.

Results

Marker genes and contrasting genes identified from Genevestigator are listed in Tables I - VII. The corresponding target sequences relating to the genes are listed in Table VIII.

Fig. 1 shows the expression evidence for the selected marker genes, based on results obtained from Genevestigator (Nebion AG, Switzerland). Under Fig.1(a)

for Acute Lymphocytic Leukemia (ALL); Fig.1(b) for Chronic Lymphocytic Leukemia (CLL); Fig.1(c) for Acute Myelocytic Leukemia (AML); Fig.1(d) for Chronic Myelocytic Leukemia (CML); Fig.1(e) for Hairy Cell Leukemia (HCL); and Fig.1(f) for Multiple leukemia subtypes. The numbers of the probe sets refer to the identifiers of Affymetrix Probe Sets from Affymetrix, Inc., USA, and are unique. The plots represent the expression level of individual markers across 5941 Affymetrix Human 133 Plus 2.0 arrays, with the following order of the samples: 1-1344: Non-malignant (blood, 1344 samples); 1345-1590: ALL (blood, 246 samples); 1591-1688: CLL (blood, 98 samples); 1689-2338: AML (blood, 640 samples); 2339-2343: JMML (blood, 5 samples); 2344-2347: Hairy Cell Leukemia (blood, 4 samples); 2348-2947: Lymphomas (blood, 600 samples); 2948-4027: ALL (bone marrow, 1079 samples); 4028-4449: CLL (bone marrow, 422 samples); 4450-5753: AML (bone marrow, 1304 samples); 5754-5788: APL (bone marrow, 35 samples); 5789-5843: CML (bone marrow, 55 samples); 5844-5876: JMML (bone marrow, 33 samples); and 5877-5941: CML (bone marrow, 65 samples).

Fig. 2 shows a heat map representing the average log-expression values of all marker genes across different blood tissues, leukemia types and subtypes, and different types of lymphatic tissue neoplasms. The number of arrays taken to process each average is indicated on the right of the heat map. The circles and the roman numbers indicate the marker genes specific for types and subtypes of leukemia (Table I to V).

Output from quality control:

As a proof of principle, one novel marker gene (C1orf228, see Table VI), two published marker genes (IGLL1 and VPREB1) and three reference genes (NACA, RPLP0, TPT1) were selected for quality control experiments and validation using RT-qPCR.

For each target gene, Ct and Standard Deviation (StDev) for positive control (POS) were calculated and results are shown in Table 1. The very low variations observed between the 4 independent runs show high analytical reproducibility using the defined protocol.

Table 1: Ct and StDev values for POS and four independent runs. Over-expressed biomarkers (first three columns) and reference genes (last three columns).

	C1orf228	IGLL1	VPREB1	NACA	RPLP0	TPT1
Run01	27.2	26.7	34.5	19.2	18.7	16.7
Run02	27.2	27.0	35.8	19.8	18.7	16.8
Run03	27.2	26.8	34.6	19.7	18.8	16.7
Run04	27.5	27.3	34.9	19.6	18.9	16.9
Average POS	27.3	26.9	34.9	19.6	18.8	16.8
StDev POS	0.166	0.252	0.606	0.257	0.116	0.098

Gene expression results:

Table 2 P-values of t-tests between the expression in the clinical groups comprising healthy samples (CON) and leukemic samples (ALL, CLL, AML, CML). The test

was performed on the gene expression results from one tested marker gene selected from Table VI (C1orf228), two known biomarker genes (IGLL1 and VPREB1) and three reference genes (NACA, RPLP0, TPT1). The normalization of the biomarker expression was obtained by Delta Ct methodology using the average of these 3 reference genes.

Comparison	C1orf228	IGLL1	VPREB1
Myeloid leukemia:			
CON vs AML+CML	2.E-11	2.E-05	0.001
CON vs AML	3.E-11	1.E-04	0.001
CON vs CML	5.E-04	1.E-07	0.355
Lymphoblastic leukemia:			
CON vs ALL + CLL	0.085	3.E-06	1.E-08
CON vs ALL	0.005	1.E-06	2.E-10
CON vs CLL	0.210	0.002	3.E-05
Leukemia:			
CON vs AML+CML+ALL+CLL	4.E-08	5.E-06	2.E-04

References:

(1) Hruz T, Laule O, Szabo G, Wessendrop F, Bleuler S, Oertle L, Widmayer P, Gruissem W and Zimmermann P (2008) Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. Adv Bioinformatics vol. 2008, Article ID 420747, 5 pages. DOI: 10.1155/2008/420747

Tables I to VII: Marker genes and contrasting genes are represented by probe sets that match to a target DNA sequence (target sequence; cf. Table VIII) from said marker or contrasting genes. The identifiers of the probe sets of the present invention refer to the identifiers of Affymetrix Probe Sets from Affymetrix, Inc., USA, and are unique. The sequences of the target sequences (SEQ ID NO: 1 to 49) of Tables I to VII are listed in Table VIII.

Table I: ALL type leukemia

Probe set ID	Gene Identifier	Gene	Gene Description	Target Sequence
220389_at	ENSG00000149201	CCDC81	coiled-coil domain containing 81	SEQ ID NO:1
215117_at	ENSG00000175097	RAG2	recombination activating gene 2	SEQ ID NO:2
208302_at	ENSG00000158497	HMHB1	histocompatibility (minor) HB-1	SEQ ID NO:3
213060_s_at	ENSG00000064886	CHI3L2	chitinase 3-like 2	SEQ ID NO:4
240179_at	LOC100505801	--	uncharacterized LOC100505801	SEQ ID NO:5
1556598_at	ENSG00000172995	ARPP21	cAMP-regulated phosphoprotein, 21kDa	SEQ ID NO:49

Table II: AML type leukemia

Probe set ID	Gene Identifier	Gene	Gene Description	Target Sequence
1553613_s_at	ENSG00000054598	FOXC1	forkhead box C1	SEQ ID NO:6
240766_at	Hs.449281, Entrez ID: 80314	--	Homo sapiens cDNA FLJ43834 fis, clone TESTI4005801	SEQ ID NO:7

Probe set ID	Gene Identifier	Gene	Gene Description	Target Sequence
236738_at	ENSG00000180044	C3orf80	chromosome 3 open reading frame 80	SEQ ID NO:8
236892_s_at	ENSG00000233101, Hs.660088	HOXB-AS3	HOXB cluster antisense RNA 3	SEQ ID NO:9

Table III: CLL type leukemia

Probe set ID	Gene Identifier	Gene	Gene Description	Target Sequence
1562587_at	ENSG00000109684	CLNK	Cytokine-dependent hematopoietic cell linker	SEQ ID NO:10
241483_at	ENSG00000109684	CLNK	Cytokine-dependent hematopoietic cell linker	SEQ ID NO:11
205414_s_at	ENSG00000006740	ARHGAP44	Rho GTPase activating protein 44	SEQ ID NO:12
236184_at	GenBank AI021923, Hs.687695	--	no description	SEQ ID NO:13
1562754_at	Hs2.436003.1, Entrez ID: 339260	--	hypothetical LOC339260	SEQ ID NO:14
239185_at	ENSG00000154258	ABCA9	ATP-binding cassette, subfamily A, member 9	SEQ ID NO:15
241278_at	GenBank AI674679	--	no description	SEQ ID NO:16

Table IV: CML type leukemia

Probe set No.	Gene Identifier	Gene	Gene Description	Target Sequence
236031_x_at	ENSG00000164946	FREM1	FRAS1 related extracellular matrix 1	SEQ ID NO:17
204624_at	ENSG00000123191	ATP7B	ATPase, Cu++ transporting, beta polypeptide	SEQ ID NO:18
205984_at	ENSG00000145708	CRHBP	corticotropin releasing hormone binding protein (CRHBP)	SEQ ID NO: 29

Table V: Hairy cell type leukemia (HCL)

Probe set No.	Gene Identifier	Gene	Gene Description	Target Sequence
238009_at	Hs.434948	--	no description	SEQ ID NO:19
207336_at	ENSG00000134532	SOX5	SRY (sex determining region Y)-box 5	SEQ ID NO:20

Table VI: Multiple types of leukemia. The leukemia types of column 5 indicate the types of leukemia in which the expression is highest, but it does not exclude expression in further types of leukemia.

Probe set No.	Gene Identifier	Gene	Gene Description	Leukemia type	Target Sequence
242520_s_at	ENSG00000198520	C1orf228	chromosome 1 open reading frame 228	ALL/AML	SEQ ID NO:21
1566482_at	Hs.684006	--	no description	AML/CML	SEQ ID NO:22

Probe set No.	Gene Identifier	Gene	Gene Description	Leukemia type	Target Sequence
214575_s_at	ENSG00000172232	AZU1	azurocidin 1	ALL/AML/CML	SEQ ID NO:23
230551_at	ENSG00000171435	KSR2	kinase suppressor of ras 2	CLL	SEQ ID NO:24
213714_at	ENSG00000165995	CACNB2	calcium channel, voltage-dependent, beta 2 subunit	CLL	SEQ ID NO:25
1558871_at	clone IMAGE:4105785	--	--	AML/CML	SEQ ID NO:26
219790_s_at	ENSG00000113389	NPR3	natriuretic peptide receptor C/guanylate cyclase C	AML/CML	SEQ ID NO:27
238206_at	ENSG00000171509	RXFP1	relaxin/insulin-like family peptide receptor 1	AML/CML	SEQ ID NO:30
1552715_a_at	ENSG00000171509	RXFP1	relaxin/insulin-like family peptide receptor 1	AML/CML	SEQ ID NO:31
236632_at	ENSG00000248890	--	hypothetical LOC100505470	ALL/AML	SEQ ID NO:32
1559315_s_at	ENSG00000246985	--	hypothetical protein LOC144481	ALL/AML	SEQ ID NO:33

Table VII: Contrasting markers or reference genes. The leukemia types of column 5 indicate the types of leukemia for which the contrasting gene is especially well suited.

Probe set No.	Gene Identifier	Gene	Gene Description	Leuk. Type	Target Seq.
204897_at	ENSG00000171522	PTGER4	prostaglandin E receptor 4	hairy cell leukemia	SEQ ID NO:34
215096_s_at	ENSG00000139684	--	esterase D mRNA, complete	hairy cell leukemia	SEQ ID NO:35

Probe set No.	Gene Identifier	Gene	Gene Description	Leuk. Type	Target Seq.
			cds		
212074_at	ENSG00000164828	SUN1	Sad1 and UNC84 domain containing 1	CML	SEQ ID NO:36
212124_at	ENSG00000108175	ZMIZ1	zinc finger, MIZ-type containing 1	CLL	SEQ ID NO:37
225673_at	ENSG00000179820	MYADM	myeloid-associated differentiation marker	CLL	SEQ ID NO:38
229560_at	ENSG00000101916	TLR8	toll-like receptor 8	ALL	SEQ ID NO:39
220005_at	ENSG00000181631	P2RY13	purinergic receptor P2Y, G-protein coupled, 13	ALL	SEQ ID NO:40
221766_s_at	ENSG00000112773	FAM46A	family with sequence similarity 46, member A	ALL	SEQ ID NO:41
204912_at	ENSG00000110324	IL10RA	interleukin 10 receptor, alpha	ALL	SEQ ID NO:42
227167_s_at	Hs.643605	--	CLONE=IMAGE:504304	t-ALL	SEQ ID NO:43
202270_at	NM_002053	GBP1	guanylate binding protein 1, interferon-inducible, 67kDa	all leukemias	SEQ ID NO:44
226272_at	ENSG00000117602	RCAN3	RCAN family member 3	all leukemias	SEQ ID NO:45
201877_s_at	ENSG00000078304	PPP2R5C	protein phosphatase 2, regulatory subunit B', gamma	AML	SEQ ID NO:46
208091_s_at	ENSG00000154978	VOPP1	vesicular, overexpressed in cancer, prosurvival protein 1	AML	SEQ ID NO:47
212149_at	ENSG00000132294	EFR3A	EFR3 homolog A (S. cerevisiae)	CML	SEQ ID NO:48

Probe set No.	Gene Identifier	Gene	Gene Description	Leuk. Type	Target Seq.
201877_s_at	ENSG00000078304	PPP2R5C	protein phosphatase 2, regulatory subunit B', gamma	AML	SEQ ID NO:28

Table VIII: target sequences corresponding to the genes of Tables I to VII:

SEQ ID NO	Target Sequence
1	gaacgggcttcagacaagctgtttctcctagaccagtgtgagaagtatcggcgctgcaag caatgccagaggcgcacctccaatgtgggcgagagcaacctgtggcccctgaacaagttc ctgcctggctcccggttgcttgtgtaaaactcaaagtttggctcttcgtttcccggggaa agtttttatcttttacatgtttgggggtgattgcgaaactgcgtatttttacctcagaga aaaaaatcattgtttaggtttggtgctttcattagattgcttgttaagcccttattgaat tcactcctgctttcctcccacccccaattatttcctatactagtttctgatggcagtgaa ggtgtctgaatggtcctgagggctagaacctgctgcacaggggctgggaatgggatccag cttcattatggctgctggggtgctgctgacccagccctgctgctgctccatgactttgtg tacaaaa
2	gtattactgcaatgagcatgtggagatagcaagagctctacacactccccaaagagtcct acccttaaaaaagcctccaatgaaatccctccgtaaaaaaggttctggaaaaatcttgac tcctgccaagaaatccttttcttagaaggttgtttgattagttttgcaaaagcctttcaga ttcaggtgtatggaatttttgaatctatttttaaaatcataacattgattttaaaaatac atttttgtttatttaaaatgcctatgttttcttttagttacatgaattaagggccagaaa aaagtgtttataatgcaatgataaataaagtcattctanaccctananattttgaaaata ttttacccaaatactcaatttactaatttattcttcactgaggatttctgatctgatttt ttattcaacaaaccttaaacacccagaagcagtaataatcatcgaggtat
3	tgccccgcatctcagaagccgggcaggccctgagccttctgacctcacatcctctgccac accacagtggagaaaccagaactggaggagcagccagaatgcagagaagaaaaaagaggt tctctgcatgtttggaagtcggaattggttgaagttgaagatgatgtgtatctgaggcac agctcttccctgacttataggctttgacactgctgttgaggtttgactcgaagcccagag ttttggtgtggatgagcagggacaaattgctgagcatgaagaagagtaaaattaagcaag tggaacatatgcccttt
4	ctctgcagaaaccaccgtgggggcccctgcctctggccctggagctgctggacccatcac agagtcttcaggcttcctggcctattatgagatctgccagttcctgaaaggagccaagat cacgcggctccaggatcagcaggttccctacgcagtcaaggggaaccagtgggtgggcta tgatgatgtgaagagtatggagaccaaggttcagttcttaaagaatttaaacctgggagg agccatgatctggtctattgacatggatgacttcactggcaaatcctgcaaccagggccc ttaccctcttgtccaagcagtcaagagaagccttggctccctgtgaaggattaacttaca gagaagcaggcaagatgaccttgctgcctggggcctgctctctcccaggaattctcatgt gggattcccttgccaggnnggcctttggatctctcttccaagcctttcctgacttcctc ttagatcatagattggacctggttttgttttcctgcagctgttgacttgttgccctgaag ta

SEQ ID NO	Target Sequence
5	cttggcctcacccattagaaggtaagctccataaaaggtgactatgctgtatcaaggaat cttgtatagtgtctggtatttaggtaacagacaataaatattaaagaatgaatagcatat ctctatttttttaacttgagttaaaccaaactccataggtgagacttctaagagcctttaa aatgctttttgaattttcataatggaggcatagtaagaattatagcctaaacatagaatc ctggtagcatcttagatggtggttaatgctttgtgtaacacactctggcacttttaaaat tgtaaatcagtttaatttttatcacaaatgtccaacacttttctcttatgaa
6	aacctgcaagccatgagcctgtacgcggccggcgagcgcgggggccacttgcagggcgcg cccgggggcgcgggcggctcggccgtggacgacccccctgcccgactactctctgcctccg gtcaccagcagcagctcgtcgtccctgagtcacggcggcggcggcggcggcgggggga ggccaggaggccggccaccaccctgcggcccaccaaggccgcctcacctcgtggtacctg aaccaggcgggcggagacctgggccacttggcgagcgcggcggcggcggcggcggccgca ggctacccgggccagcagcagaacttccactcggtgcgggagatgttcgagtcacagagg atcggcttgaacaactctccagtgaacgggaatagtagctgtcaaatggccttcccttcc agccagtctctgtaccgcacgtccggagctttcgtctacgactgtagcaagttttgacac accctcaaagccgaactaaatcgaaccccaaagcaggaaaagctaaaggaacccatcaag gcaaaatcgaaa
7	agtgcatatcttggtaaggctctgtgaggaaatgacctaatctggagagaagctgctggg acactcntctgtgtgtggcctgtggctcacaggggcaggtattctcagaacgcacatggg ataatcaccatcctggtgtgtccaggctctgggaacaggctcttcctcgtggctttcaca tgagtttcctggtgggctatgtagtcacttttctgcatccttcttttaaaggctcatgaa tgtcaaagtaacacagaccctgagatgaggcaggaaagttgtatcggaatgttttcagac tatcaaccagaccaaacgttctggaatccataa
8	gtaatgtacatatgcatattgtctatgtactatatacacattgtttataatattgttatt acagtttgtcattcacctgaaaggagaaggaaaagtacgaaaggtttctctgcttgggaa aaggtgaatgttgttatatcaagaacgattacacacatggntctcatacatgnttttaga acattgttcttctgattgaagaagtctgatgctcctgaaaaatcttaaaatatctgactt gtattgaagaaaattatttaattaaatttttaaaggctggttgaaaaagtctgacagttc tgagaatttttttaaatgtctgaattgtaataaaaaaatggtttactttaaacttctaaaa actaatgaccttgtgactaa
9	ggtagacttcgcttgaaaagatcgagaatggaggaggggatcggcactcacttgcagaca tcaacagtttccagaaaattcggttcaattccttcaccacggacccaccaaccaaggagc tggcaaaacccactaaccaaccaggagaacacagacccggtttgtctttttgacaaccgag tggaactgacggg
10	aacattctagtaactggccactggaaaataaacaaaaataaaaactagggttttaaaagt atcttctaaaaaacaacaacaaaaaatactataaacatagccattatgctcatgatacag gcgagcagcaaagggcaccagaagctgttgcttaaatgtttgcagccagtgcaagacaag tctatgggaaattcccaaatctgtgctctttacaggacactgcgctgcctttatgtcagt tgttgggccttacctatctacaatgtgtg
11	gtactgcatgtgtattcattactaagcactattgtagatgttgaaaatagagcagtgaat aagagagatagatgacaaataaatatacaaacaaatnagataannnaanncagtttccca ttactggattggatttctattttaacacttatttctgagtacttgtggggtgatgccaac attaaatgttttgaggaagctcctttgagatggctttgaaaagtcctaagtgagaaaant ttaaaantttcccaaacttatttatgtggggtgtgcatgtgtgtgcctgtgcctgtgtg cgtgtgtgtgtgtgtgtttttaagtaaccgncatcatggctgtaatcagttaagtctcct gttctaattactacggaggaatcaaatgaaccccattgtcgaatgtagccaattatgtat gtgtgtacacaaatacaaatttttatgatagcctatattgtatattagcattatatttgt atatctttttattctgagcaatatttgttctagctaagca

SEQ ID NO	Target Sequence
12	attacgttacccaaagcatgtggaggaaaagtgaaaccagccacggagacgctggcccac ggctcggcctgcggtgtggcctgctttgctcaccagcgtcagccgctcatttccttctca tgaagtcccatctggtcatggggacgagggccgggagggcaccgggtagccttttcacac ttggggattaggggagtgagaaaagatttgggccatgcatgcaaagtcaaagtttaaaat tttatccttttcaaatagatgatataatatacctatacatgatataatatttgtatatat gaaatctctctatatttgtttaatttgagccattcaatctaaaccaatgtacaggtgtac aatgaaaaatttaaatgcttagttatttttcccaacacagtgtaaagtcaccctcctctg agagtgggatgtgcagagttttgatgttgcagctttgctcacttcctggcaagggcaggt catgcctcaatttgtaatgggagtctggggtaagggtg
13	aaatcagtcttaacgctcagagataaataaccacattcatgacaatgagattatgtaatt gtttttttattcacaaatattttctggcacacacattaattcacagaaagaaagaacagta gtcacgtgggcacaatatttaacccttttcagatactgaggtttatcaccgcttgatttct tagttggaaatgttgcagtttggtttcaggcaggatttccaggctatcacagcctgctcc gatggcatgattattaggagctcatctgcatcactggtgaatttctttggtcacttctgg acagtttttaaacataagtccatgcccttctttt
14	aactttcaggtctcggcccaaacatcgggggtctcctctgaccacacagtagtactgggc cctcctttgtactgctagtgttactttcttttgaaacttttaattacagaacttaccaca catatacatagagaggactgcacaatgatgccctcctcatccattaccagcttcacaagt catttaaattctgccttgtactttcccaaccagttgttttggcttgtgtattttaaagca aagccaagacctcatatcattttacctgcagatactgggccataaagatctgtaaaacat aacccaaatatcctatcactgccagcaataatttattaacatcatctaatattcaggcca tatccaatttttccctgattgtcttgaaaactatccattgcttgctgttctttttcataat ctgtctttttccttgatagaac
15	taccatttaagtttgcatctgaatgaaaggtgtgatccagagagtataacatcactggtt aagcagcacatctctgatgccaaattgacagcncaaagtgaagaanaacttgtatatatt ttgcctttggaaaggacaaacaaatttccagaacttnacagggatcttgatagatgttct aaccaaggcattgaggattatggtgtttccataacaactttgaatgaggtgtttctgaaa ttagaaggaaaatcaactattgatgaatcagatattggaatttggggacaattacaaact gatggggcaaaagatataggaagccttgttgagctggaacaagttttgtcttccttccac gaaacaaggaaaacaatcagtggcgtggcgctctgg
16	gtaataattcccagtcacacaaacccatacattctctggattttgccatggactcttgtt ctccaatgattccagactattagaacacttatgacattttggttaataataatccttcct ataggccaaacttgggtaatgatttaaggaggtattaaaatctgggaatgccagacagtg tgctgaaacagagactcgagtcttgaacaagaaaatatcatttgtttaaccttgaataag tcttcttactagtcttgtgcttattttcctcccttataaagtaattgagtcagcaaatca tgggggtataagaattctggatctcgcaccactctccatcaaaagtatttaacacatacc ccacatatat
17	taccacctcaggaaaaggcccactgtgtccaggggcgtcagtccagctcaatagtgtcta gcttttcagaatcttctgattgtttggtctttaacaaaaacatagcatctatttgctaac aataaacttaagcaacatatgataatcatttcatctccttccaagaagtggtgaaaatga accaggtgacattaatttaacttctgtttttgtttatttttaatacaaccaaaaagatat gcacatattttaatttacattatggtactgtataatttaaaccataaactggtgctgatt aatttctgcaacaaagcagcatattaaatgtactacatctagtgttgta
18	gtgaaggcttcactcttattgtccttctaccttgaatagaagttttcctgataagaataa acgaggaaaaggtccttgcctcctggaagaacaaatctaccaggtgatctattcattgtt tcaactcagaatgcacttgattcaggaggtcatctgaccttcaccttggatggttagttt cacttttacatatagtttttgcagggttttatttataaaatccaagcgcgctgttgat

SEQ ID NO	Target Sequence
	tgtgttttccttgttttcagcccccgactccagcccgcagcacatttccgctgtccgtc agtaattgtgtcctctctttatgcttgcttggggaatgttgttttctgactaggctgatc
19	gtatcctataactatcaacttcccaggttgaagacgatgtgttgagtttcctactgattg attgattcctgtcctccccagtgtttccgtcactggttcactaaaacagtatttatatag ctccactggctctaaagctcttagtccttctaatattttggattttacaagtaaaaatgg aaaaaaaatagaaaagagncaatcaaatgcctggagcttaaaacaaagtatgtgcaacct accatctcacttgaaatttaataaaataataagtaattatgtaaatataacatagagtta tagatttatattttgttcataacacatagtgtaatataagttgtatattttcatgttttt ggttttatgttatcattcatgccacaataaaaataaaacaggagtttatgtgctcttaaa aaaaagatgtgggttgccaccaacctgtttttcgtttttgtttttttgtttatttttatttt atttttttgcattctccttttttcagtattactgccatg
20	agccatcgccatggctgggatgccctcccctcacctgccctcggagcactcaagcgtgtc tagcagcccagagcctgggatgcctgttatccagagcacttacggtgtgaaaggagagga gccacatatcaaagaagagatacaggccgaggacatcaatggagaaatttatgatgagta cgacgaggaagaggatgatccagatgtagattatgggagtgacagtgaaaaccatattgc aggacaagccaactgataagggtcaaaagattgttgtgaccttaggacttaaagaagccc taactggttcatccttaccagtggccaagcacattaactttctcatacactgactgttac tttaactgttagtcttaaatagttgggacatcagctgactaatagacctcagcctcaaaa ggcttggaaagaaaaaacaaatacaacaagcaaacaacaatatcaacaacaagagattga aataagctatgggtaaaataatgccagtaattcagctgctacatccaagcactgaagtct tacccgtcaacttt
21	agccatcgagttgatcaaagacctggtcggttacgatgtgcgccaggcgctgctcaaggg cctcgtggcgctgctgataccgtcggtcaaggagatctccaaactgcaggcc
22	gtctatggacacttgatttcttgctttatggatactctcaggaagttctggtctcaactt tgagtatctcattgtcctcttactttagaatgtggggtatgggggaaaacttctatttga aaatttgagccatagcggaaatttggaatgacagggaaacaccttctcagaatcctctt acaataaaatttatatatgttatgccagacaatgtactttatactaagactggacttttta aatagtctcatggattagaagtcagatacaataaataaatacttttaaaaattatattt ttaggtatatgagcattgttgttgggactacagtgctattttaaagtgaaatttagttga tttgcatgtgagtatacacctacaggtctctgttatagccgtgacatgtcttctgttact ataaattacataacaaatccttatggtttcttgggtcacttattttatttaaagacaaaa acatctttcagtacatatgtaagacttaaggcattattgaaaataagcctgtgctacatg gta
23	atgagcgagaatggctacgacccccagcagaacctgaacgacctgatgctgcttcagctg gaccgtgaggccaacctcaccagcagcgtgacgatactgccactgcctctgcagaacgcc acggtggaagccggcaccagatgccaggtggccggctgggggagccagcgcagtgggggg cgtctctcccgttttcccaggtttgtcaacgtgactgtgacccccgaggaccagtgtcgc cccaacaacgtgtgcaccggtgtgctcacccgccgcggtggcatctgcaatggggacggg ggcacccccctcgtctgcgagggcctggcccacggcgtggcctccttttccctggggccc tgtggccgaggccctgacttcttcacccgagtggcgctcttccgagactggatcgatggt gttctcaacaacccgg
24	acagaccaacgcggaccgatggtccatattcttttctccatgtttgtttgtgtcgactgt tgcagagtgagttgatgtaagaaggtggtttcatgttttcttttgggggccacttcttt cttttcctttccgtttctatctctgtctgtctccgtctctgtctctctctttctcacttg ttgccaatgtcttgttgctgagaaggaacactctattcagatgccatccaggaaagaaag gccagatggatggagggtccactgtgtatatttgtacagaatcatttataaagttttcta cacttctaaaaaaagtatttctaacttgggctgttggacagtttgtgtgcctctgagccg gcaagcctgttgtttcttgtgtttagtgc

SEQ ID NO	Target Sequence
25	gttctactccatacagttcacactgattgtgacacattcttagtagctagtgtctgttct agtcactgcactggagtctacgagccggaactcgctatatgcacgtgtgtgtgtccgtat gtaagaaagtgtgcaccgagtgactgaatggttgagatgaattggaatgctgaagactaa cgaagaaactagagactgatatcgagcattctgcccacctcgctctgtatttaattaatt gtgctatatgttgctttaacaacccattgagcagtcagggaatgtgagtaagcttgctgc ca
26	tggtttcaagtgaccctcacacgtcagcctcctaaagagntgaagttataagcatgaacc atntgtggccagcggtgggactgatccttacnnngntcttgtgtgcttcttatctctctt ggaaccacgtcatggccatgaaaacaagctcaagctagcctactgaaggatacaagaacc atagaggaaagctgagtcactttagctgaggccataatagactgatncagccccctgcca acaaccagtgatggcagatgcttgatcgggccccagaaagttcatcagaaatgnagtgag tcaaggtcacnaataactgatcagctgaaccatggactcaggaaaaatttttttaaagg tagttgtttgttttgtttttaacataccagtagctacatagcccatattgcaaatga
27	ctctacgtcttggctctacatgaagtactcagagctggttacagcaaaaaggatggaggg aaaattatacagcagacttggaacagaacatttgaaggtatcgccgggcaggtgtccata gatgccaacggagaccgatatggggatttctctgtgattgccatgactgatgtggaggcg ggcacccaggaggttattggtgattattttggaaaagaaggtcgttttgaaatgcggccg aatgtcaaatatccttggggccctttaaaactgagaatagatgaaaaccgaattgtagag catacaaacagctctccctgcaaatcatgtggcctagaagaatcggcagtgacaggaatt gtcgtgggggctttactaggagctggcttgctaatggccttctacttttttcaggaagaaa tacagaataaccattgagaggcgaacccagcaagaagaaagtaaccttggaaaacatcgg gaattacgggaagattccatcagatcccatttttcagt
28	tccgtctttagctgatagattctcaaatatccttgattttggatgttcagtatgtttgtg agagaggtttctgggaagactctctttttgccctcgggaaaaagcaaaatatcaatgttt gggtgactgtgtaaagctcagtgtgtaagaacatctttttgtctaggttttctttctgct ctttattgaagacaaacactcaccaaaaagaaaaataaaagttttcagagaaactaattt tctttggcaagagtattacttaatattttggcctcctaaagtttccctagttagtactcg gactcctgtgctaattgtcagcttacatatcattgtatagagactgtttattctgtacca aactgatttcaaaagtactacattgaaaataaaccggtgactgtttttcttcataaagtt ctgcgtttggcatcttcactctttccaaaatgtatctgtacatcagaaatgtcactattc caagtgtctttttagtgtggcctttagtatggcttccttttaatat
29	gaaactgcagcttctccataatttatcctgtggtgatcaaaatatctgatcttaccctgg gacacgtaaatggtcttcagttaaagaaatcctcagcaggttgcgagggaataggagact ttgtggagctgctggagggaactggattggacccttccaagatgacgcctttagctgatc tctgctaccccttttcatggcccggcccagatgaaagttggctgtgacaacactgtggtgc gcatggtctccagtggaaaacacgtaaatcgtgtgacttttgagtatcgtcagctggagc cgtacgagctggaaaacccaaatggaaacagtatcggggaattctgtttgtctggtcttt gaataaccaacccagtgatttacatgctgatagctaagtgagtttttaatggccattgtg tatgattttgatgcacaactagttaaaagcctttcataccagtcagtatttcccagcc
30	tggtttaattcacccactttagatgggtgaatgttatggtgtgtgaaatatctcagtaaa gcagttaaaaggaaaaagagctggaatgcactgattcaggaacttaatttcaggaaggaa aggtctgtatgtacacatttcactttaagcagaaaatctttcttcaagaaatgactttac tttctctttgcactgccagcacgtgagatactaacttttttaactagttgttcttctctag tctctacgttattagaatttttgctttcataatgtgaaacctttaagcaggagaagaaa atgttttcagatagtttcaaatacaccaaaaatgtttgaaacacaaaaatactggaatca aaccataatgcacttatt

SEQ ID NO	Target Sequence
31	aatggagtatgcttccctcttcattcagaagatacagaaagtattggagcccagatttat tcagtggcaatttttcttggtattaatttggccgcatttatcatcatagtttttttcctat ggaagcatgttttatagtgttcatcaaagtgccataacagcaactgaaatacggaatcaa gttaaaaaagagatgatccttgccaaacgtttttttctttatagtatttactgatgcatta tgctggatacccatttttgtagtgaaatttctttcactgcttcaggtagaaataccaggt accataacctcttgggtagtgatttttattctgcccattaacagtgctttgaacccaatt ctctatactctgaccacaagaccatttaaagaaatgattcatcggttttggtataactac agacaaagaaaatctatggacagcaaaggtcagaaaacatatgctccatcattcatctgg gtggaaatgtggccactgcaggagatgccacctgagttaatgaagccggacctttttcaca tac
32	ccagacaaatcaaatcggccctttccagcaaggagtggaacagcccaagggtggctagat cacatcttctgctcacaccaccactgagcaaccctctcnaggactgccacctangcatag gtcatgccctgtgcaagagcacataaagtgggagctgaagtcctgcccaggctccattca ccaagcagagcttatgttgttgtgtgtctgtctacccagaggatgtgggtccttttta aaatgtgcaccaatgcatcttgtatgagctggcaccagctcggaccatcccaccctcata ccctggtcctatccagggttatattccc
33	aacaatttacaacagtaagcacctgtacacaagccccatctccagtcctgtttctagggc ttccaatttaacacacctagaggctgggcata
34	agcagcttattgtttctctgaaagtgtgtgtagttttactttcctaaggaattaccaaga atatcctttaaaatttaaaaggatggcaagttgcatcagaaagctttattttgagatgta aaaagattcccaaacgtggttacattagccattcatgtatgtcagaagtgcagaattggg gcacttaatggtcaccttgtaacagttttgtgtaactcccagtgatgctgtacacatatt tgaagggtctttctcaaagaaatattaagcatgttttgttgctcagtgtttttgtgaatt gcttggttgtaattaaattctgagcctgatattgatatg
35	tgatcatagctactacttcattgcaacctttattactgaccacatcagacatcatgctaa atacctgaatgcatgaaaaaactccaaataagagaatctcttcaggattataaaagttgt aaaatgcaactgtattgctgagcaaaa
36	aatgagatatctacaaggcacttaaagtgttacagatgttttaccttaagaattatttaa gttgtgttgggttaagacagttttcagtgtaccgtaaatgttgtgttttcagaaaaagac aaaacgatggtgctgactggttttctgtatattgcacaacagtcctcaaatacactgatg tatgaaactattcatacatcaagcagcatttttttcactctccttagaattggaactatg cagttaaggcagataaaatgtacagatgtttcatatattacaggttacatatataaatca aaatttcctatataaaactgatttgggatttggggtggaaatattttgaatattaattta tttttaaagatgcaagataggactttgtgcaatgtatttttgtaaatgcttttcaaaata tctgtctttggtagtgcttctgctgctgccaccaaattgataagatg
37	tgtctctctcgctcatgtaatatactctgaccctgagtggaaaggggttttngttctgtt tttattttacctacatgtactatttagcttcagtgtactagtcctgccacctgtgtattt ttagggtgctatggaaataatgaaaagaaacggggatttcagaagaaaattgtaaccaaa ttcatactttgtataatttttgatatcatgatcacaggtgattcacacgtacacacataa acacacccaccagtgcagcctgaagtaactcccacagaaaccatcatcgtctttgtacat cgtatgtacaatgcaatcatttcatactttaaactggtcaaaaaactaattgtgatttct agtcttgcaaagctg
38	gagggacagcgatgtggccctctgtgttaagaataacgtgtcctgctttggcagagagaa gaaaatagccactgcccgctttcaaggcaagatcgacctttttctgttttgttttgttttt ctttcttttttcctggccatgaggacaaaaattactgagtggcccttaaagagggaagttt

SEQ ID NO	Target Sequence
	gttttcagctgttctcttttgcccgtaggtgggagggtggggattgctgcgtcctagcta gaggaatggctttgcttgaatgtgtagtgcacacgcacgggtgtttctgtgtgctagttg cttcttgctgctgcttcctgcttgtctgggactcacatacataacgtgnnatatatatat atatatataaatgtataaatatatattttatttttttttaaatccttggagcttctggtt cctatcagttcctgttgttaatcgtaga
39	ggtgagctctgattgcttcagttggtcatcaactattttcccttgacntgctgtcctngg gatggcctgctatcttgatgatagattgtgaatatcaggaggcagggatcactgtggacc atcttagcagttgacctaacacatcttcttttcaatatctaagaacttttgccactgtga ctaatggtcctaatattaagctgttgtttatatttatcatatatctatggctacatggtt atattatgctgtggttgcgttcggttttatttacagttgcttttacaaatatttgctgta acatttgacttctaaggtttagatgccatttaagaactgagatggatagcttttaaagca tcttttacttcttaccatttttttaaaagtatgcagctaaattcgaagcttttggtctata ttgttaattgccattgctgtaaatctt
40	ttcttgcatccttctgtgattcaaaaaagtaaaatgtggctttctgaaatgatggataag agtctacatcttctagaaaaaatacataaaggagtagttaagctctgtaaatgtgccacg agctccaacacgaccatcgtagggtgaagcccacgttttcttccatggcctcaaaggccc tagaacttgcctacctttctggccttacctcctagctacttatccatctcttgaacttta tactcttgtataaatttctaactttcagaaaatgccatactctgttttggcaccacacat gtatatttccccctggtacacttggaagactcttatccatctgtgaaaccctatgttgtc atcacttggtccatgaaatattacctggccaatatcccaccatcacctcaaacccaatca ccccctcctctgtatgctgtcacacctatattattaaacttatcacattgcattgtaatt acttcctgacctttgtatctactcttttagt
41	gatttccaggaagcctttgatcacctttgtaacaagatcattgccaccaggaacccagag gaaatccgaggggggaggcctgcttaagtactgcaacctcttggtgaggggctttaggccn gcctctgatgaaatcaagacccttcaaaggtatatgtgttccaggtttttcatcgacttc tcagacattggagagcagcagagaaaactggagtcctatttgcagaaccactttgtggga ttggaagaccgcaagtatgagtatctcatgacccttcatggagtggtaaatgagagcaca gtgtgccntgatgggacatgaaagaagacagactttaaaccttatcaccatgctggctat ccgggtgttagctgaccaaaatgtcattcctaatgtggctaatgtcacttgctattacca gccagccccctatgtagcagatgccaactttagcaattactacattgcacaggttcagcc agtattcacgtgccagcaacagacctactccacttggctaccctgcaattaaga
42	taggccatttggactctgccttcaaacaaaggcagttcagtccacaggcatggaagctgt gaggggacaggcctgtgcgtgccatccagagtcatctcagccctgcctttctctggagca ttctgaaaacagatattctggcccagggaatccagccatgacccccacccctctgccaaa gtactcttaggtgccagtctggtaactgaactccctctggaggcaggcttgagggaggat tcctcagggttcccttgaaagctttatttatttatttgttcatttatttattggagagg cagcattgcacagtgaaagaattctggatatctcaggagccccgaaattctagctctgac tttgctgtttccagtggtatgaccttggagaagtcacttatcctcttggagcctcagttt cctcatctgcagaataatgactgacttgtctaattcatagggatgtgaggttctgctgag g
43	gattgtctgaataggcatcctcatctatatttacccaaaacctcgcttactgtcatgtgc actacaaattgcaatttggaaacctactgtattgaaattctgtcagtttatggttcttga agactgatgtcctttcccaaacactggttactgcagcagcatttttaatgtgtaagtgaa gaaaaaaggccactaaggccaaagatttttaagaatcattgtacaaatcattatgttaa actatctaagctttgctgtaatactgttttctcttcaatatgtgatggtacaggaaggat gttaaatgaaggggtggtattgcaggagagca

SEQ ID NO	Target Sequence
44	atcctcactgatgatttcaagctaaagcaaaccaccttatacagagatctagaatctctt tatgttctccagaggaaggtggaagaaaccatgggcaggagtaggaattgagtgataaac aattgggctaatgaagaaaacttctcttattgttcagttcatccagattataacttcaat gggacactttagaccattagacaattgacactggattaaacaaattcacataatgccaaa tacacaatgtatttatagcaacgtataatttgcaaagatggactttaaaagatgctgtgt aactaaactgaaataattcaattacttattatttagaatgttaaagcttatgatagtctt ttctaattcttaacactcatacttgaaatctttccgagtttccccagaagagaatatggg attttttttgacatttttgacccatttaataatgctcttgtgtttacctagtatatgtag actttgtcttatgtgtcaaaagtcctaggaaagtggttgatgtt
45	ctgttttctgtttagtcctcaatcttcctaactcaaattggggactgaggagagagaaag gtggttacccctgttaccgtgccatattcttcttgctgcttttcaaccccacgtgattgt tgattgacggttctgctataatgtgcgtgcccttcaagtttcagaaaactttcccaatca tttcacttcaatcttaattgaacccaagagtcaaagttattattttctccgaacgtgttt gtgatcttctgttatattttggggcatgttacctttatggtatataagctgtagtgcata ctctttgtattgcaaaaaactggtcagtaatttatgtacatgtattccacattttagtgt gcttgaagtgacaatccatagtttgtagtagtttgttatttgtcaactttaccctgtgtt ttaaggacatctaaacattccttgtcctatcaagatgacaaaagcagaatgtaatttttt tttggaagcttcgtgattacctgtaaca
46	tccgtctttagctgatagattctcaaatatccttgattttggatgttcagtatgtttgtg agagaggtttctgggaagactctcttttgccctcgggaaaaagcaaaatatcaatgttt gggtgactgtgtaaagctcagtgtgtaagaacatctttttgtctaggttttctttctgct ctttattgaagacaaacactcaccaaaaagaaaaataaaagttttcagagaaactaattt tctttggcaagagtattacttaatattttggcctcctaaagtttccctagttagtactcg gactcctgtgctaattgtcagcttacatatcattgtatagagactgtttattctgtacca aactgatttcaaaagtactacattgaaaataaaccggtgactgtttttcttcataaagtt ctgcgtttggcatcttcactctttccaaaatgtatctgtacatcagaaatgtcactattc caagtgtctttttagtgtggcctttagtatggcttccttttaatat
47	gacaactgcgtgggtccaaacactcctcttcctccaggtcatttgttttgcatttttaat gtctttatttttttgtaatgaaaaagcacactaagctgcccctggaatcgggtgcagctga ataggcacccaaaagtccgtgactaaattccgtttgtctttttgatagcaaattatgtta agagacagtgatggctagggctcaacaattttgtattcccatgtttgtgtgagacagagt ttgttttcccttgaacttggttagaattgtgctactgtgaacgctgatcctgcatatgga agtcccactttggtgacatttcctggccattcttgtttccattgtgtggatggtgggttg tgcccacttcctggagtgagacagctcctggtgtgtagaattcccggagcgtccgtggtt cagagtaaac
48	atcattctctttttgctgtatagaattgcttatatcactctttctttcatgacattggtt aacatttaaatgttcctcctgtacttgtgttgtctgtgaccgcttatagagttttattgt tattggtgtttacctgaatacctatgcgtacacacacacatatttccttaatacttctga actcattatctttttagaataataatactactacactttaccagcaattaacttctcccta cccaaaatgtttttcttcctgtctgaaaatggaactaatttgtcttattcgtgcttatat ctgtat
49	ttgtgtatgtattaccacagggcttnaatatcttaataagattttagttaatgtcatttt catgaaagacagtcatttggaagccaagagctctatctatgttttgttcaaatagccgag taatctactatcttggcagtaattacttctgctgcat

Claims

1. A method for detecting leukemia from patient samples and to preferably classify it into specific types of leukemia, comprising the steps of:

a. determining the expression level of at least one marker gene; and

b. comparing the expression level of the at least one marker gene to a reference, wherein an up-regulated expression level of the at least one marker gene predicts leukemia, preferably a specific type of leukemia;

wherein the at least one marker gene corresponds to at least one of the genes listed in one or more of Tables I to VI and/or is represented by at least one probe set or at least one target sequence listed in one or more of Tables I to VI.

2. Method according to claim 1, wherein ALL type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table I and/or is represented by at least one probe set or at least one target sequence listed in Table I.

3. Method according to claim 1, wherein AML type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed

in Table II and/or is represented by at least one probe set or at least one target sequence listed in Table II.

4. Method according to claim 1, wherein CLL type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table III and/or is represented by at least one probe set or at least one target sequence listed in Table III.

5. Method according to claim 1, wherein CML type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table IV and/or is represented by at least one probe set or at least one target sequence listed in Table IV.

6. Method according to claim 1, wherein hairy cell leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table V and/or is represented by at least one probe set or at least one target sequence listed in Table V.

7. Method according to claim 1, wherein multiple types of leukemia are detected and the at least one marker gene corresponds to at least one of the genes listed in Table VI and/or is represented by at least one probe set or at least one target sequence listed in Table VI.

8. Method according to claim 1, wherein the expression levels of a group of at least two marker genes are determined, wherein each marker gene of the

group of at least two marker genes corresponds to one of the genes listed in Tables I to VI and/or is represented by one probe set or one target sequence listed in Tables I to VI.

9. Method according to any one of the preceding claims, wherein at least one contrasting gene corresponding to at least one of the genes listed in Table VII and/or being represented by at least one probe set or at least one target sequence listed in Table VII is used to increase the discriminating power of the method for detecting leukemia from patient samples.

10. Method according to any one of the preceding claims, wherein the patient sample is blood sample, preferably peripheral blood mononuclear cells (PBMC), or bone marrow.

11. Method according to any one of the preceding claims, wherein the reference is a control sample from a healthy person and the marker gene in the patient sample is up-regulated compared to the control sample.

12. Method according to any one of the preceding claims, wherein the reference is an internal reference gene within the patient sample, which is invariant between leukemic and non-leukemic samples or which is down-regulated in leukemic samples.

13. Method according to claim 8, wherein the reference is a template signature of the group of marker genes derived from expression levels of the group of marker genes from non-leukemic samples.

14. A kit for detecting leukemia from patient samples and to preferably classify it into specific types of leukemia, the kit comprising at least one reagent for determining the expression level of at least one marker gene and optionally at least one contrasting gene, wherein the at least one marker gene corresponds to at least one of the genes listed in one or more of Tables I to VI and/or is represented by at least one probe set or at least one target sequence listed in one or more of Tables I to VI and wherein the at least one contrasting gene corresponds to at least one of the genes listed in Table VII and/or is represented by at least one probe set or at least one target sequence listed in Tables VII.

15. Kit according to claim 14, comprising at least two, preferably three, marker genes and preferably one or more contrasting gene.

16. Use of at least one marker gene for detecting leukemia from patient samples, preferably a specific type of leukemia, wherein the at least one marker gene corresponds to at least one of the genes listed in one or more of Tables I to VI and/or is represented by at least one probe set or at least one target sequence listed in one or more of Tables I to VI.

17. Use according to claim 16, wherein ALL type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table I and/or is represented by at least one probe set or at least one target sequence listed in Table I.

18. Use according to claim 16, wherein AML type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table II and/or is represented by at least one probe set or at least one target sequence listed in Table II.

19. Use according to claim 16, wherein CLL type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table III and/or is represented by at least one probe set or at least one target sequence listed in Table III.

20. Use according to claim 16, wherein CML type leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table IV and/or is represented by at least one probe set or at least one target sequence listed in Table IV.

21. Use according to claim 16, wherein hairy cell leukemia is detected and the at least one marker gene corresponds to at least one of the genes listed in Table V and/or is represented by at least one probe set or at least one target sequence listed in Table V.

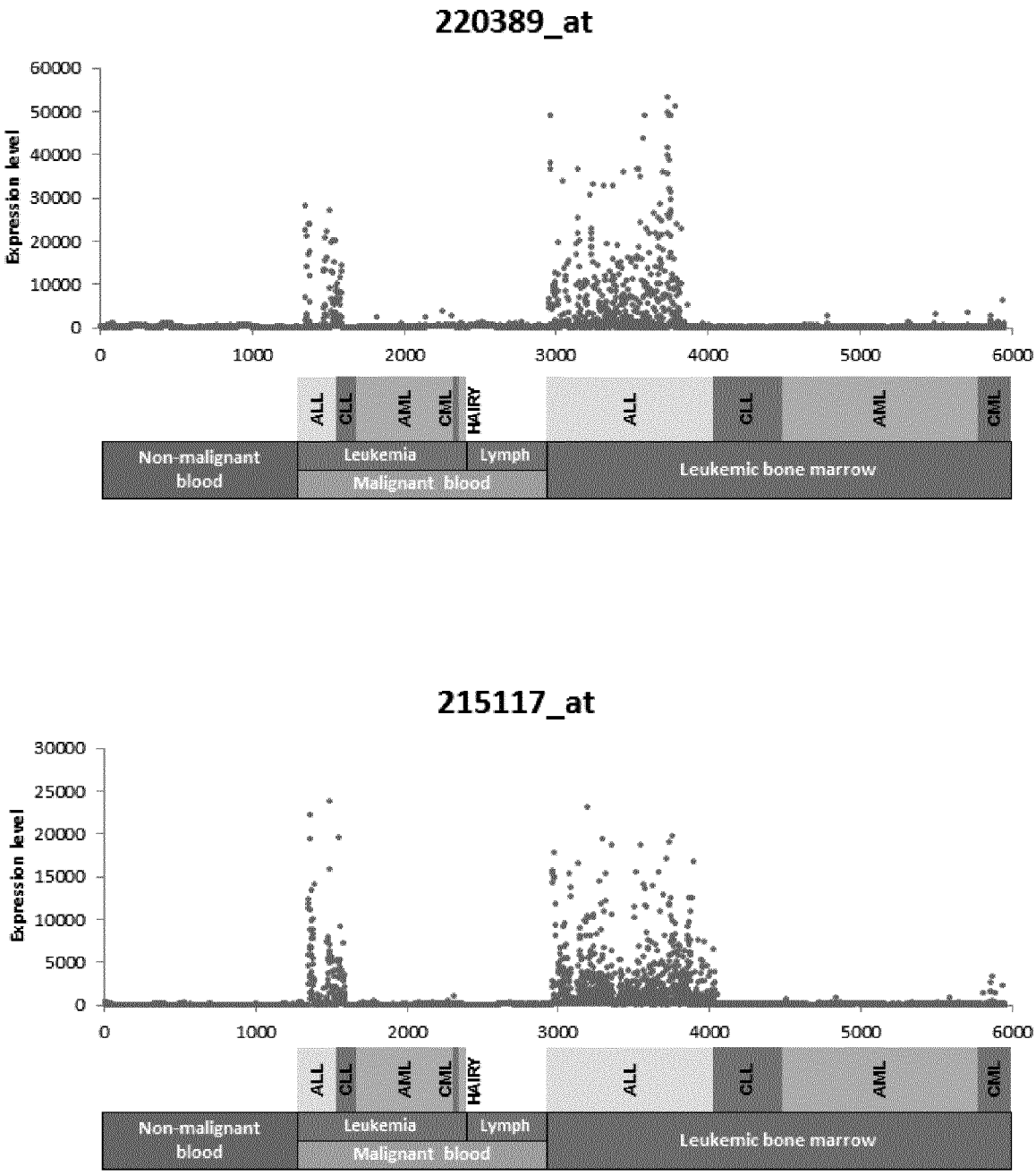
22. Use according to claim 16, wherein multiple types of leukemia are detected and the at least one marker gene corresponds to at least one of the genes listed in Table VI and/or is represented by at least one probe set or at least one target sequence listed in Table VI.

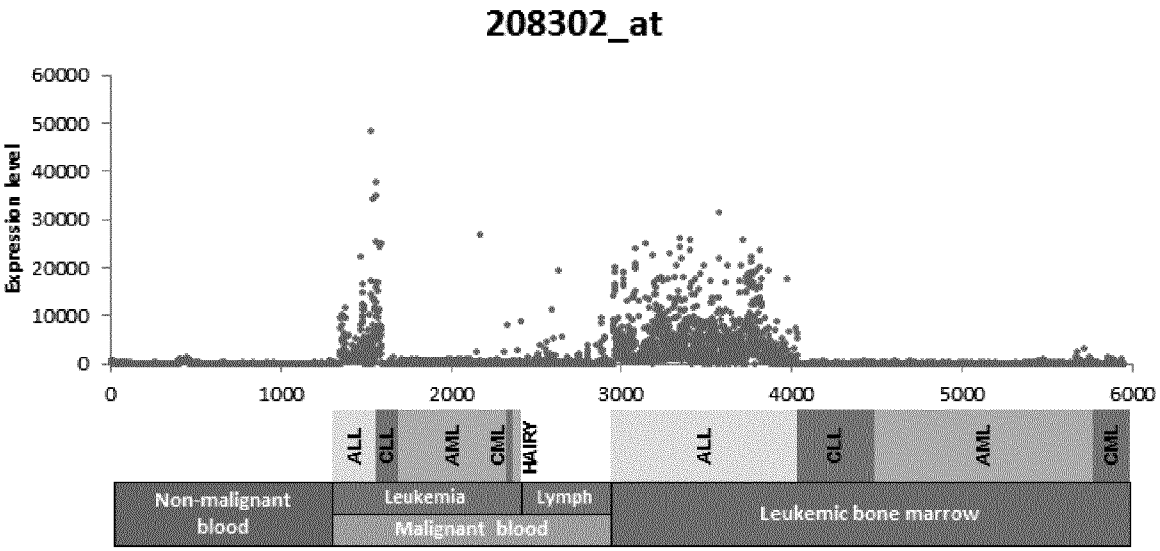
23. Use according to claim 16, wherein a group of at least two marker genes are selected, wherein each marker gene of the group of at least two marker genes corresponds to one of the genes listed in Tables I to VI and/or is represented by one probe set or one target sequence listed in Tables I to VI

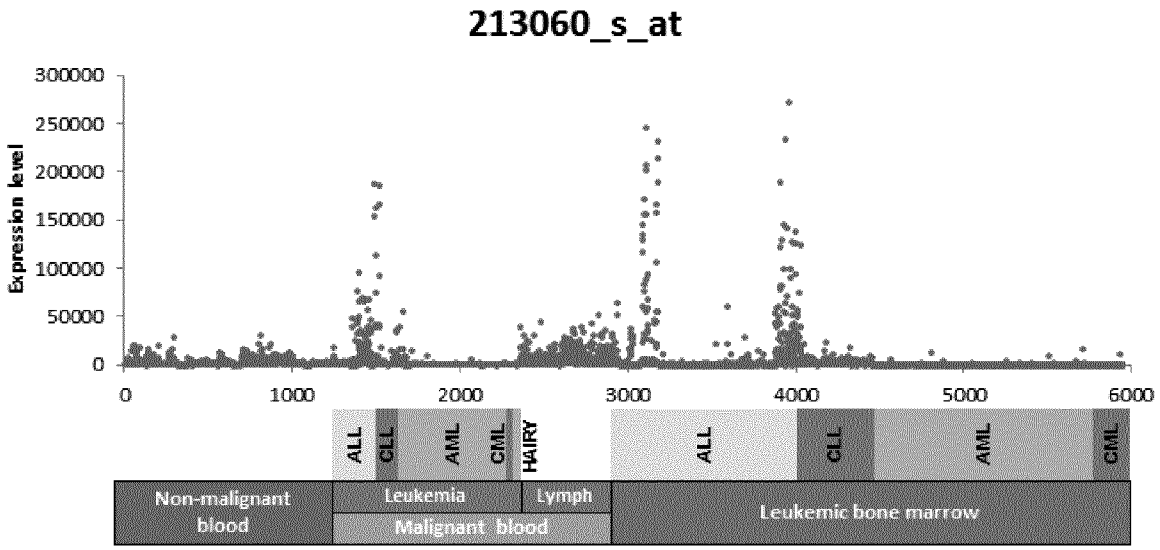
24. Use of at least one contrasting gene corresponding to at least one of the genes listed in Table VII and/or being represented by at least one probe set or at least one target sequence listed in Table VII to increase the discriminating power of a method for detecting leukemia from patient samples.

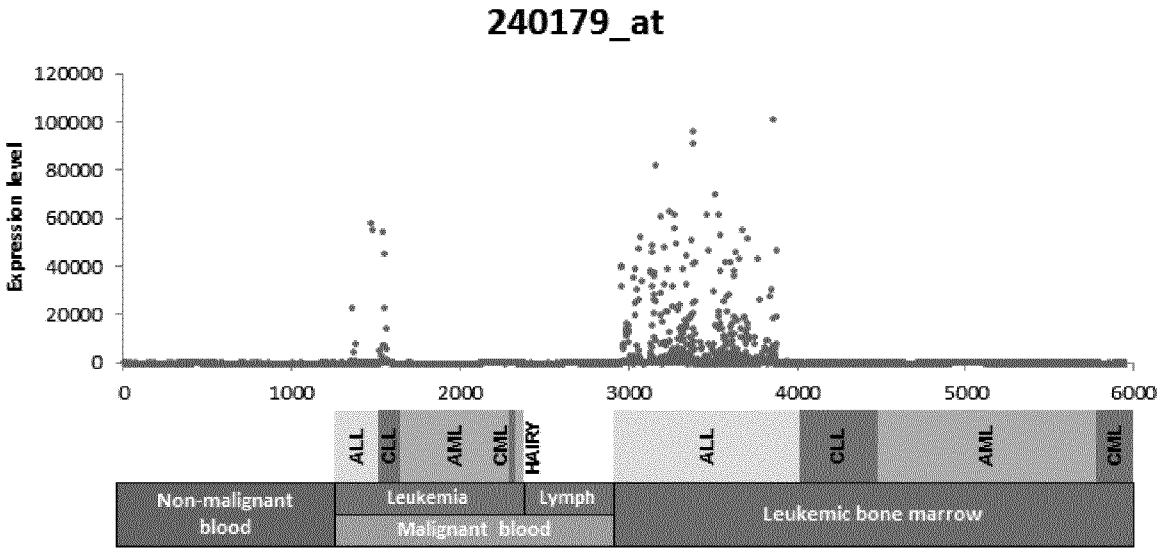
25. A reagent for detecting leukemia from patient samples and to preferably classify it into specific types of leukemia, the reagent being suitable for determining the expression level of at least one marker gene, wherein the at least one marker gene corresponds to at least one of the genes listed in one or more of Tables I to VI and/or is represented by at least one probe set or at least one target sequence listed in one or more of Tables I to VI.

Fig. 1(a)


220389_at
Expression level
60000
50000
40000
30000
20000
10000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
215117_at
Expression level
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow


208302_at
Expression level
60000
50000
40000
30000
20000
10000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow


213060_s_at
Expression level
300000
250000
200000
150000
100000
50000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow


240179_at
Expression level
120000
100000
80000
60000
40000
20000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Lymph
Malignant blood
Leukemic bone marrow

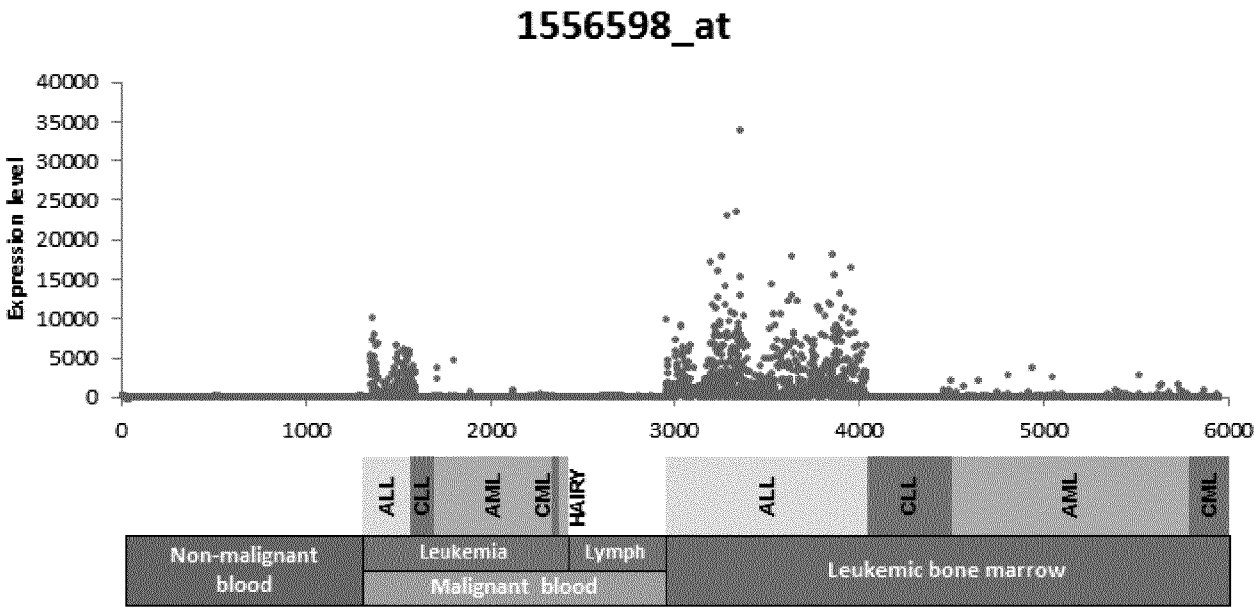
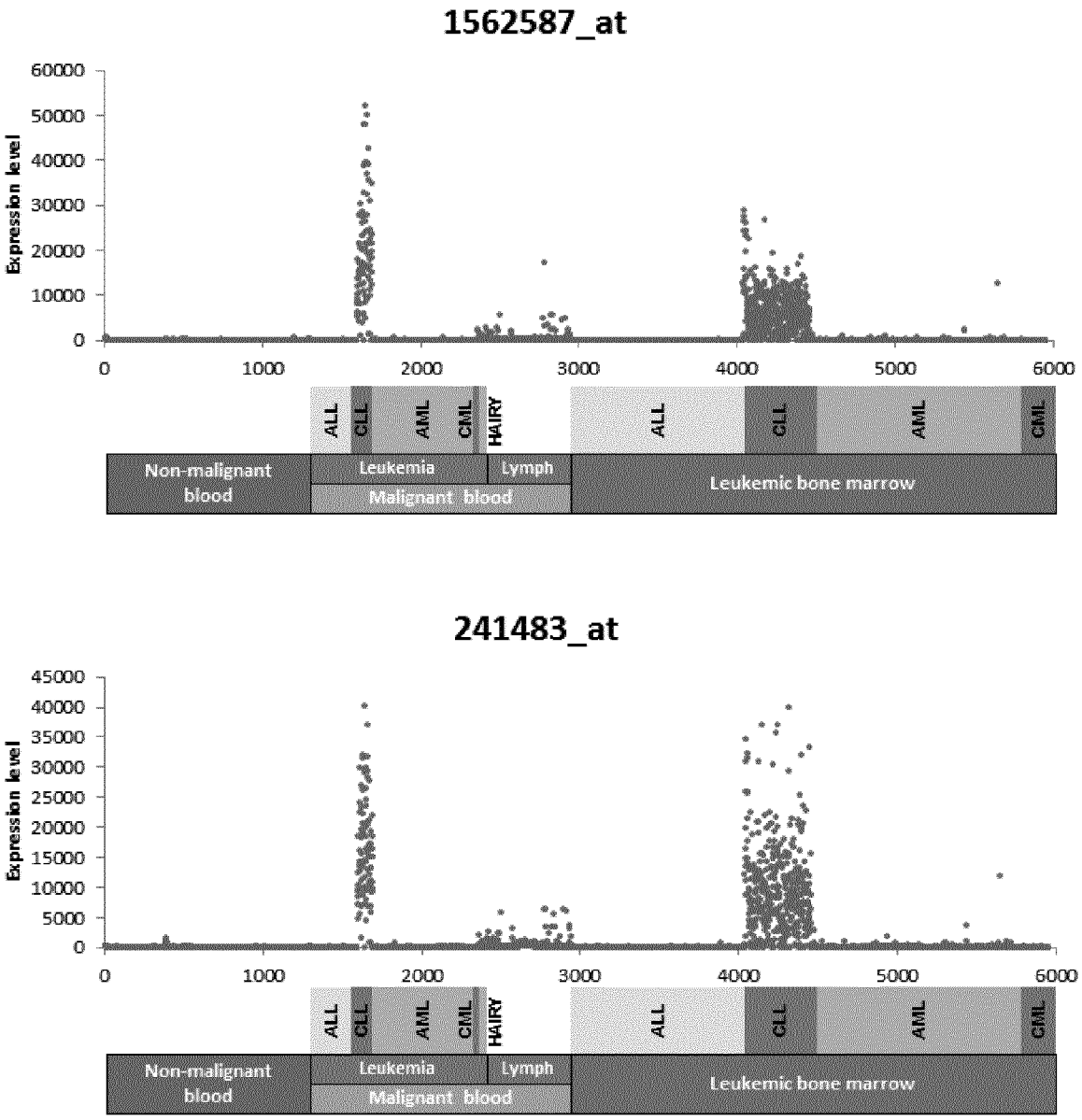
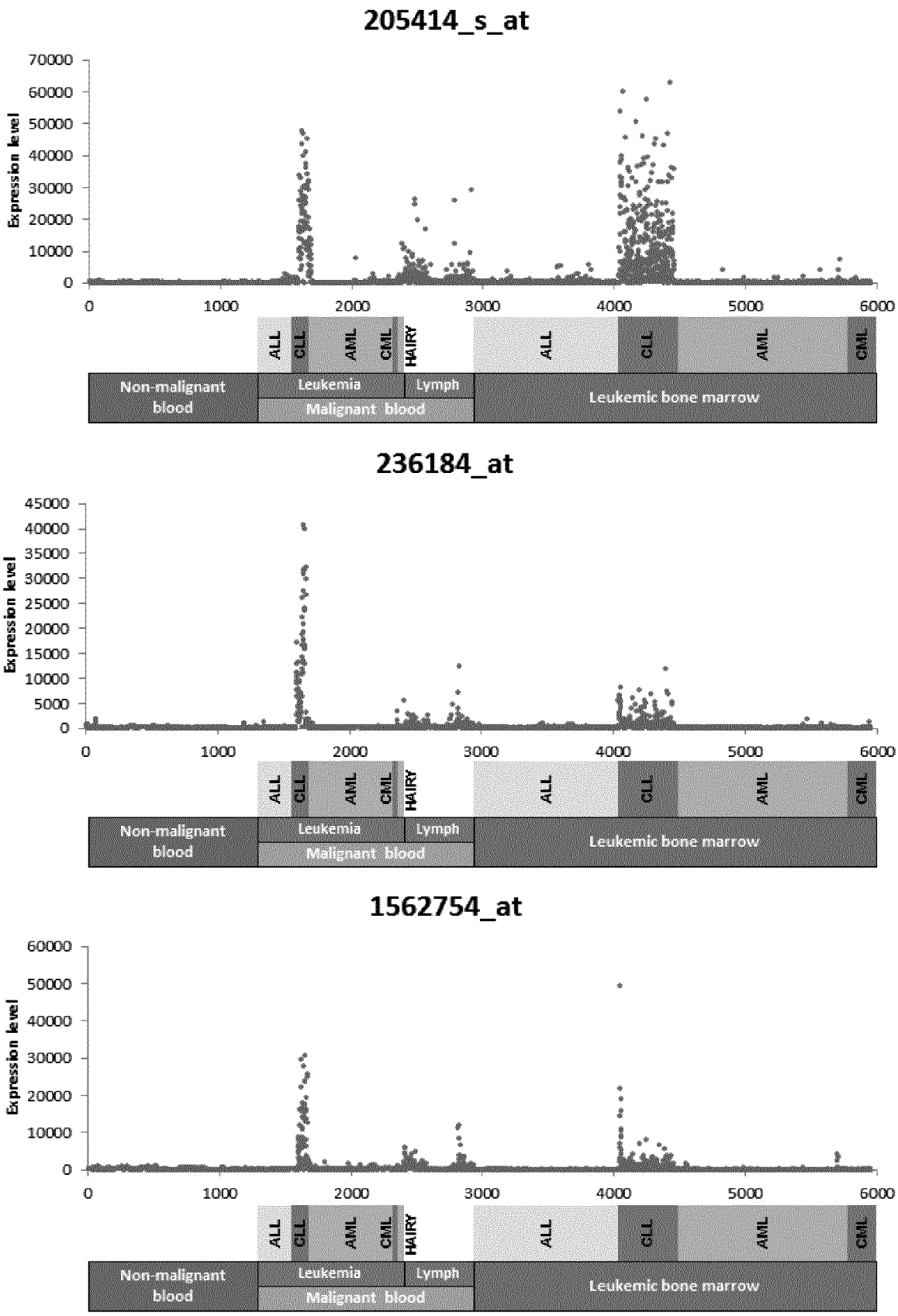
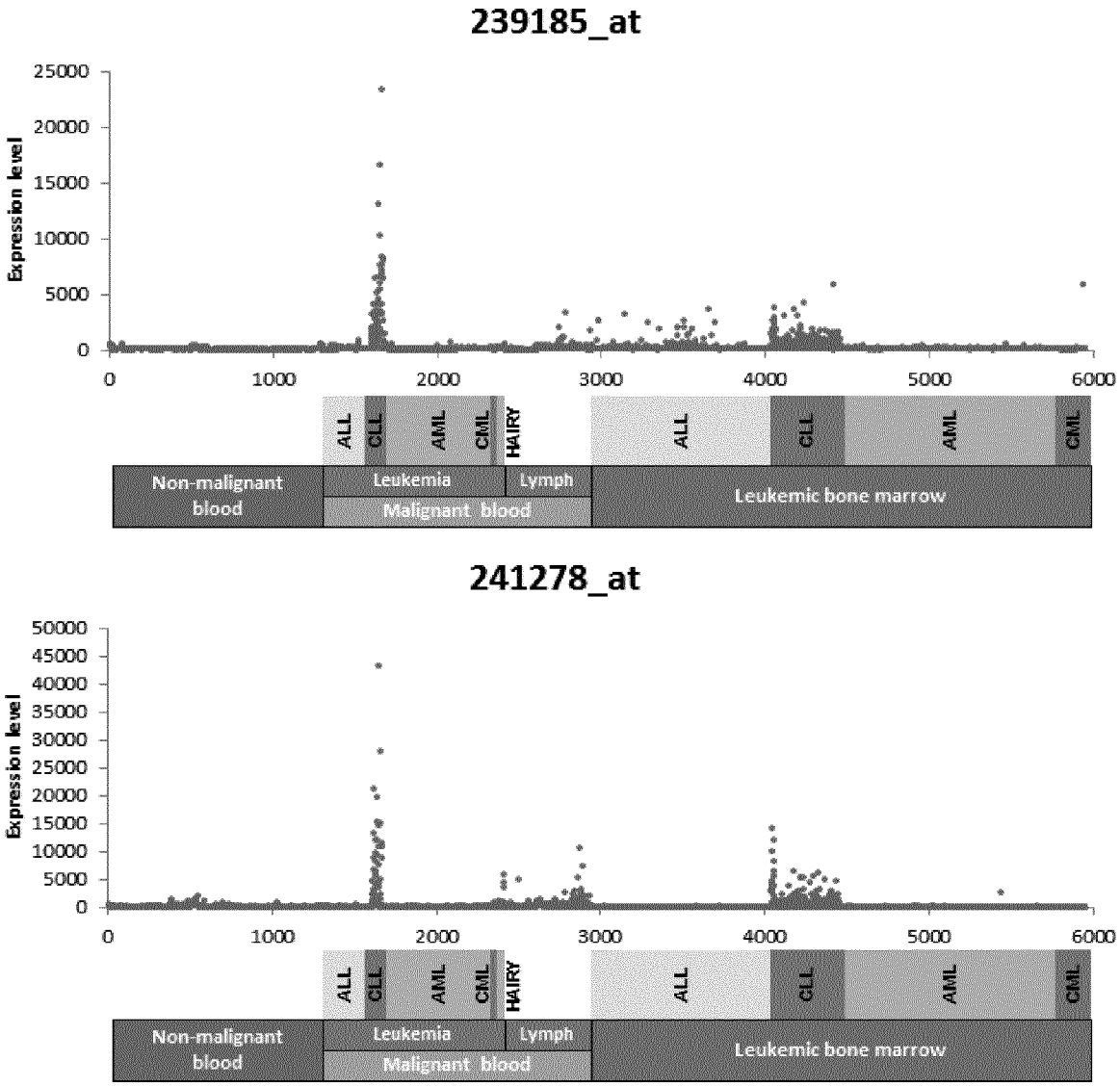

1556598_at
Expression level
40000
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Lymph
Malignant blood
Leukemic bone marrow

Fig. 1(b)

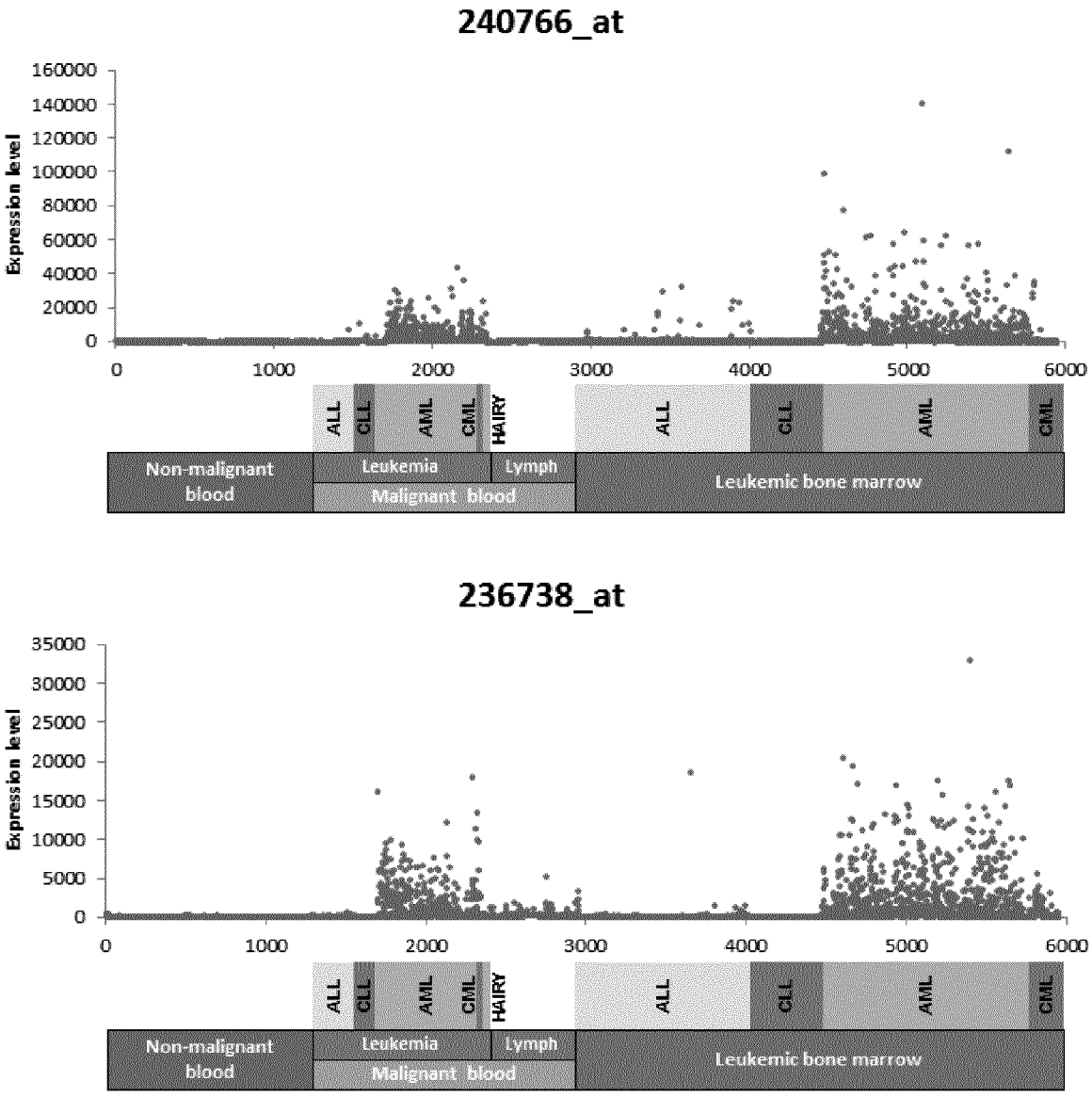

1562587_at
Expression level
60000
50000
40000
30000
20000
10000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
241483_at
Expression level
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow


205414_s_at
Expression level
70000
60000
50000
40000
30000
20000
10000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
236184_at
Expression level
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
1562754_at
Expression level
60000
50000
40000
30000
20000
10000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow


239185_at
Expression level
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow

241278_at
Expression level
50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow

Fig. 1(c)


240766_at
Expression level
160000
140000
120000
100000
80000
60000
40000
20000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
236738_at
Expression level
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow

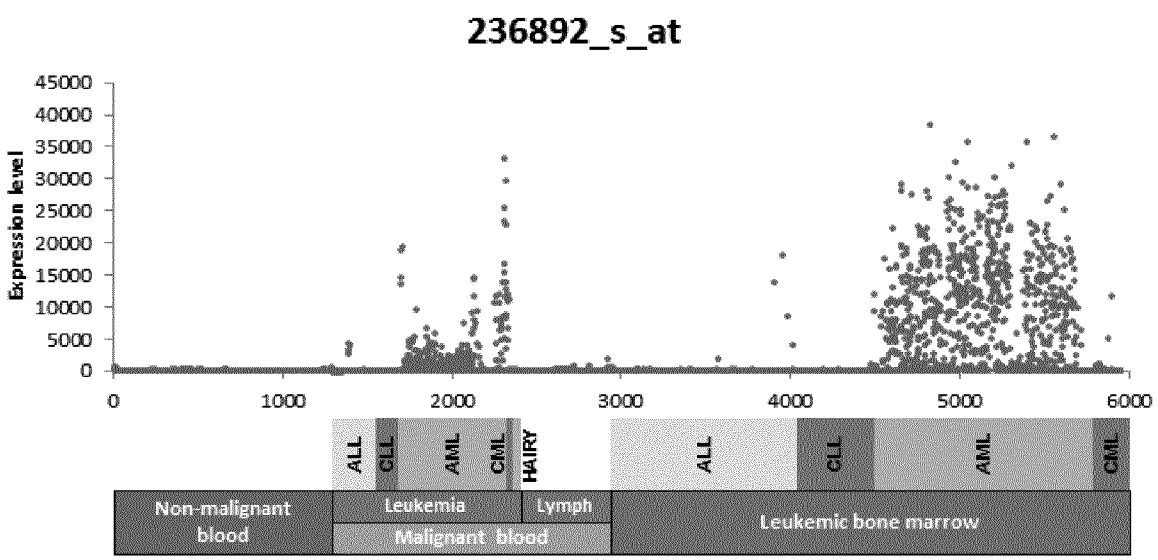

236892_s_at
Expression level
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow

Fig. 1(d)

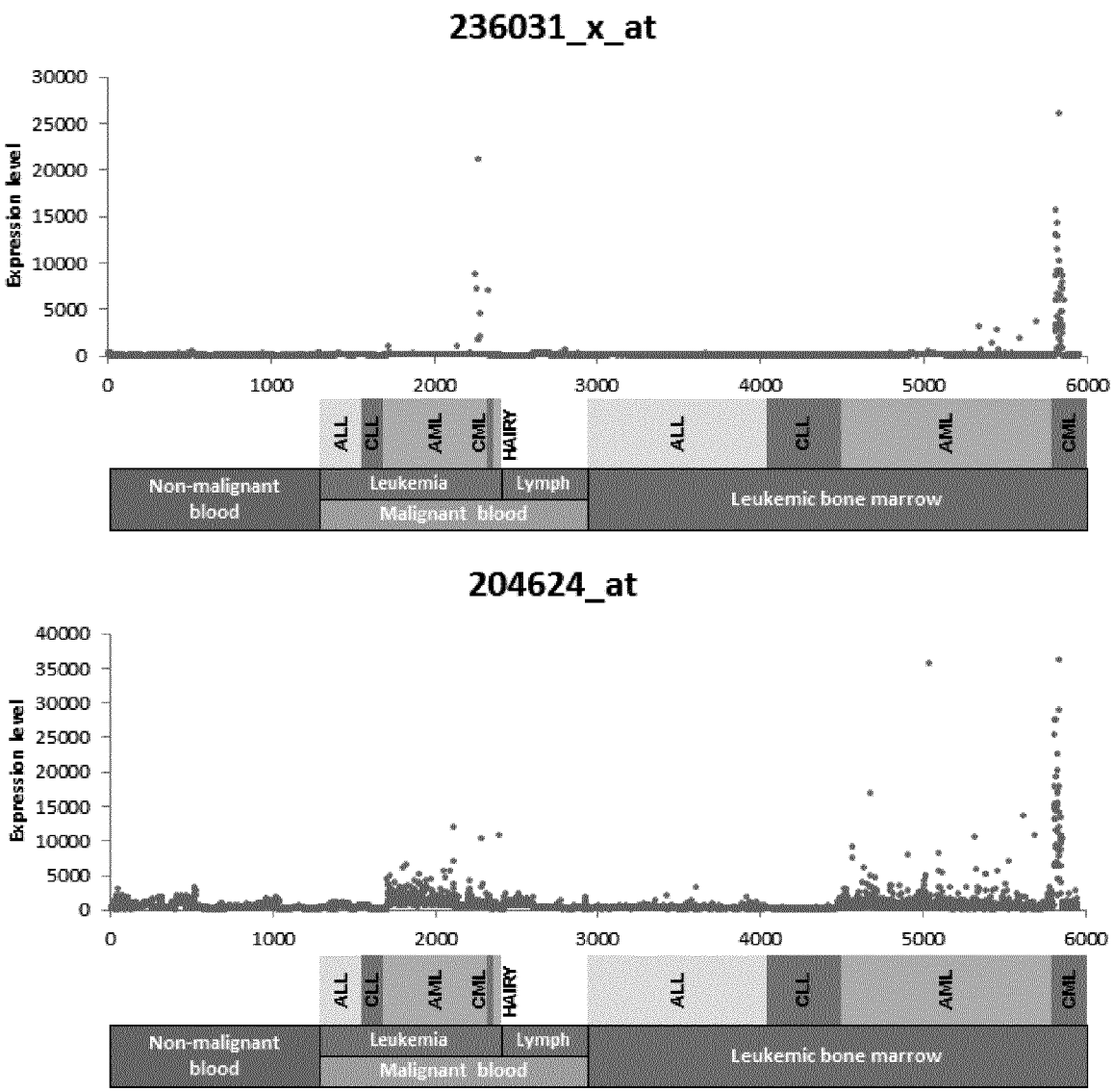

236031_x_at
Expression level
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Lymph
Malignant blood
Leukemic bone marrow
204624_at
Expression level
40000
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Lymph
Malignant blood
Leukemic bone marrow

Fig. 1(e)

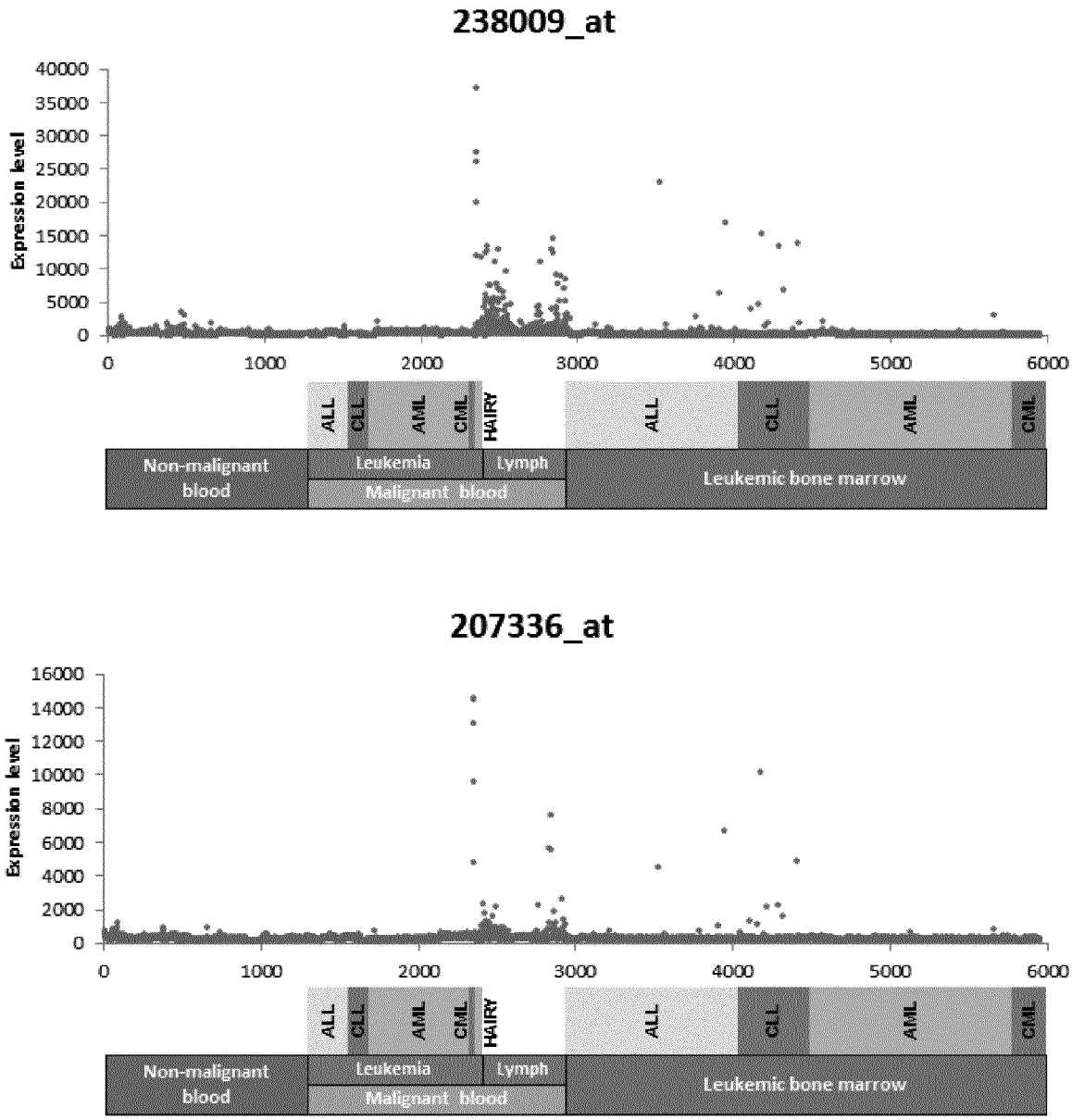
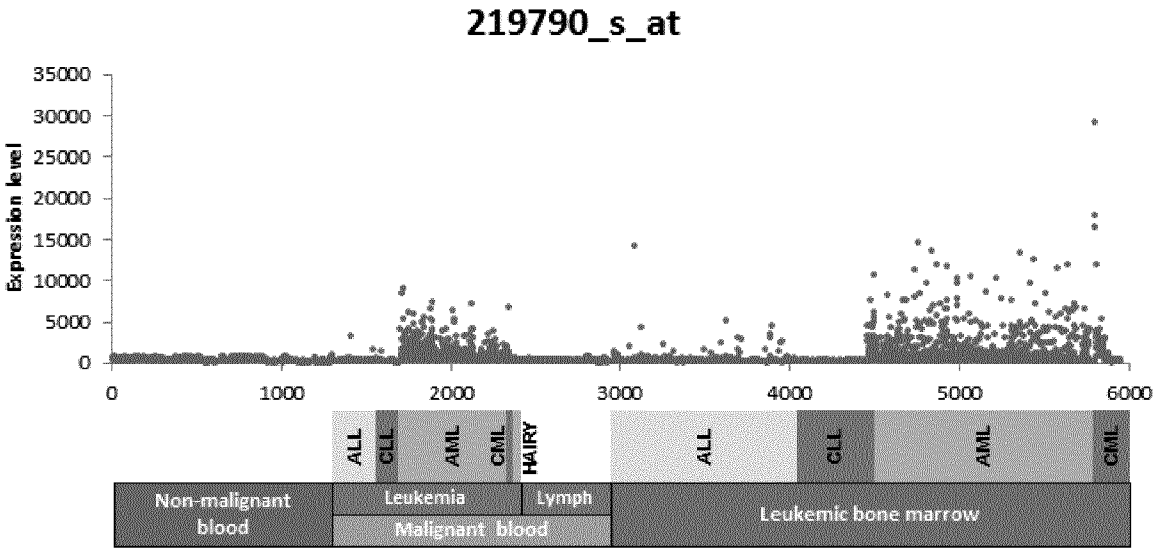
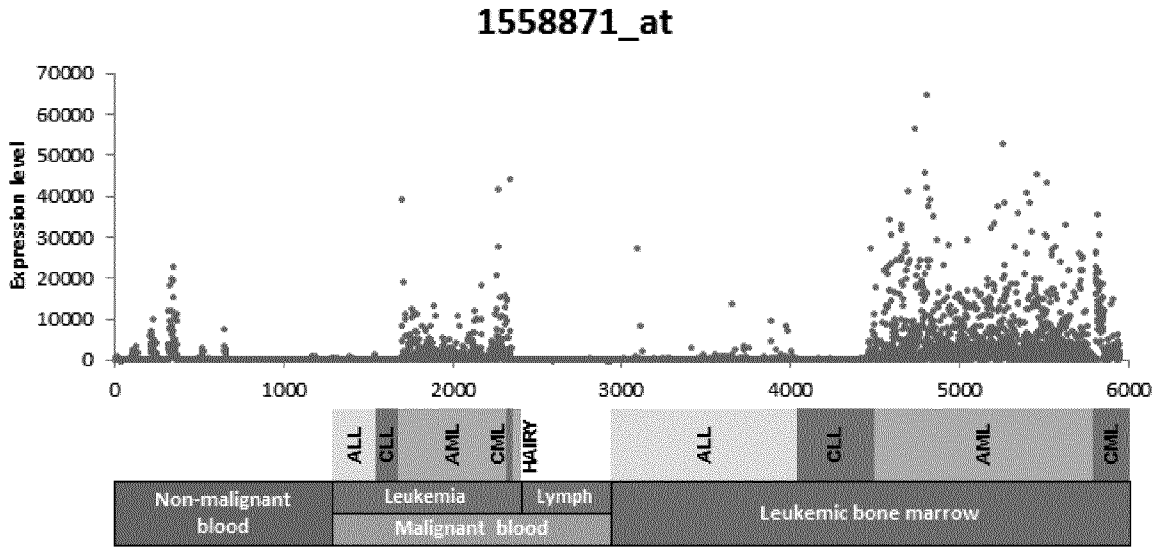
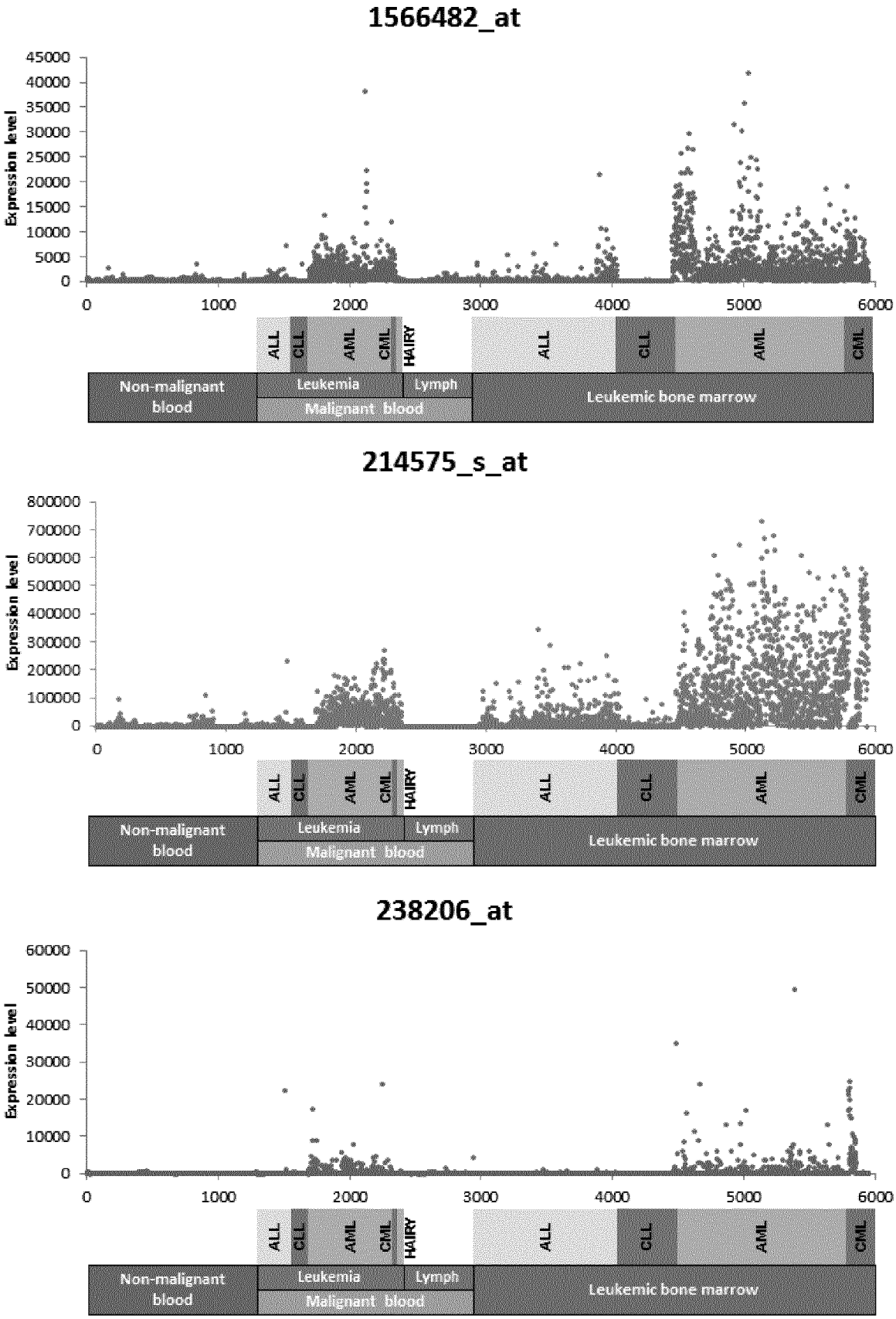
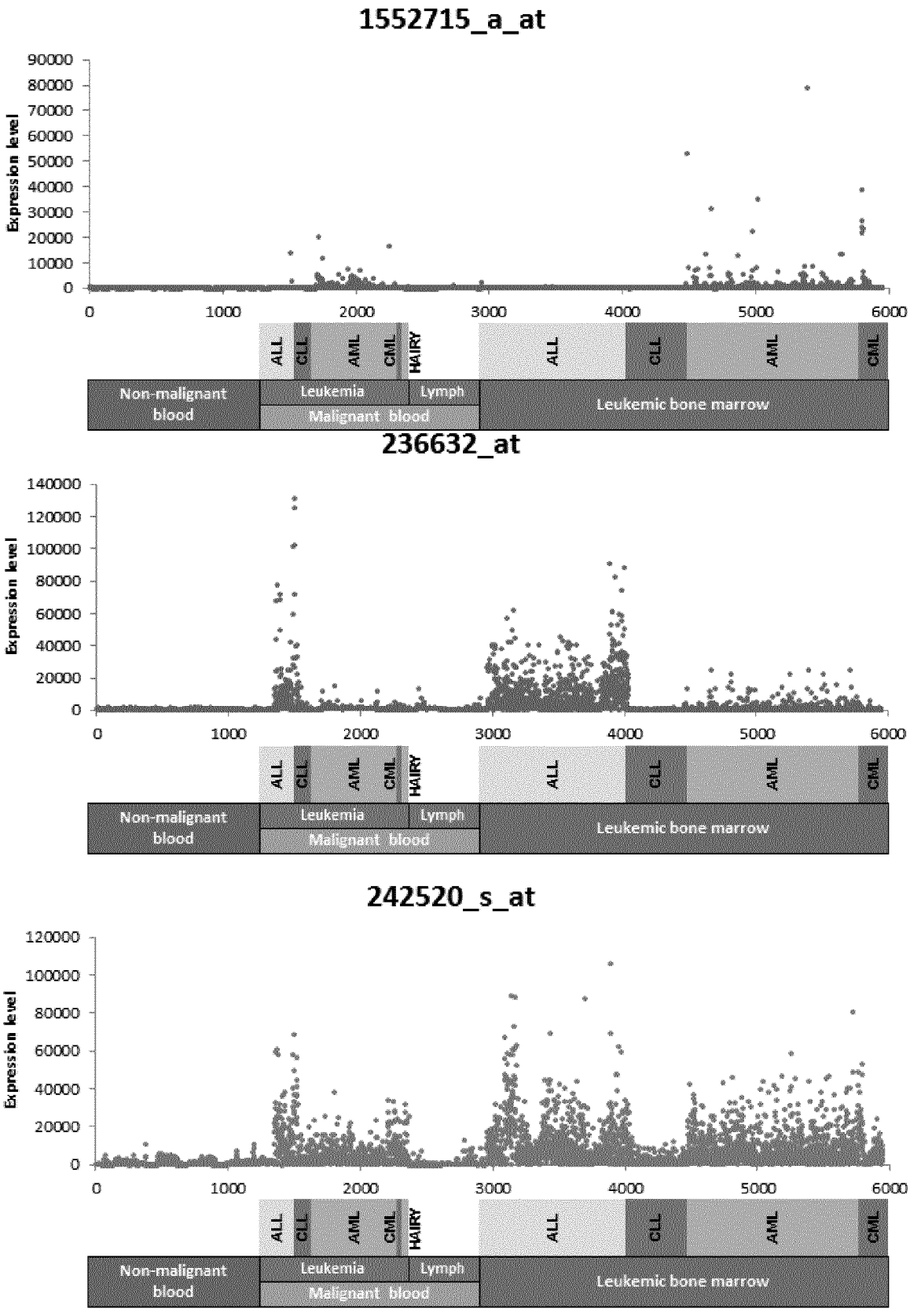

238009_at
Expression level
40000
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Lymph
Malignant blood
Leukemic bone marrow
207336_at
Expression level
16000
14000
12000
10000
8000
6000
4000
2000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Lymph
Malignant blood
Leukemic bone marrow

Fig. 1(f)


219790_s_at
Expression level
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow


1558871_at
Expression level
70000
60000
50000
40000
30000
20000
10000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow


1566482_at
Expression level
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
214575_s_at
Expression level
800000
700000
600000
500000
400000
300000
200000
100000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
238206_at
Expression level
60000
50000
40000
30000
20000
10000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow


1552715_a_at
Expression level
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
236632_at
Expression level
140000
120000
100000
80000
60000
40000
20000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
242520_s_at
Expression level
120000
100000
80000
60000
40000
20000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow

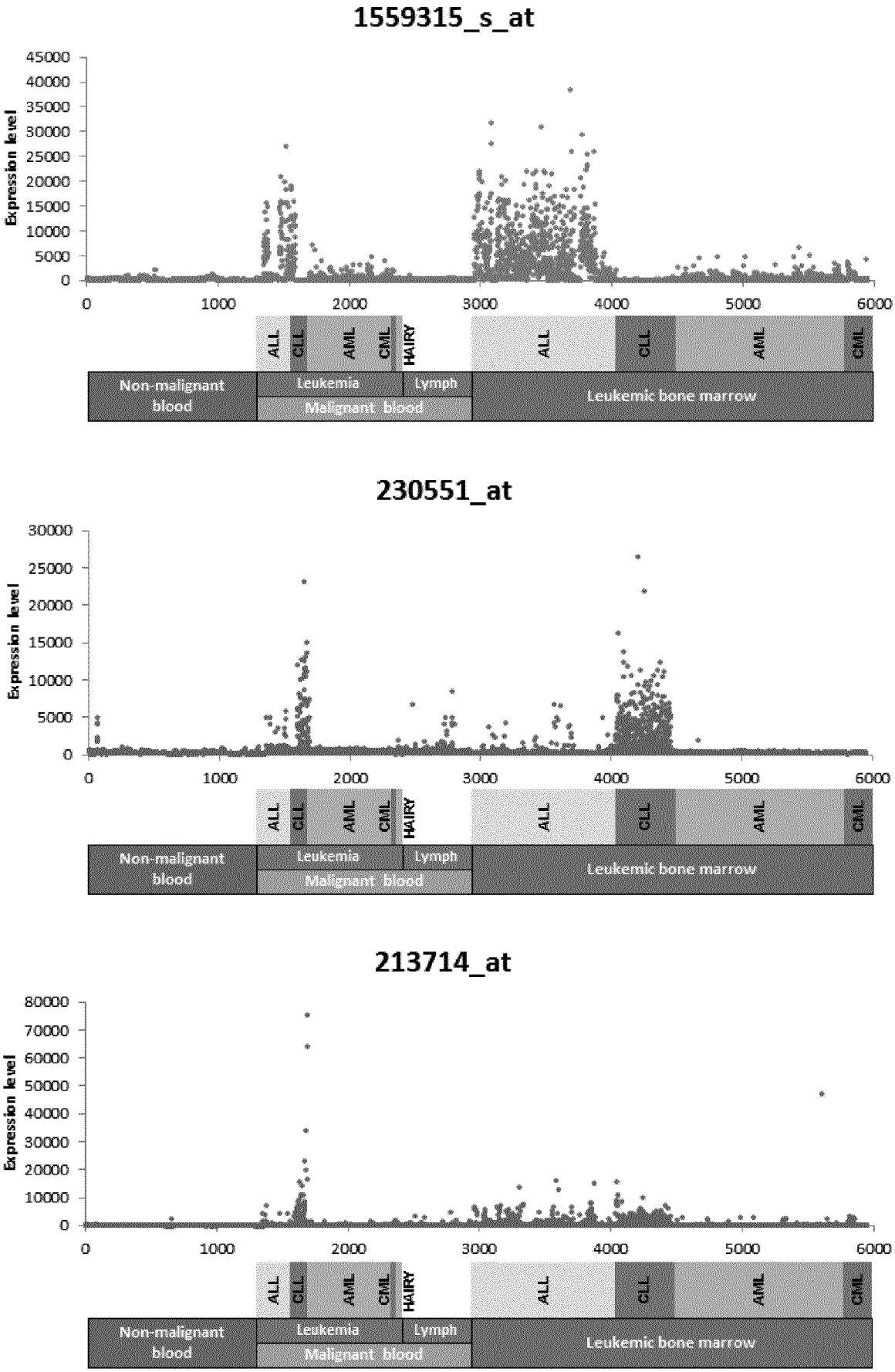
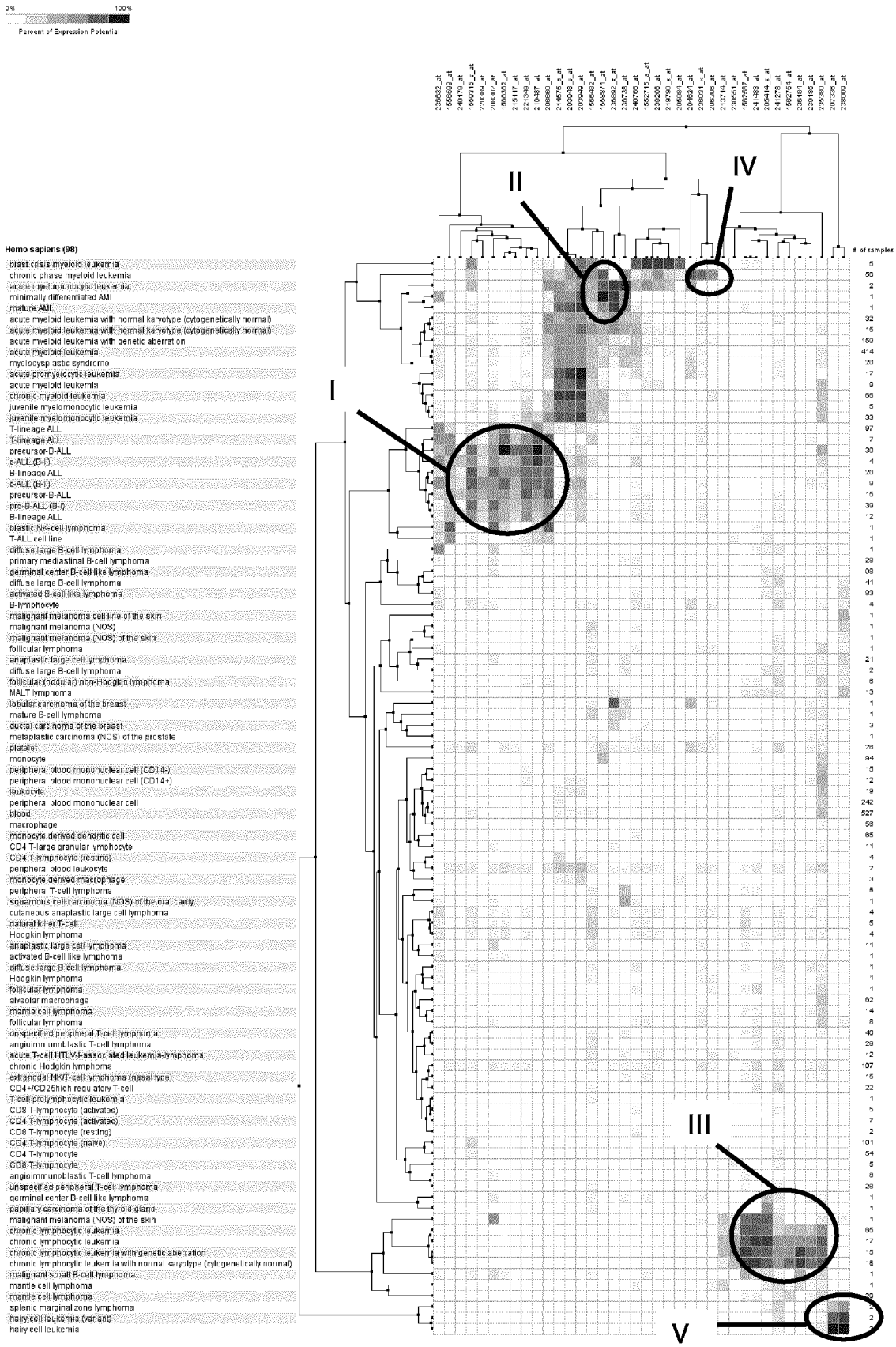

1559315_s_at
Expression level
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
230551_at
Expression level
30000
25000
20000
15000
10000
5000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow
213714_at
Expression level
80000
70000
60000
50000
40000
30000
20000
10000
0
0
1000
2000
3000
4000
5000
6000
ALL
CLL
AML
CML
HAIRY
ALL
CLL
AML
CML
Non-malignant blood
Leukemia
Malignant blood
Lymph
Leukemic bone marrow

Fig. 2


0%
100%
Percent of Expression Potential
Homo sapiens (98)
# of samples
I
II
III
IV
V
blast crisis myeloid leukemia
chronic phase myeloid leukemia
acute myelomonocytic leukemia
minimally differentiated AML
mature AML
acute myeloid leukemia with normal karyotype (cytogenetically normal)
acute myeloid leukemia with normal karyotype (cytogenetically normal)
acute myeloid leukemia with genetic aberration
acute myeloid leukemia
myelodysplastic syndrome
acute promyelocytic leukemia
acute myeloid leukemia
chronic myeloid leukemia
juvenile myelomonocytic leukemia
juvenile myelomonocytic leukemia
T-lineage ALL
T-lineage ALL
precursor-B-ALL
c-ALL (B-II)
B-lineage ALL
c-ALL (B-II)
precursor-B-ALL
pro-B-ALL (B-I)
B-lineage ALL
blastic NK-cell lymphoma
T-ALL cell line
diffuse large B-cell lymphoma
primary mediastinal B-cell lymphoma
germinal center B-cell like lymphoma
diffuse large B-cell lymphoma
activated B-cell like lymphoma
B-lymphocyte
malignant melanoma cell line of the skin
malignant melanoma (NOS)
malignant melanoma (NOS) of the skin
follicular lymphoma
anaplastic large cell lymphoma
diffuse large B-cell lymphoma
follicular (nodular) non-Hodgkin lymphoma
MALT lymphoma
lobular carcinoma of the breast
mature B-cell lymphoma
ductal carcinoma of the breast
metaplastic carcinoma (NOS) of the prostate
platelet
monocyte
peripheral blood mononuclear cell (CD14-)
peripheral blood mononuclear cell (CD14+)
leukocyte
peripheral blood mononuclear cell
blood
macrophage
monocyte derived dendritic cell
CD4 T-large granular lymphocyte
CD4 T-lymphocyte (resting)
peripheral blood leukocyte
monocyte derived macrophage
peripheral T-cell lymphoma
squamous cell carcinoma (NOS) of the oral cavity
cutaneous anaplastic large cell lymphoma
natural killer T-cell
Hodgkin lymphoma
anaplastic large cell lymphoma
activated B-cell like lymphoma
diffuse large B-cell lymphoma
Hodgkin lymphoma
follicular lymphoma
alveolar macrophage
mantle cell lymphoma
follicular lymphoma
unspecified peripheral T-cell lymphoma
angioimmunoblastic T-cell lymphoma
acute T-cell HTLV-I-associated leukemia-lymphoma
chronic Hodgkin lymphoma
extranodal NK/T-cell lymphoma (nasal type)
CD4+/CD25high regulatory T-cell
T-cell prolymphocytic leukemia
CD8 T-lymphocyte (activated)
CD4 T-lymphocyte (activated)
CD8 T-lymphocyte (resting)
CD4 T-lymphocyte (naive)
CD4 T-lymphocyte
CD8 T-lymphocyte
angioimmunoblastic T-cell lymphoma
unspecified peripheral T-cell lymphoma
germinal center B-cell like lymphoma
papillary carcinoma of the thyroid gland
malignant melanoma (NOS) of the skin
chronic lymphocytic leukemia
chronic lymphocytic leukemia
chronic lymphocytic leukemia with genetic aberration
chronic lymphocytic leukemia with normal karyotype (cytogenetically normal)
malignant small B-cell lymphoma
mantle cell lymphoma
mantle cell lymphoma
splenic marginal zone lymphoma
hairy cell leukemia (variant)
hairy cell leukemia