

(51) International Patent Classification:
G01N 33/50 (2006.01)(21) International Application Number:
PCT/EP2011/053314(22) International Filing Date:
4 March 2011 (04.03.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10075100.7 4 March 2010 (04.03.2010) EP(71) Applicant (for all designated States except US): **CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN** [DE/DE];
Charitéplatz 1, 10117 Berlin (DE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **REINKE, Petra** [DE/DE]; Hegelplatz 1, 10117 Berlin (DE). **VOLK, Hans-Dieter** [DE/DE]; Hegelplatz 1, 10117 Berlin (DE). **BABEL, Nina** [DE/DE]; Schonacher Str. 21, 13469 Berlin (DE).(74) Agent: **JUNGHANS, Claas**; Schulz Junghans Patentanwälte, Chausseestraße 5, 10115 Berlin (DE).

(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, RO, RS, RU, 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, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, 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).

Declarations under Rule 4.17:

- as to the identity of the inventor (Rule 4.17(i))
- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: HIGH THROUGHPUT ANALYSIS OF T-CELL RECEPTOR REPERTOIRES

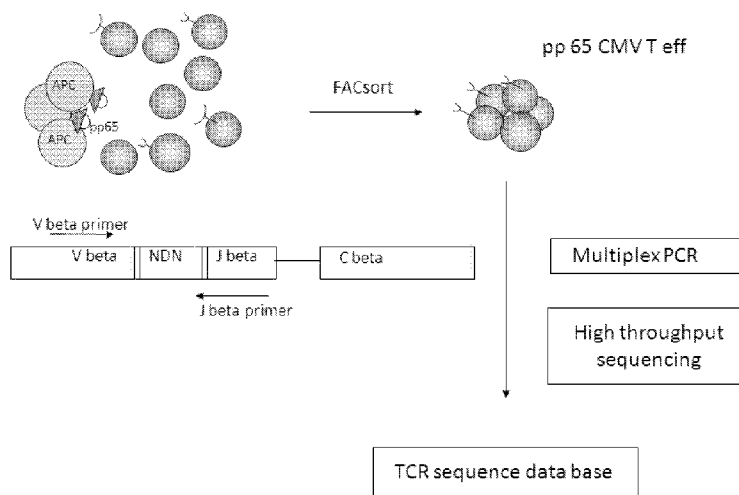


Fig. 1

(57) **Abstract:** A method for analysis of a T cell repertoire is provided, whereby T cell receptor protein sequences are obtained from lymphocytes, TCR sequences are aligned with a TCR sequence data base, and TCR sequences aligning to data base sequences are selected and their frequency distribution is determined. The TCR sequence data base is obtained by stimulating lymphocytes in vitro with antigens, selecting T cells responding to stimulation and obtaining TCR sequences from the selected cells, or comprises TCR specific for a tumour or an adoptively transferred T cell. Additionally, a method to obtain antigen-specific T cells for adoptive transfer is provided, whereby T cells responding to stimulation with antigens are selected, T cells comprising certain TCR sequences are obtained by culturing individual clones or by transferring a transgenic TCR into a T cell.

High throughput analysis of T-cell receptor repertoires

Description

Quantitative analysis of antigen-specific T cells in peripheral blood and tissue promises new insights into the aetiology and pathogenesis of diseases. Monitoring and manipulation of T cell receptor distribution among antigen-specific T cells may also allow for novel approaches in settings ranging from transplant management to cancer therapy.

Current techniques for T-cell monitoring include ELISPOT, cytokine flow cytometry, and MHC class I and II tetramer analysis. These methods however suffer from drawbacks with regard to their practical utility in the clinical setting. Monitoring of T cells in peripheral blood by standard methods requires fresh blood samples that have to be processed immediately. In addition, methods restricted to cells obtained from blood preclude analysis of tissue infiltrates, an important source of immune status information for transplant monitoring and other applications. Obtaining quantitative results requires laborious procedures and expensive reagents and machinery, and is confined to addressing problems where immunodominant epitopes and their HLA restriction are known and reagents exist for their analysis.

A method for the analysis of clonotypic composition in cytomegalovirus (CMV) specific CTL by manual sequencing of the variable V β chain CDR3 region of T cell receptors (TCR) is shown in Babel et al (J. Am. Soc. Nephrol. 20; 344-352, 2009). Therein, mono- or oligoclonal restriction of the TCR repertoire in CMV-specific T cells is analyzed by subcloning and traditional dideoxy sequencing after stimulation with a whole CMV pp 65 peptide library. The methods used in this work, however, are not satisfactory for clinical T-cell response monitoring, as only a limited number of sequences can be analyzed by conventional cloning, and sequencing of the TCR-CDR3 region might lead to biased results or deliver incomplete information.

Hence, methods found in the state of the art are dissatisfactory when applied to diagnostic objectives where sensitivity higher than

a few percent is required. One prominent example is the monitoring of T cell lymphoma patients, where quantification of tumour cells is important for monitoring both treatment efficiency and relapse occurrence. Especially for the latter objective, sensitivity in the percent range, as attained by current diagnostic methods, is of little clinical value. Sensitivity in the sub-per-mill range is desirable for relapse monitoring in order to commence relapse treatment as early as possible.

Conventional analysis of T cell heterogeneity has been limited to obtaining around one hundred sequences, making the method unsuitable for diagnostic objectives that require higher sensitivity or result accuracy than a few percent.

High-throughput or "deep" sequencing methods have recently been developed, delivering hundreds of millions bases of sequence information in a single day. High throughput sequencing uses massively parallel sequencing reactions, reading out large numbers of relatively short sequences which are aligned by computing.

Different technology platforms may be chosen for this aim. Examples are so-called pyrosequencing and nanopore sequencing. The main difference between the three platforms most commonly employed consists of the method of in vitro clonal amplification, and of the sequencing procedure itself. The commercially available Roche 454 and Solid ABI platforms use emulsion PCR for isolation of individual DNA molecules from each other. These single molecules are coated by primer-beads in aqueous droplets within an oil phase. Amplification of clonal copies of the DNA molecule by PCR is followed by immobilization for later sequencing. In contrast, the Illumina Genome Analyzer platform uses so-called bridge PCR for in vitro clonal amplification, where fragments are amplified upon primers attached to a solid surface.

Sequencing by ligation is a further method available for high-throughput sequencing tasks.

The objective of the present invention is to provide a method for analysis of the T cell receptor repertoire in lymphocytes obtained from a patient. Such method should allow clonal characterization of said repertoire, should not be restricted to blood lymphocytes and

should allow for a sensitivity of detection greater than available from methods of the state of the art.

This objective is attained by a method as specified in the independent claims.

- 5 According to a first aspect of the invention, a method for analysis of a patient's T cell repertoire is provided, whereby
- a. nucleic acid molecules encoding T cell receptor protein sequences are obtained from lymphocytes isolated from a patient;
 - b. at least one thousand of said nucleic acid molecules are
10 sequenced, rendering sample TCR sequences;
 - c. the sample TCR sequences are aligned with a TCR sequence data base, and sample TCR sequences positively aligning to data base sequences are selected as selected sample TCR sequences,
 - d. the frequency distribution of at least some of the selected
15 sample TCR sequences is determined.

In a first alternative, a TCR sequence database according to this aspect of the invention is constituted of sequences derived from antigen-specific TCRs. According to this alternative, the TCR database is obtained by stimulating lymphocytes isolated from a
20 patient in vitro with one or several antigens; selecting T cells responding to stimulation from said lymphocytes; and obtaining a plurality of nucleic acid sequences encoding T cell receptor proteins from at least some of the T cells selected in this step. In this case, the selected sample TCR sequences will be antigen-
25 specific TCR sequences. Such TCR sequence data base characteristic of antigen-specific T cells will be heterogeneous and will comprise a great variety of different TCR sequences, according to the heterogeneity of the T cell response mounted against the antigen.

Other applications however, such as T cell lymphoma monitoring, or
30 monitoring of adoptively transferred T cells, will use a TCR sequence data base comprised of only a few sequences or one single TCR sequence characteristic of the T cell clone, the frequency of which is to be monitored.

Hence, according to a second alternative of this first aspect of the
35 present invention, the TCR sequence data base comprises TCR sequences specific for either a T cell derived tumour having been

obtained from said patient previously, or for an adoptively transferred T cell previously transferred into said patient.

Nucleic acid molecules may be DNA or RNA, for example messenger RNA or genomic DNA encoding T cell receptor protein sequences. In the context of this invention, T cell receptor protein sequences are sequences that confer antigen specificity on the T cell receptor, i.e. the hypervariable parts of the TCR protein chain that are generated during T cell maturation.

TCR nucleic acid sequences arise during T cell maturation by recombination of V, D and J gene segments for the generation of beta and delta chains, V and J segments in the case of alpha and gamma chains. Sequence variations arising from somatic hypermutation give rise to further complexity of possible sequences. Different sequences may be grouped according to their use of germline V-, D- and J-sequences as V-, D- or J-families.

Correspondingly, nucleic acid molecules encoding T cell receptor protein sequences are genomic or mRNA-sequences that encode these antigen-specificity-conferring TCR peptide sequences. The greatest variety is found in beta chain sequences, however alpha, gamma or delta chain sequences can be employed to practice the invention.

Aligning in the context of the present invention is the process of comparing each sample TCR sequence base by base with each TCR data base sequence, and quantifying the result in a comprehensible manner, preferably by a single parameter of similarity. Methods of alignment are known in the art; the FASTA and BLAST programme packages available publicly are examples. Sample TCR sequences that positively align to data base sequences show a degree of similarity to the latter that indicates that they have the same or highly similar antigen specificity.

The plurality of nucleic acid sequences that is sequenced in step b according to the above description of this aspect of the invention must be a number sufficient to give rise to a result that provides statistically relevant information on frequency distributions of TCR sequences in the sub-percent range. A number that is sufficient to provide information in the sub-per-mill-range is preferred. Hence, the plurality of sequences may signify a number greater than one thousand, for example ten thousand, twenty-five thousand or fifty

thousand. In experiments carried out by the inventors, it was found that in order to retrieve information about the major immunodominant T cell clones within a patient's response to CMV, obtaining one hundred thousand sample TCR sequences did not materially improve the information in comparison to a second sequencing reaction during the course of which only thirty thousand sample TCR sequences were obtained.

The number above which significant information is not added by increasing the number of sample TCR sequences however may vary depending on the specific medical context. If recurrence of T cell tumour entities or transplant rejection monitoring is the objective of the method, far higher sensitivities will be warranted than if obtaining a general picture of the T cell response is intended.

A surprising advantage of the method according to the invention is the fact that sample TCR sequence distributions obtained by methods of the state of the art are qualitatively different from TCR sequence distributions obtained by the present invention. The methods of the state of the art make use of recombinant insertion of PCR amplicates into plasmids (subcloning), and subsequent clone picking and individual sequence analysis. This method introduces bias on a number of levels, most prominently during the steps of the subcloning process, but also during bacteria colony growth and selection. It was found that when comparing sample TCR sequence distribution obtained by high throughput sequencing to a smaller sample TCR sequence set obtained by subcloning, the high-throughput method according to the present invention afforded a more diverse sequence set than the traditional method. The dominant clone in CMV-specific TCR molecules was present in between 13% and 22% of the sample TCR sequences obtained by a method according to the invention, in comparison to 38% to 60% of clones analysed by the traditional method. These figures relate to the same original T cell pool, obtained from the same patient.

Sample TCR sequences may be in the form of nucleic acid (NA) sequences or amino acid (AA) sequences. The sequencing reaction will render NA sequences, which can then be translated to AA sequences using the known triplet code. Comparison is performed advantageously on the level of AA sequences, as the biologically meaningful

comparison to determine TCR specificity is performed on the AA sequence level.

One important advantage of this aspect of the present invention is the significant improvement in sensitivity when applied to monitoring of T cell derived tumours. Both monitoring of treatment response and, even more importantly, relapse monitoring, require the detection of small changes in tumour cell frequency, or very small total numbers of tumour cells, to provide optimal information to the physician. Current methods that provide information in the small percent range do not satisfy this requirement, as a tumour cell frequency of a few percent is equivalent to detecting a full relapse, whereas detection of recidivistic tumour cells in the sub-per-mill range, as enabled by the present invention, facilitates the opportunity to restart aggressive treatment regimens early on during the relapse.

A TCR sequence database is preferred for comparison, which comprises a set of TCR sequences that are representative of TCR of antigen-specific T cells. Especially preferred is a TCR sequence data base that is generated by the steps of

- a) stimulating lymphocytes isolated from a patient in vitro with one or several antigens;
- b) selecting T cells responding to stimulation from said lymphocytes;
- c) obtaining a plurality of nucleic acid sequences encoding T cell receptor proteins from the selected T cells.

The number corresponding to the plurality of nucleic acid molecules sequenced in step c) may vary according to type of immune response or T cell population studied. It must be sufficiently large to allow for a statistically relevant attribution of distribution frequencies to several dozen sequences in the sub-percent range, preferably in the sub-per-mill range. Hence, a plurality of sequences may signify one thousand, preferably a number significantly greater than one thousand, for example ten thousand, twenty-five thousand or fifty thousand.

Current methods of high throughput sequencing allow for the parallel sequencing of thousands of short sequences; such high numbers are

certainly within the range of numbers intended. According to one preferred embodiment of the invention, at least one thousand, even more preferred more than ten thousand sample TCR sequences are obtained from one patient. The sample TCR sequences may be aligned to known TCR sequences in order to assign them to a known TCR V- or J-family, for example.

When applied in the context of immune response monitoring, the method according to the present invention provides significant advantages over methods known in the art, by facilitating an analysis of the clonal diversity of the immune response. This may be applied to monitor and manage immune responses that are desired, from a medical point of view, or responses that are to be avoided. One example of an application to monitor a desired immune response is vaccination. In vaccination, immunization regimes are preferred that elicit a deep and broad response, as defined by the number and variety, respectively, of T cells specific for the antigen used for vaccination. The breadth of T cell responses has proven predictive for the amount of protection achieved in field trials of HIV, and the principle is applicable to all other pathogens where T cell responses are required for mounting a protective immune response, such as plasmodium, mycobacterium or leishmania. By applying the method of the present invention to analysis of the T cell repertoire specific for the antigen, a clonal analysis of the diversity of the T cell response is facilitated. Moreover, by applying the present invention repeatedly over time, information about the temporal evolution of the T cell mediated immune response is facilitated at a molecular level.

Similarly, the method of the present invention greatly improves the monitoring of immune functions that are detrimental for the patient. A prominent example is the monitoring of allergen-specific responses during allergen immunotherapy (induction of tolerance). Another application is the monitoring of transplant recipients (graft rejection or acceptance/tolerance).

According to a second aspect of the invention, a method to obtain antigen-specific T cells for adoptive T cell transfer is provided, whereby lymphocytes isolated from a patient are stimulated in vitro with one or several antigens, the antigens preferably being present at low concentration during stimulation.

According to this aspect of the invention, T cells responding to stimulation by antigen(s) are selected, and a plurality of nucleic acid sequences are obtained from a sample or at least a subsample of the selected T cells. Subsequently, nucleic acid molecule sequences
5 obtained in the previous step are selected according to their frequency among the plurality of sequences obtained in that step, giving rise to a set of preferred sequences. T cells comprising specific preferred T cell receptor molecule sequences are obtained by culturing individual clones originating from a sample of the
10 cells selected according to their stimulation, analysis of the T cell receptor protein sequence of the clone, and selecting a clone comprising a desired T cell receptor protein sequence, or by transferring an nucleic acid sequence encoding the desired T cell protein sequence under control of an RNA polymerase promoter
15 operable in a mammalian cell, into a T cell.

A low concentration of an antigen in this context is a concentration where only a fraction of T cells naturally capable of responding to the antigen respond. Preferred as "low concentration" is a concentration of antigen at which only 10-25% of the numbers of
20 cells respond to the antigen concentration compared to a "normal" concentration at which the entire population of cells that may be activated do respond. If, by way of example, T cell stimulation is effected in cell culture at a "normal" concentration of 1 µg/µL, a "low" concentration would be 0.1 µg/µL or lower. If low
25 concentrations of target protein peptides are employed for stimulation, only TC with highly affine T-cell receptors will react.

A preferred mode of practicing the invention employs obtaining at least one thousand, more preferred ten thousand or thirtythousand nucleic acid sequences.

30 The person skilled in the art will be able to chose from a large number of technologies for transferring nucleic acids encoding desired protein sequences into T cells. These include viral gene transfer, for example by vectors derived from the adenovirus, adeno-associated virus, lentivirus or herpes virus families, and insertion
35 of DNA expression constructs into cells.

This aspect of the invention greatly improves over methods of generation of antigen-specific T cells for adoptive immunotherapy,

such as published by Hammer (Eur. J. Immunol. 2005, 35:2250-2258). One important aspect of improvement is the generation of sequence data on TCR, leading to the possibility of transferring a representative set of dominant T cell clones adapted both to the nature of antigen and the specific biological background of the patient.

According to one preferred embodiment of any aspect of the invention, the lymphocytes are isolated from blood. In a more preferred embodiment, the lymphocytes are peripheral blood mononuclear cells (PBMC), especially T lymphocytes, or lymphocytes obtained by tissue biopsy.

According to another preferred embodiment of the invention, the isolated lymphocytes are regulatory T cells. In an alternative preferred embodiment, the isolated lymphocytes are effector T cells. One may also practice the invention using a combination of both effector and regulatory T cells.

According to a further preferred embodiment of all aspects, lymphocytes are stimulated with a peptide library representing antigenic peptides, or whole protein antigen, e.g. overlapping peptide pools spanning the whole CMV-specific protein pp 65, or lysates of cytomegalovirus (CMV-), BK-virus (BKV-), or Epstein-Barr-Virus (EBV-) specific proteins.

According to yet another preferred embodiment, lymphocytes are selected according to their expression of stimulation markers. According to a more preferred embodiment, selection is achieved through sorting by fluorescence activated cell sorting (FACS) or magnetic cell sorting (MACS) according to the cells' expression of activation markers on the cell surface or within the cells. Examples for such activation markers are IFN γ , IL-2, TNF α , CD69, and CD40L.

The nucleic acid molecules encoding T cell receptor protein sequences that are analyzed according to the invention, encode regions of the TCR that contribute to TCR specificity; they preferably encode a peptide chain contributing to the diversity of CDR3 variants that generate the vast number of possible TCR structures. According to another preferred embodiment, the nucleic

acid molecules analyzed encoding T cell receptor protein sequences encode all or part of the CDR3 region of the TCR beta chain.

The examples of this specification illustrate a method for the analysis of TCR-CDR3 repertoire in CMV-specific T cells. A marked
5 clonal restriction was observed among the effector memory T cells generated from individual donors using a CMV pp 65 peptide library. Sequencing data on more than 100 000 different CMV-specific TCR sequences for each donor attest to oligoclonality in CMV-specific TCR repertoire, while a very high degree of inter-individual clonal
10 diversity exists within the CMV T cell recognition spectrum.

One embodiment of the method according to the invention similarly allows for the tracing of the clonal composition of viral antigen-specific T cells in the clinical course of patients with or without virus re-activation, as well as in patients receiving adoptive
15 virus-specific T cell therapy. A preferred embodiment is directed towards monitoring the CMV infection and re-activation. Analysis of CMV-specific cytotoxic T cell (CTL) responses is essential for the prevention and therapy of CMV re-activation.

Another embodiment of the present invention allows for the
20 monitoring of naturally occurring or adoptively transferred antigen-specific T cells.

The utility of the information obtained by the present method resides in its qualitative and quantitative accuracy as regards the T cell receptor repertoire of the immune response within a specific
25 individual, unprecedented by methods of the state of the art. The interpretation of this information will be left to the physician, as well as possible conclusions with view towards treatment of the patient from whom this information was obtained.

Short description of the figures

30 Fig. 1 shows an example of a method to obtain the sequence data base for CMV pp65 specific TCR sequences.

Fig. 2 shows the results of an assay determining the sensitivity of clonal quantification over varying dilutions.

Fig. 3 shows frequency data in percent for dominant and subdominant clones obtained from 4 different preparation/sequencing runs from one donor.

Fig. 4 shows the Morisita Horn similarity index of the whole TCR repertoire obtained by 2 independent analyses from one sample.

Fig. 5 shows tracking of AG-specific T cells by analysis of the CD3+ fraction from healthy donors.

Fig. 6 shows the frequency (percent, Y-axis) of dominant and subdominant clones obtained from one donor at 0 (x-Axis labeled "1") and 3 years (x-axis labeled "2").

Examples

Material and methods

Peripheral blood lymphocytes were isolated from a CMV-positive healthy donor. RNA was isolated as described previously (Babel et al, J. Am. Soc. Nephrol. 20; 344-352, 2009). The hypervariable regions of TCR V β chain was amplified by a 2-tube multiplex PCR with a primer mix that covers all V β TCR gene rearrangements on genomic DNA or cDNA samples as described previously (Babel et al, *ibid.*).

RNA was isolated from separated cells by RNA Miniprep Kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions. RNA was quantified by the Agilent 2100 Bioanalyzer System (Agilent Technologies, Santa Clara, CA) and used for cDNA synthesis. DNA was isolated using a commercially available DNA extraction kit (QIAamp Blood Kit, Qiagen, Hilden, Germany) according to the manufacturer's instructions. PCR products were separated on a 1.7 % agarose gel, excised, and purified using the Gel Extraction Kit (Eppendorf, Hamburg, Germany) following the manufacturer's instructions.

The purified product of multiplex PCR underwent nucleotide sequence analysis by high-throughput sequencing.

Example 1:

CDR3 junction sequences were obtained from 100.000 sequences. The frequency of the most prevalent sequences is given in table A below.

Table A: CMV-specific sequences, 100.000 runs

SEQUENCE NAME	CDR3 JUNCTION (AA)	Length (nt)	# of reads	%
FXTBR9C02G4SL	C A S S P Q T G V G N Y G Y T F	48	6648	22
FXTBR9C02GWS	C A S S L L G Q G V N Q P Q H F	51	1874	6
FXTBR9C02GXBS	C S V S T G G W D Q P Q H F	42	517	1.7
FXTBR9C02G7SDM	C A S S F L T T T D T Q Y F	42	508	1.69
FXTBR9C02JEMV8	C A S S L T S G A T Y N E Q F F	48	348	1.16
FXTBR9C02IW5QK	C A S S S G Q G Q V T Q Y F	42	301	1.0

In a second experiment, CMV-specific sequences were obtained from a second individual in 30.000 runs (table B):

5 Table B: CMV-specific sequences, 30.000 runs

SEQUENCE NAME	CDR3 JUNCTION (AA)	Length (nt)	# of reads	%
FXTBR9C02G4SL	C A S T W T G P D Q P Q H F	42	12918	13
FXTBR9C02I90TN	C S V E V G V E G Y T F	36	4952	5
FXTBR9C02H7EFH	C A S S R G A G N T E A F F	42	4524	4.5
FXTBR9C02I0ONJ	C A S S A G L L Q E T Q Y F	42	1570	1.57
FXTBR9C02JDN7O	C A S S V Q G Y T E A F F	39	1225	1.23
FXTBR9C02I2LII	C A S S Q E V Q A P Y G Y T F	45	803	0.8

Example 2: Sensitivity of the method

In order to define sensitivity of the method allowing identification of the clones of interest presented at low frequency, pp65-specific gDNA was diluted at different ratios in a third party non-specific

gDNA from a CMV-negative donor (a CD3+ fraction of an irrelevant donor was used in order to introduce large amount of recombined CDR3 regions of a irrelevant specificity). In parallel to undiluted samples, several dilution series were set up with the highest one of 1:25.000. Figure 2 demonstrates clonal distribution of analyzed samples at different dilutions. The applied method is highly sensitive: Ag-specific dominant and subdominant clones persisted relative in all serial dilutions up to 1: 25.000. A high correlation was also observed between the frequency of single dominant and their dilutions ($r = 0.9992 - 0.999989$; dependent on the analyzed clone; $p < 0.001$)

Example 3. Stability/variability of the method

Figure 3 represents frequency data of dominant and subdominant clones obtained from 4 different preparation/sequencing runs from one donor. The method appears to be robust with a very low coefficient of variability for inter-assay variability.

Figure 4 represents a similarity index of the whole TCR repertoire (the x axis represents the number of different clones which frequencies were compared to each other) obtained by 2 independent analyses from one sample. The high Morisita-Horn similarity index (about 0.94) remains stable even if the whole repertoire (all obtained TCR sequences from the whole sequencing run) was included in the analysis.

Example 4: Tracking of antigen-specific cells in peripheral blood

In order to determine whether tracking of AG-specific T cells is possible by analysis of their TCR sequences in whole blood without previous antigen stimulation and sorting steps, CD3+ fraction from healthy donors with known clonal distribution of pp65+ CTL were analyzed. Figure 5 demonstrates that all dominant and subdominant Ag-specific clones were detected in peripheral blood without antigen stimulation and sorting steps.

Example 5. Dynamic of TCR repertoire

Figure 6 represents the frequency of dominant and subdominant clones obtained from one donor at different time points (0, and 3 years). As presented in the figure 6 the TCR repertoire remains also stable in follow up

Claims

1. A method for analysis of a patient's T cell repertoire, whereby

a) nucleic acid molecules encoding T cell receptor protein sequences are obtained from lymphocytes isolated from a patient;

b) at least one thousand of said nucleic acid molecules obtained in step a) are sequenced, rendering sample TCR sequences;

c) the sample TCR sequences obtained in step b) are aligned with a TCR sequence data base, and sample TCR sequences positively aligning to data base sequences are selected as selected sample TCR sequences,

d) the frequency distribution of at least some of the selected sample TCR sequences obtained in step c) is determined

whereby the TCR sequence data base either

i) is obtained by stimulating lymphocytes isolated from a patient in vitro with one or several antigens; selecting T cells responding to stimulation from said lymphocytes; and obtaining a plurality of nucleic acid sequences encoding T cell receptor proteins from at least some of the T cells selected in step i)), or

ii) whereby the TCR sequence data base comprises TCR sequences specific for

- a T cell derived tumour having been obtained from said patient previously, or
- an adoptively transferred T cell previously injected into said patient.

2. A method according to claim 1, characterized in that the TCR sequence data base comprises TCR sequences of TCR molecules specific for antigens from any of the pathogens of the group including Epstein-Barr virus, Cytomegalovirus, plasmodium falciparum, mycobacterium tuberculosis, and/or Human Immunodeficiency Virus.

3. A method according to at least one of the above claims,
characterized in that at least ten thousand sample TCR
sequences are obtained.
4. A method according to at least one of the above claims,
5 characterized in that the TCR sequence data base comprises TCR
sequences of TCR molecules specific for allergens or autoimmune
antigens.
5. A method to obtain antigen-specific T cells for adoptive T cell
transfer, whereby
 - 10 a) lymphocytes isolated from a patient are stimulated in
vitro with one or several antigens and T cells responding
to stimulation are selected from said lymphocytes;
 - b) a plurality of nucleic acid molecule sequences encoding T
cell receptor protein sequences are obtained from the T
15 cells selected in step a);
 - c) nucleic acid molecule sequences obtained in step b) are
selected according to their frequency among the plurality
of sequences obtained in step b), rendering preferred
sequences;
 - 20 d) T cells comprising preferred sequences are obtained by
 - culturing individual clones originating from a sample
of the cells selected in step a), analysis of the
nucleic acid molecule sequences encoding T cell
receptor protein sequences of the clone, and
25 selecting a clone comprising a preferred sequence, or
 - transferring an operable nucleic acid sequence
encoding a T cell protein sequence comprising a
preferred sequence into a T cell.
6. A method according to claim 5, characterized in that the
30 lymphocytes isolated from a patient are stimulated in vitro
with one or several antigens at an antigen concentration equal
or less than 25% of the antigen concentration employed to
optimally stimulate lymphocytes.

7. A method according to at least one of the above claims,
characterized in that the lymphocytes isolated from a patient
are PBMC or lymphocytes obtained by tissue biopsy.
- 5 8. A method according to at least one of the above claims,
characterized in that lymphocytes are stimulated by exposition
to a peptide library representing antigenic peptides, or whole
protein antigen.
- 10 9. Method according to at least one of the above claims,
characterized in that lymphocytes are selected by fluorescence
activated cell sorting (FACS) or magnetic cell sorting (MACS)
according to the cells' expression of activation markers and/or
cytokines on the cell surface or within the cells.
- 15 10. Method according to at least one of the above claims,
characterized in that the nucleic acid molecules encoding T
cell receptor protein sequences encode all or part of the CDR3
region of the TCR beta chain.
11. Cells obtained by a method according to at least one of claims
4 to 10.
12. Use of cells according to claim 11 as a pharmaceutical.
- 20 13. Use of cells according to claim 11 as a pharmaceutical in
cancer therapy, treatment of transplant rejection, graft-
versus-host-disease, viral infection, or vaccination.

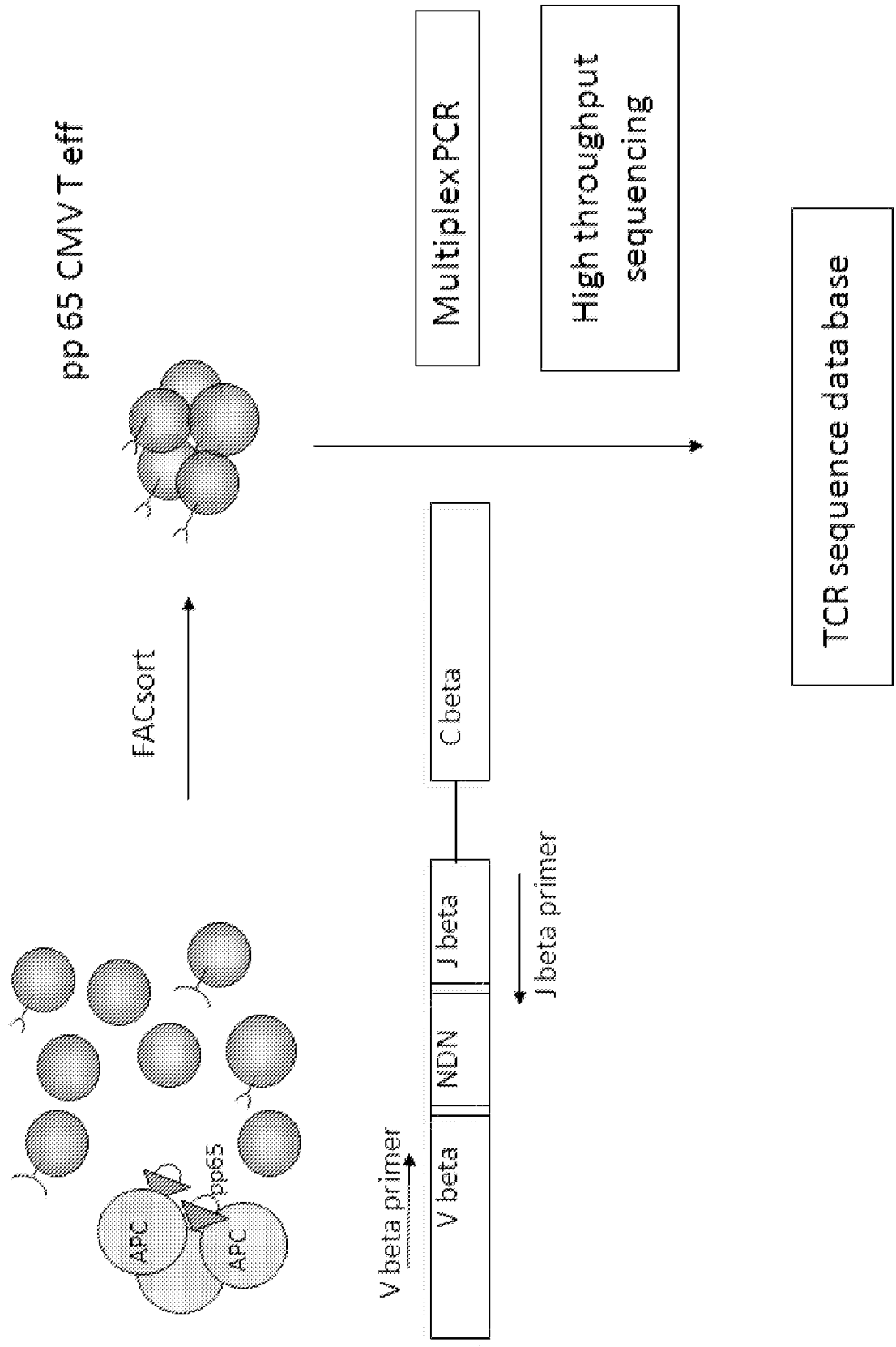


Fig. 1

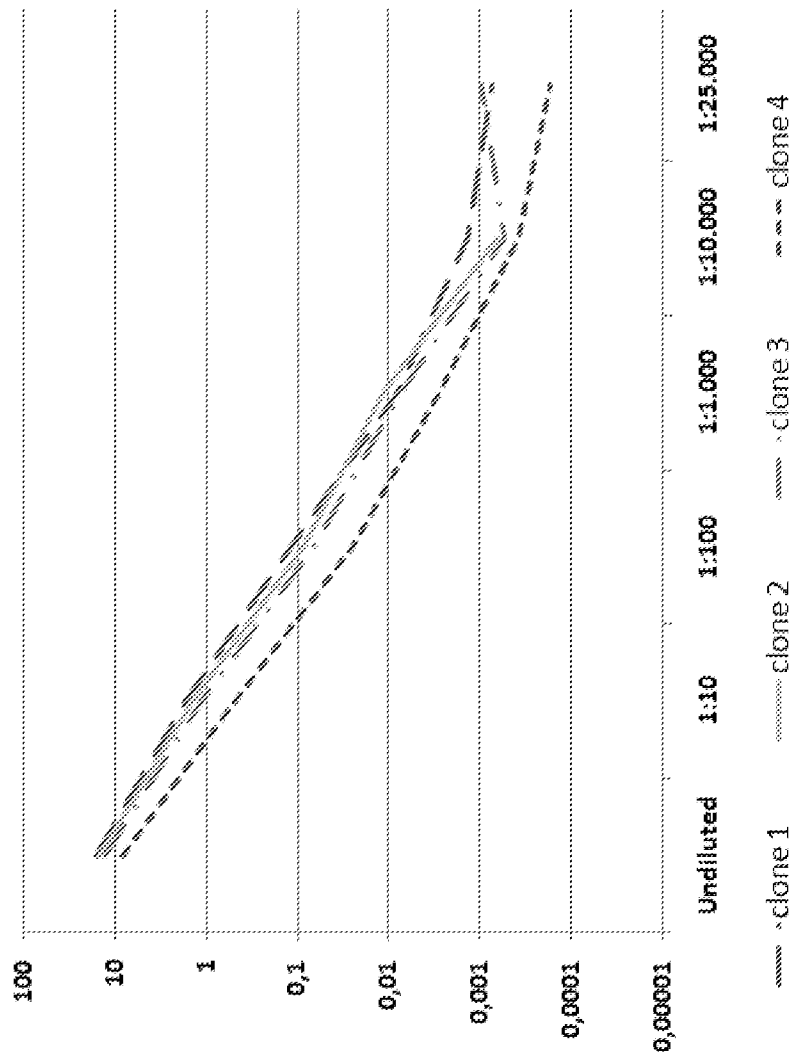


Fig. 2

	Sample#1	Sample#2	Sample#3	Sample#4	Mean	SD	CV in %	Median CV in %
V6-3,V6-2:J2-7	2,69	3,21	2,67	2,56	2,78	0,21	7,6	8,825
V6-1:J2-5	1,66	0,97	1,26	1,43	1,33	0,22	16,2	
V12-4,V12-3:J1-6	1,39	1,84	1,26	1,18	1,42	0,21	14,8	
V6-1:J2-5	1,21	0,80	0,95	1,10	1,01	0,14	13,8	
V25-1:J2-3	1,20	1,12	1,17	1,12	1,16	0,03	3,0	
V13:J2-3	1,13	0,95	1,50	1,48	1,27	0,23	17,8	
V6-5:J1-1	0,99	0,84	0,85	0,94	0,90	0,06	6,7	
V18:J1-3	0,74	0,83	1,26	1,12	0,99	0,20	20,4	
V27:J2-7	0,64	0,53	0,61	0,56	0,59	0,04	6,6	
V12-4,V12-3:J2-7	0,38	0,47	0,36	0,36	0,39	0,04	10,0	
V6-5:J1-6	0,36	0,44	0,26	0,26	0,33	0,07	21,3	
V12-4,V12-3:J1-5	0,35	0,32	0,37	0,34	0,34	0,02	4,5	
V6-3,V6-2:J2-7	0,32	0,37	0,30	0,27	0,31	0,03	9,1	
V29-1:J2-1	0,30	0,26	0,42	0,41	0,35	0,07	19,0	
V6-3,V6-2:J2-5	0,29	0,26	0,29	0,26	0,28	0,02	5,5	
V10-3:J1-1	0,25	0,19	0,25	0,24	0,23	0,02	8,4	
V6-3,V6-2:J2-5	0,24	0,27	0,25	0,21	0,24	0,02	6,6	
V20-1:J1-5	0,23	0,29	0,29	0,26	0,26	0,02	8,5	

Fig. 3

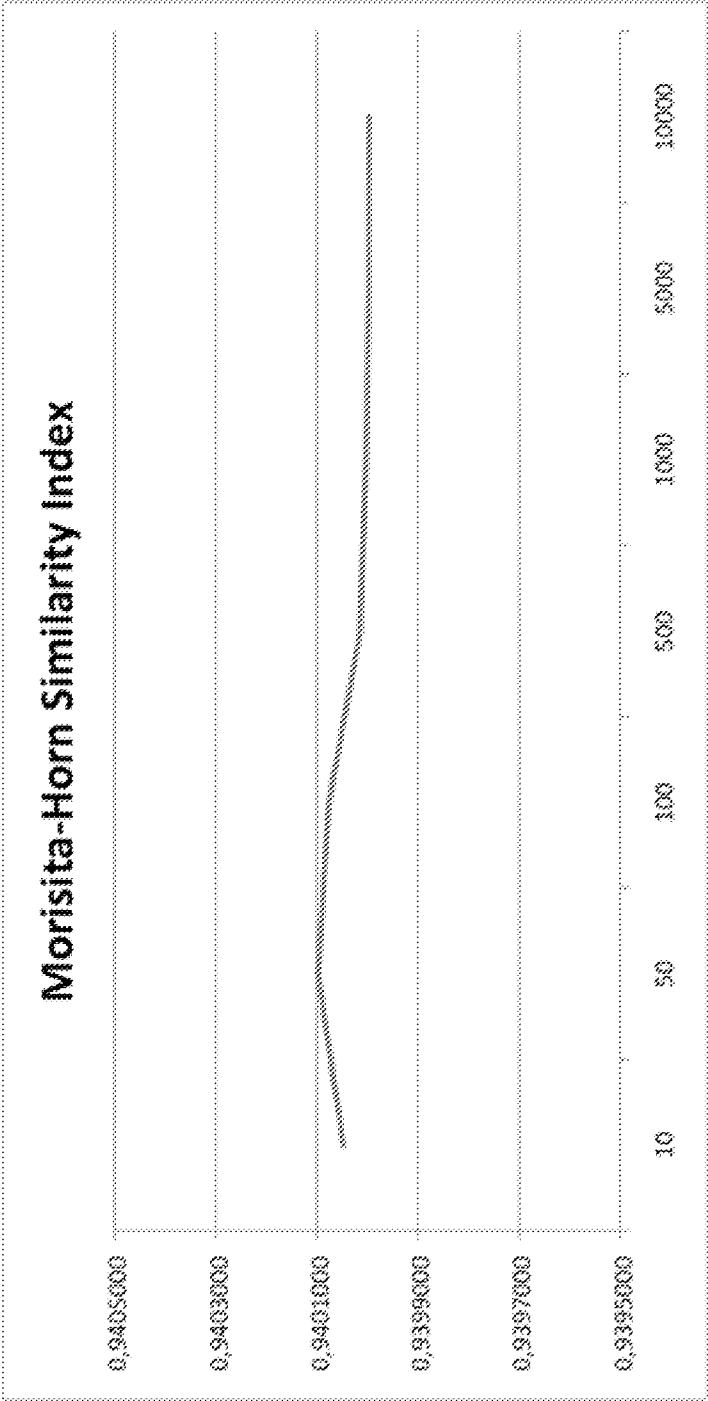


Fig. 4

Tracking of CMV-specific T-cells in blood

	pp65+ Cell Line	CD3+ gDNA in %
V6-3,V6-2:J1-1	13,5%	0,04%
V27:J1-1	12,5%	0,05%
V27:J1-1	10,9%	0,11%
V4-3:J1-1	7,1%	0,04%
V12-4,V12-3:J1-2	4,9%	0,06%
V6-3,V6-2:J1-1	4,0%	0,06%
V6-5:J1-2	3,3%	0,01%
V20-1:J2-7	3,3%	0,06%
V27:J1-5	2,2%	0,004%
V12-4,V12-3:J1-2	1,8%	0,03%
V12-4,V12-3:J2-7	1,5%	0,02%

Fig. 5

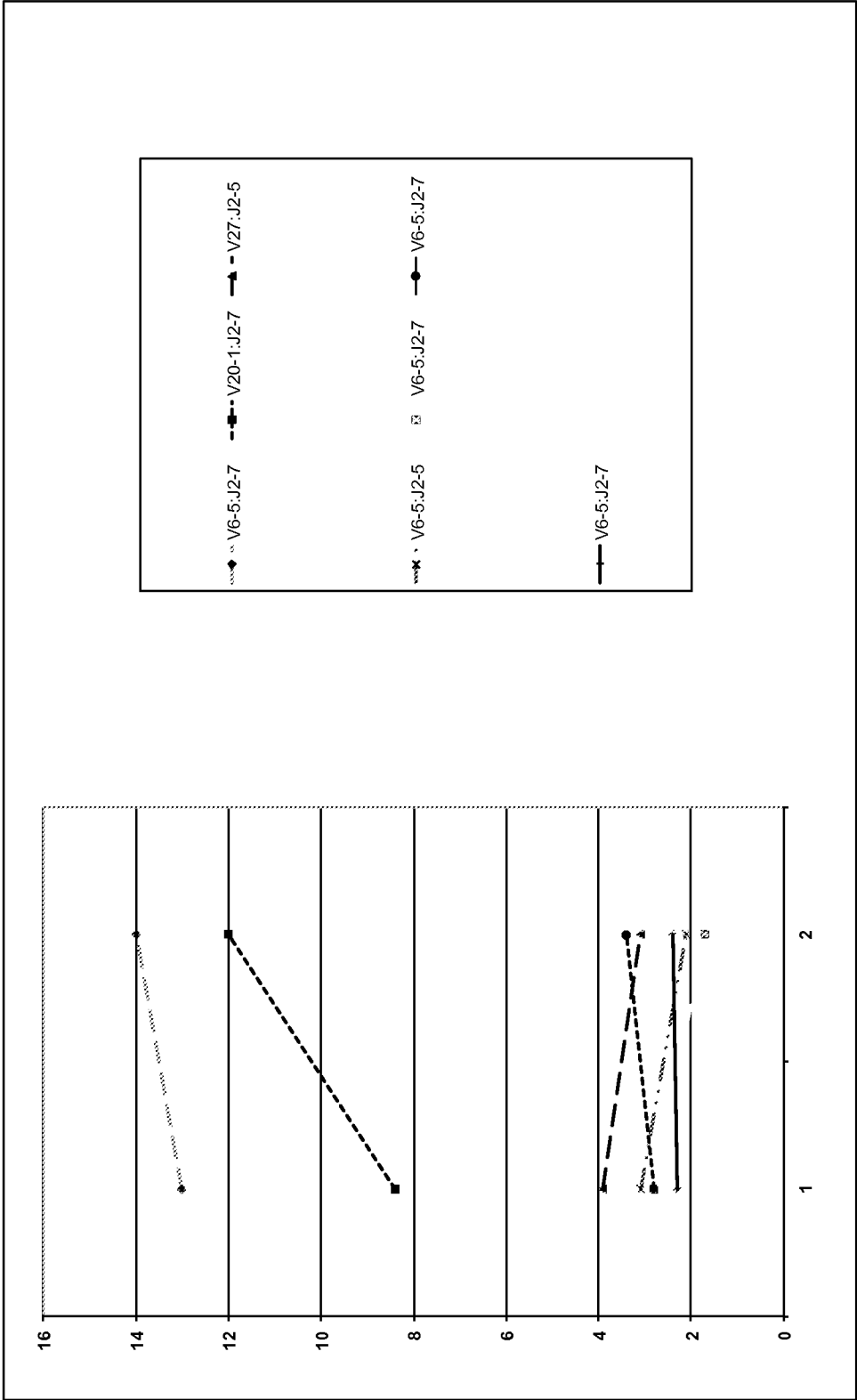


Fig. 6

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/053314

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01N33/50
ADD.

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

B. FIELDS SEARCHED

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

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

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

EPO-Internal, BIOSIS, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Marcin Wlodarski: "Analyse der T-Zell Klonitäten in T-LGL Leukämie und Anderen Erkrnakungen mit erworbenem Knochenmarkversagen", 27 March 2009 (2009-03-27), Medizinische Fakultät Charité- Universitätsmedezin Berlin, Berlin, XP002587604, pages 1-23, page 8 - page 13; figure 3 ----- -/--	1-13



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

11 April 2011

Date of mailing of the international search report

27/04/2011

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Gonçalves Mauger, M

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/053314

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LUE C ET AL: "An automated method for the analysis of T-cell receptor repertoires. Rapid RT-PCR fragment length analysis of the T-cell receptor beta chain complementarity-determining region 3", AMERICAN JOURNAL OF CLINICAL PATHOLOGY, AMERICAN SOCIETY FOR CLINICAL PATHOLOGY, US, vol. 111, no. 5, 1 May 1999 (1999-05-01), pages 683-690, XP009134895, ISSN: 0002-9173 page 683 - page 686 -----	1-13
X	O'KEEFE C L ET AL: "Molecular TCR diagnostics can be used to identify shared clonotypes after allogeneic hematopoietic stem cell transplantation", EXPERIMENTAL HEMATOLOGY, ELSEVIER INC, US LNKD- DOI:10.1016/J.EXPHEM.2004.07.013, vol. 32, no. 10, 1 October 2004 (2004-10-01), pages 1010-1022, XP004613169, ISSN: 0301-472X page 1011 - page 1018 -----	1-13
X	BABEL NINA ET AL: "Clonotype Analysis of Cytomegalovirus-Specific Cytotoxic T Lymphocytes", JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY, vol. 20, no. 2, February 2009 (2009-02), pages 344-352, XP002587605, ISSN: 1046-6673 cited in the application page 346 - page 351; figures 2,3,4 -----	1-13
X	HAMMER MARKUS H ET AL: "HLA type-independent generation of antigen-specific T cells for adoptive immunotherapy", EUROPEAN JOURNAL OF IMMUNOLOGY, vol. 35, no. 7, July 2005 (2005-07), pages 2250-2258, XP002587606, ISSN: 0014-2980 page 2250 - page 2252; figure 1 ----- -/--	5-13

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/053314

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WLODARSKI ET AL: "Clonal predominance of CD8<+> T cells in patients with unexplained neutropenia", EXPERIMENTAL HEMATOLOGY, ELSEVIER INC, US LNKD- DOI:10.1016/J.EXPHEM.2007.11.011, vol. 36, no. 3, 13 February 2008 (2008-02-13), pages 293-300, XP022480050, ISSN: 0301-472X page 294 - page 296 -----	1-13
X	HUMME DANIEL ET AL: "Dominance of Nonmalignant T-Cell Clones and Distortion of the TCR Repertoire in the Peripheral Blood of Patients with Cutaneous CD30(+) Lymphoproliferative Disorders", JOURNAL OF INVESTIGATIVE DERMATOLOGY, vol. 129, no. 1, January 2009 (2009-01), pages 89-98, XP002587607, ISSN: 0022-202X page 96 - page 98 -----	1-13
X	BREWER ET AL: "An improved methodology to detect human T cell receptor beta variable family gene expression patterns", JOURNAL OF IMMUNOLOGICAL METHODS, ELSEVIER SCIENCE PUBLISHERS B.V.,AMSTERDAM, NL, vol. 302, no. 1-2, 1 July 2005 (2005-07-01), pages 54-67, XP005008041, ISSN: 0022-1759 page 56 - page 57 -----	1-13