

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

(19) World Intellectual Property Organization

International Bureau



(10) International Publication Number

WO 2017/013103 A1

(43) International Publication Date  
26 January 2017 (26.01.2017)

(51) International Patent Classification:  
*B01L 3/00* (2006.01)

(74) Agent: LC PATENTS; Kempische Steenweg 542A, 3500 Hasselt (BE).

(21) International Application Number:  
PCT/EP2016/067149

(22) International Filing Date:  
19 July 2016 (19.07.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
15178158.0 23 July 2015 (23.07.2015) EP

(71) Applicant: BIOCARTIS NV [BE/BE]; Generaal De Wittelaan 11 B3, 2800 Mechelen (BE).

(72) Inventors: VERGAUWE, Nicolas; Biocartis NV, Generaal De Wittelaan 11 B3, 2800 Mechelen (BE). MEERSSEMAN, Geert; Biocartis NV, Generaal De Wittelaan 11 B3, 2800 Mechelen (BE).

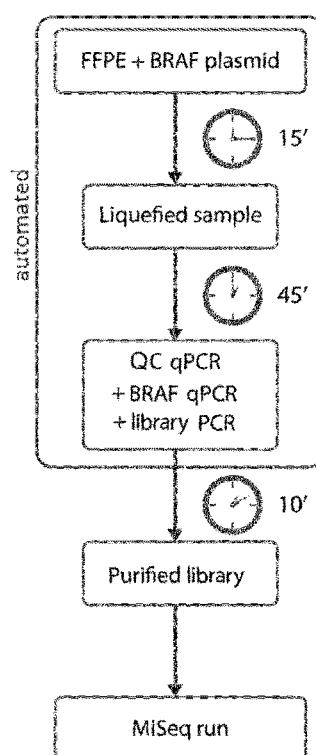
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP,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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,

[Continued on next page]

(54) Title: AUTOMATED SAMPLE TO NGS LIBRARY PREPARATION

Figure 9



(57) Abstract: The present invention generally relates to cartridges for automated systems and automated methods for fully automated processing of biological samples to next-generation sequencing-ready nucleic acid libraries. In particular, the present invention concerns a fluidic cartridge comprising a compartment for receiving a biological sample, means for liberating and/or purifying nucleic acids from the received sample and for transporting said nucleic acids to a compartment wherein NGS library can be prepared from said nucleic acids using provided therein reagents. In a particular aspect, the present invention also concerns cartridges for automated systems and automated methods for preparing an NGS nucleic acid library from a biological sample simultaneously or sequentially with a qPCR assay performed on the same biological sample.



SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,  
GW, KM, ML, MR, NE, SN, TD, TG).

— *of inventorship (Rule 4.17(iv))*

**Declarations under Rule 4.17:**

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))* — *with international search report (Art. 21(3))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))* — *with sequence listing part of description (Rule 5.2(a))*

**Published:**

## AUTOMATED SAMPLE TO NGS LIBRARY PREPARATION

### TECHNICAL FIELD

[0001] The present invention generally relates to cartridges for automated systems and automated methods for fully automated processing of biological samples to next-generation sequencing-ready nucleic acid libraries. In particular, the present invention concerns a fluidic cartridge comprising a compartment for receiving a biological sample, means for liberating and/or purifying nucleic acids from the received sample and for transporting said nucleic acids to a compartment wherein NGS library can be prepared from said nucleic acids using provided therein reagents. In a particular aspect, the present invention also concerns cartridges for automated systems and automated methods for preparing an NGS nucleic acid library from a biological sample simultaneously or sequentially with a qPCR assay performed on the same biological sample.

### BACKGROUND OF THE INVENTION

[0002] The ever-decreasing costs of high-throughput sequencing, also known as second generation- or next generation-sequencing (NGS) in contrast to much slower classical Sanger sequencing, are gradually bringing this powerful high-coverage diagnostic technique into clinical practice. Currently, there exist many different NGS strategies and platforms, including e.g. the commercially known Illumina (Solexa) sequencing, Roche 454 sequencing, Ion torrent (Proton / PGM) sequencing, or SOLiD sequencing. All of them differ in details but in general involve generation of many short overlapping nucleic acid fragments or “reads”, often modified to include sequencing platform-specific adapter sequences. A collection of such fragments prepared from a specific source of a nucleic acid is called a NGS nucleic acid library. A process of preparing such nucleic acid library from a clinical sample is time-consuming and requires multiple preparatory steps. Namely, not only nucleic acids retrieved from a sample have to be sufficiently pure and recovered in a sufficient amount but also, in order to provide satisfactory genetic information coverage, they have to meet often stringent quality criteria and be free of contaminations that could impede either the library preparation step or the sequencing reaction. As a result, NGS library preparation is not only still labor-intensive on its own, but it also requires additional steps of assessing the quantity and quality of either the purified nucleic acid or constructed NGS library prior to sequencing step in order to ensure high-quality results and their correct interpretation. Thus, even despite the tremendous improvements made to make NGS even faster, cheaper, and requiring less manpower, the technique is still relatively expensive and much slower as compared to diagnostic screening via e.g. target specific quantitative or real-time PCR (qPCR).

[0003] To date, processing of a clinical sample taken from a patient to an NGS-ready is largely done manually and at least includes the steps of:

1. cell lysis or liquefaction;

2. liberation or purification of the nucleic acids;
3. determination of the quality and quantity of the nucleic acids by any known means in the art, for example absorbance (optical density) measurement, agarose gel electrophoresis, fluorometry with a fluorescent DNA-binding dyes like picogreen, or different types of quality control PCRs (cf. Buehler et al. 2010);
4. preparation of an NGS library according to a strategy depending on the NGS platform of choice; and
5. often but optionally, determination of the quality and quantity of the DNA library by any of the means known in the art.

10 [0004] Some NGS providers like Illumina make their best efforts to simplify the above-explained workflow, example of which is the introduced in January 2015 by Illumina NeoPrep that automates the library preparation step (step 4) from a manually isolated and quality-verified DNA. However, to date there is no protocol for direct, fast, and automated sample-to-NGS-ready library processing, and especially not such that provides for concomitant nucleic acid quality and quantity assessment.

15 [0005] The present invention addresses this and other needs by providing a fluidic cartridge for automated processing of a biological sample of practically any type (including blood, stool, urine, FFPE, swab, etc.) to an NG sequencing ready library. The cartridge of the present invention comprises all the means necessary for automated sample liquefaction, followed by nucleic acid liberation and/or purification, and then preparation of an NGS-ready library from at least a part of the liberated and/or 20 purified nucleic acid without any manual human intervention from the moment the sample is introduced into the cartridge. Importantly, the cartridge of the invention also provides all the means necessary for automated qPCR analysis, which can be performed in parallel with or sequentially with respect to the NGS library preparation from another portion of the same liberated and/or purified nucleic acid of from the NGS library itself. Such qPCR can be a quality control (QC) qPCR for 25 assessing the quality of nucleic acids purified from the sample provided into the cartridge prior to library preparation, or for assessing the quality of library after it has been prepared and before subjecting it to NGS analysis. Alternatively, such qPCR can be an assay qPCR for screening target actionable mutations in a clinical sample.

[0006] By automatically and concomitantly generating (a) an NGS-ready library, together with (b) its 30 quality evaluation, and performing (c) target qPCR screening from one and identically-processed biological sample, one would not only save time and trained human resources but also would ensure better management of precious and limited in amount clinical samples. A further advantage of such solution would be the ability to immediately use the information obtained from a real-time-monitored QC qPCR reaction for an indication of nucleic acid quantity and quality, and thus its suitability for NGS 35 library generation. A further advantage would be the ability to directly compare the assay qPCR results and the results obtained from an NGS analysis of the library concomitantly generated with and on the same machine as said qPCR, either at the same time of (i.e. simultaneously) or just shortly after (i.e.

sequentially), and, importantly, by using the same pool of identically processed nucleic acid. Furthermore, by thus removing the usual temporal separation between a qPCR analysis and stabilization of the free nucleic acids in a form of a library, one could further minimize any potential analytical dichotomies between qPCR and NGS results, which stem from prolonged storage and 5 potential degradation of biological samples or free nucleic acids.

[0007] Thus, the cartridges and methods of the present invention not only allow automated sample-to-NGS-library-generation together with potential verification of the sample's or library's quality, but further provide means for immediate high-depth discovery of driver mutations in advance of NGS-in-depth genomic analysis. This and other advantages of the present invention are presented in 10 continuation.

### **SUMMARY OF THE INVENTION**

[0008] The present invention is defined in the appended independent claims. Preferred embodiments are defined in the dependent claims. In particular, the present invention concerns a cartridge for 15 automated processing of a biological sample, the cartridge comprising

- a sample compartment for receiving a biological sample;
- means for liberating or purifying nucleic acid from the biological sample received in the sample compartment, said means capable of entering in fluid communication with said sample compartment;
- 20 - a library compartment for preparing a nucleic acid library positioned downstream of the sample compartment and of the means for liberating or purifying nucleic acid, the library compartment configured to receive at least a portion of the liberated or purified nucleic acid and comprising reagents for preparing a nucleic acid library.

[0009] In a preferred embodiment, the above-described cartridge is characterized in that that the 25 library compartment is adapted for:

- - generation of nucleic acid fragments from the portion of the liberated or purified nucleic acid received into the library compartment, and for
- - attaching oligonucleotide adapters to at least one, preferably both ends of said generated nucleic acid fragments.

30 [0010] In a further aspect, the present invention also provides an automated method for preparing a next-generation sequencing (NGS) nucleic acid library from a biological sample, said method comprising the steps of:

- a) receiving a biological sample;
- b) liberating or purifying at least a part of nucleic acids from the received biological sample;
- 35 c) providing at least a portion of the liberated or purified nucleic acid into a library compartment for preparing an NGS nucleic acid library and comprising reagents for preparing a nucleic acid

library; and

d) preparing an NGS nucleic acid library, said preparing not requiring any human intervention after the receiving of the biological sample;  
wherein at least the steps b) to d) are performed on an automated system.

5 [0011] Finally, the invention also provides uses of cartridges according to the invention for automated preparation of a next-generation sequencing (NGS) library from a biological sample.

## DEFINITIONS

10 [0012] The term "*quantitative PCR*" or simply "*qPCR*" is herein given the definition of a laboratory technique based on the polymerase chain reaction (PCR), which is used to amplify and simultaneously detect or quantify a targeted DNA molecule. In contrast to standard PCR where the product of the reaction is detected at its end, i.e. after thermocycling has finished, the key feature of qPCR is that the DNA product is being detected during thermocycling as the reaction progresses in "real time"; hence, the alternative name of qPCR "*real-time PCR*". There currently exist many different types of qPCRs. For example, when starting with a reverse transcription (RT) step, qPCR can be used to quantify numbers of messenger RNAs and is then called a reverse transcriptase qPCR or an RT-qPCR. As used herein the terms "*quantitative PCR*" or simply "*qPCR*" will be employed with preference over the term "*real-time PCR*" or "*RT-PCR*" in order to avoid confusion with *reverse transcription PCR*, also frequently abbreviated as RT-PCR. Most qPCRs use one of the two most common methods for detecting the product amplification in real-time: (a) intercalation of non-specific fluorescent dyes with any double-stranded DNA, or (2) sequence-specific DNA probes consisting of oligonucleotides that are labelled with a fluorescent reporter which permits detection only after hybridization of the probe with its complementary target sequence. The fluorescent signals generated during thermocycling are detected by an appropriate optical detection system and tracked from the moment they pass the background threshold till the reaction reaches plateau. The copy number of the target sequences can be estimated using either relative or absolute quantification strategy, typically by analyzing the shape of the obtained amplification curve (standard curve strategy) or by determining when the signal rises above some threshold value (often called the Ct value, but sometimes also Cp value or Cq value). In relative quantification, the target nucleic acid levels estimated in a given sample using the Ct or standard curve analysis are expressed as relative to values obtained for the same target in another reference sample, for example, an untreated control sample. Conversely, in absolute quantification the qPCR signal is related to input copy number using a standard curve or can also be calculated according to a more recent digital PCR method. For the moment being, the first strategy is still more prevalent and bases the estimation of the target DNA amount by comparing the obtained values with a previously made standard curve. These and other qPCR quantification strategies are

broadly known in the art and their calculation can differ in smaller or greater depending on a given application and a qPCR system.

[0013] As used herein, the term "*means for performing quantitative PCR*" shall be understood as minimum necessary arrangement of reagents and elements for performing a qPCR. They will usually

5 include any reagents allowing detectable in real time PCR thermocycling of a nucleic acid template received from a source of nucleic acid. Such reagents include but depending on the type of qPCR are not limited to a PCR-grade polymerase, at least one primer set a detectable dye or a probe, dNTPs, PCR buffer etc. Further, the "*means for performing quantitative PCR*" will usually also include any standard known in the art minimal assembly of parts, which usually includes but is not limited to the  
10 following: (1) a suitable compartment (further referred to as a "*a thermocycling qPCR compartment*") where the real time-detectable thermocycling can take place. Such compartments can e.g. be formed by a chamber suitable for amplifying nucleic acids, i.e. made from appropriate material and providing for sufficient internal temperature regulation, and also comprising at least one wall allowing real-time detection of signals generated during such amplification, e.g. a wall transparent to light. Further, (2)  
15 means for varying temperature in this chamber or other compartment, as broadly known from various existing thermocycling machines. Then, (3) means for detecting the signals generated during the qPCR thermocycling, like an optical detector coupled to a computer etc. In brief, such minimal assembly will normally include any known in the art system or systems capable of initiating and maintaining the thermocycling reaction in the thermocycling qPCR compartment, adjusting and  
20 regulating the temperature to ensure stable thermocycling conditions therein etc.; further, it will also include any appropriate detection device or devices, means for data processing (e.g. a computer alternatively connected to a database), and output systems allowing to read and monitor the thermocycling of the qPCR reaction in real-time (usu. a computer screen displaying the reaction progress in an appropriate graphic user interface); as well as any software packages suitable for  
25 operating the machinery and/or displaying and possibly also aiding the interpretation of the obtained results.

[0014] Further, as used herein, the term "*sequencing library*" or "*sequencing nucleic acid library*" refers to a set of polynucleotides, most often of DNA type, that are ready for sequence analysis, in particular using any of the currently known next-generation sequencing (NGS) strategies. In line with  
30 this, in a currently preferred embodiment of the invention, a sequencing library is composed of a plurality of PCR-amplified DNA molecules, fused with adapter molecules (or adapter sequences) compatible with a given NGS strategy of choice. A comprehensive overview of sequencing library types and ways of preparing them can be found in a review of van Dijk et al. *Exp Cell Res.* 2014.

[0015] Similarly, as used herein, the term "*means for preparing a nucleic acid library*" shall be understood as minimum necessary elements for preparing an NGS-suitable nucleic acid library, which include at least a compartment where a nucleic acid received from a source of nucleic acid can be  
35

subjected to nucleic acid library preparation; the minimum necessary reagents for the library preparation (such as appropriate enzyme or a mix of enzymes, NGS-strategy-specific adapter sequences or adapters, buffer etc.); and also a standard known in the art machinery and software suitable for of e.g. initiating and/or directing the library preparation procedure. As used herein the term

5 "reagents necessary for preparing a nucleic acid library" is to be understood as any mix of reagents known in the art that are sufficient for preparing a nucleic library that can be used for NGS. Preferably, "reagents necessary for preparing a nucleic acid library" can to be construed as any mix of reagents known in the art that are sufficient for preparing a nucleic library that can be directly used for NGS, i.e. comprising NGS-specific adapter sequences that are compatible with a given NGS strategy of choice.

10 Such reagents may comprise primer sequences wherein the NGS-specific adapter sequences are included in the sequences of primers and the reagent mix comprises enzymes, substrates, and buffering conditions which allow to perform a library PCR with such primers. Alternatively, such NGS-specific adapter sequences can be suitable for ligation and can be provided in a reagent mix also comprising enzyme ligase.

15 [0016] Further, the term "means for liberating or purifying nucleic acid from the biological sample" is to be understood as any plurality or chemical reagents and/or physical elements as known in the art that are known to be used for liberating nucleic acids from cells or other structures in a biological sample, and, in case of purification, sufficiently separating said nucleic acids from unwanted sample debris into an acceptably pure form (wherein the term "acceptably" depends on the further purpose of

20 such purified nucleic acids), usually in an aqueous solution. Chemical reagents suitable for such purpose include e.g. any known in the art detergents and/or buffers comprising detergents, chaotropic agents, nuclease inhibitors etc. that are used in tissue or cell disrupting and/or liquefying, and thus releasing nucleic acids contained therein into solution. Similarly, physical elements known in the art to be used in various methods of sample processing for the purpose of nucleic acid liberation/purification

25 include e.g. silica solid supports such as resins in spin columns, silica membranes, beads etc.; further mechanical disruptors or machines generating disruptive energy such as sonicators etc.

[0017] Then, the term "nucleic acid" and its equivalent "polynucleotide", as used herein, refer to a polymer of ribonucleosides or deoxyribonucleosides comprising phosphodiester linkages between nucleotide subunits. Nucleic acids include but are not limited to DNA and RNA, e.g. including genomic

30 DNA, mitochondrial or mtDNA, cDNA, mRNA, rRNA, tRNA, hnRNA, microRNA, lncRNA, and various modified versions thereof. Nucleic acids can most commonly be obtained from natural sources like biological samples obtained from different types of organisms. On the other hand, nucleic acids can also be synthesized, recombined, or otherwise produced in any of the known human-devised methods (e.g. nucleic acid amplification method like PCR).

35 [0018] As used herein, the term "source of a nucleic acid" is to be understood as any substance whether liquid or solid, comprising or expected to comprise nucleic acid. A source of nucleic acid can

e.g. be an artificially created solution comprising a synthetic or recombinant nucleic acid such as among many other a solution containing a ligation product, an electrophoresis marker (so called "ladder"), a primer stock etc. Most commonly however, a source of nucleic acid will be a biological sample obtained from an organism or cells forming or derived thereof, preferably a clinical sample

5 obtained from a patient.

[0019] As used herein, the term "*biological sample*", or simply "*sample*", is intended to include a variety of biological sources that contain nucleic acid and/or cellular material, for example including: cultures of cells such as mammalian cells but also of eukaryotic microorganisms, body fluids, body fluid precipitates, lavage specimen, fine needle aspirates, biopsy samples, tissue samples, cancer

10 cells, other types of cells obtained from a patient, cells from a tissue or *in vitro* cultured cells from an individual being tested and/or treated for disease or infection, or forensic samples. Non-limiting examples of body fluid samples include whole blood, bone marrow, cerebrospinal fluid (CSF), peritoneal fluid, pleural fluid, lymph fluid, serum, plasma, urine, chyle, stool, ejaculate, sputum, nipple aspirate, saliva, swabs specimen, wash or lavage fluid and/or brush specimens.

[0020] Once a biological sample is provided into the systems or during performing the methods of the invention, it will usually be contacted with a composition to provide a lysate in which nucleic acid is released. As used herein, by "*contacting*" is meant bringing together, exposing, incubating, or mixing of the sample and the composition. "*Releasing*" refers to liberating, obtaining and/or reversal of cross-linking. For liberating nucleic acid from a sample, protease activity and pH-buffering may be required

20 from the composition. Releasing may require from the composition potential precipitating activity of components other than nucleic acid present in the investigated sample and removal/dissolving of fixative. Releasing may require conditions such as heating or High-Intensity Focused Ultrasound (HIFU). In one embodiment in accordance with the spirit of the invention, a biological sample is introduced into a cartridge compatible with an automated system such as a diagnostic analyzer,

25 wherein the sample processing steps involving contacting with various solutions and releasing of nucleic acids take place.

[0021] Further, the term "*cartridge*" is to be understood as a self-contained assembly of chambers and/or channels, which is formed as a single object that can be transferred or moved as one fitting inside or outside of a larger instrument suitable for accepting or connecting to such cartridge. Some

30 parts contained in the cartridge may be firmly connected whereas others may be flexibly connected and movable with respect to other components of the cartridge. Analogously, as used herein the term "*fluidic cartridge*" shall be understood as a cartridge including at least one chamber or channel suitable for treating, processing, discharging, or analyzing a fluid, preferably a liquid. In preferred embodiments of the present invention, a cartridge can additionally be equipped in means of physical or mechanical

35 manipulation of received therein samples that aid their lysis and liquefaction. For example, an advantageous cartridge can comprise one-way inlet valve, a septum, one or more filters or affinity

membranes, an overflow, and/or a zone compatible with HIFU treatment. Further, an advantageous cartridge can comprise means for mixing and/or proper distribution of contained therein liquids, such as one or more plungers or manifolds. Next, an advantageous cartridge may comprise several reaction and/or waste chambers, may provide for keeping the received sample fully contained from the 5 external environment, or be made from affordable materials such as plastics and/or other polymers. An example of a suitable cartridge is given in WO2007004103. Advantageously, a fluidic cartridge can be a microfluidic cartridge. In the context of fluidic cartridges the terms “*downstream*” and “*upstream*” can be defined as relating to the direction in which fluids flow in such cartridge. Namely, a section of a fluidic path in a cartridge from which a fluid flows towards a second section in the same cartridge is to 10 be interpreted as positioned *upstream* of the latter. Analogously, the section to which a fluid arrives later is positioned *downstream* with respect to a section which said fluid passed earlier.

[0022] In general, as used herein the terms “*fluidic*” or sometimes “*microfluidic*” refers to systems and arrangements dealing with the behavior, control, and manipulation of fluids that are geometrically constrained to a small, typically sub-millimeter-scale in at least one or two dimensions (e.g. width and 15 height of a channel). Such small-volume fluids are moved, mixed, separated or otherwise processed at micro scale requiring small size and low energy consumption. Microfluidic systems include structures such as micro pneumatic systems (pressure sources, liquid pumps, micro valves, etc.) and microfluidic structures for the handling of micro, nano- and picoliter volumes (microfluidic channels, etc.). Exemplary microfluidic systems have been described in EP1896180, EP1904234, and 20 EP2419705 and can accordingly be incorporated applied in certain embodiments of the presented herein invention.

#### **BRIEF DESCRIPTION OF FIGURES**

[0023] For a fuller understanding of the nature of the present invention, reference is made to the 25 following detailed description taken in conjunction with the accompanying drawings in which:

**Figure 1:** shows 5 DNA bands on an electrophoretic gel, corresponding to 5 products of a 5plex qPCR performed on a liquefied FFPE sample.

**Figure 2:** shows qPCR amplification curves for the 5 products of the 5plex qPCR shown in Figure 1.

**Figure 3:** shows a Cq to copy number histogram for at least 4 replicates of each of the 5plex qPCR 30 products.

**Figure 4:** shows R squared value determination for each of the 5plex qPCR products.

**Figure 5:** shows the ability of the 5plex qPCR to distinguish between different degrees of nucleic acid fragmentation. Panel A shows results obtained from 3 FFPE samples having relatively intact DNA; Panel B shows results from another set of 3 FFPE samples with a higher level of fragmentation; lastly, 35 Panel C shows results from 6 different FFPE samples containing heavily fragmented DNA.

**Figure 6:** shows results of a BRAF-specific qPCR capable of discerning between wt and V600K/R/M BRAF mutants, performed on three FFPE samples each spiked with a plasmid containing a sequence encoding for either wt BRAF, V600M mutant BRAF, or V600K and T149C double-mutant BRAF.

**Figure 7:** shows principles of a one type of NGS-ready library preparation using library PCR with primers containing NGS-specific adapters. Sequences of wild type (Seq ID NO.: 1), V600M (Seq ID NO.: 2) and V600K + V1794C (Seq ID NO.: 3) are represented.

**Figure 8:** shows results of NGS performed on three FFPE samples each spiked with a plasmid containing a sequence encoding for either wt BRAF, V600M mutant BRAF, or V600K and T149C double-mutant BRAF.

10 **Figure 9:** shows an example of an optimized sample-to-result workflow according to the present invention.

#### **DETAILED DESCRIPTION OF THE INVENTION**

[0024] The present invention generally relates to automatic preparation of next-generation sequencing (NGS) libraries from biological samples. In particular, the invention relates to a generic way to generate an NGS-library from any type of sample (blood, stool, urine, FFPE, swab, ...) in an automated manner and with the possibility of automatically verifying quantity and quality of the thus obtained library or the nucleic acid used for its construction using a QC qPCR and of almost instantaneously obtaining clinical information from said nucleic acid using an assay qPCR. In case if the latter, it should be noted that exactly the same and identically-processed nucleic acid pool is used for NGS library preparation and qPCR, which further provides unique means for directly comparing the diagnostic information on key target mutations as screened by assay qPCR with a broader genetic landscape as later obtained from the NGS data.

[0025] In particular, the present invention provides a fluidic cartridge for automated processing a biological sample on an automated system and engageable with said system, the cartridge comprising:

- a sample compartment for receiving a biological sample;
- means for liberating or purifying nucleic acid from the biological sample received in the sample compartment, said means capable of entering in fluid communication with said sample compartment either by being connected by a channel to said sample compartment or by being entirely or partially positioned in the sample compartment;
- a library compartment for preparing a nucleic acid library positioned downstream of the sample compartment and of the means for liberating or purifying nucleic acid, the library compartment configured to receive at least a portion of the liberated or purified nucleic acid (e.g. as a result of a pumping action directing the extracted nucleic acids in a channel to the library compartment) and comprising reagents for preparing a nucleic acid library such as enzymes, NGS-specific adapters, any buffering solutions etc.

[0026] One of the objects of the present invention is to provide means for biological sample-to-NGS library processing that is fast, easy-to-use, and requires minimal or virtually no human handling at least from the moment a sample is inserted into the cartridge of the present invention, the cartridge is engaged with its operating automated system, and the automated system is instructed to run the

5 automated library preparation procedure. Thus, it is also one of the objectives of the present invention to provide a cartridge or a platform wherein the library preparation process does not require any human intervention from the moment a biological sample is provided (e.g. inserted into the cartridge or otherwise introduced into the automated system compatible with the cartridge) and the automated processing is initiated by the user.

10 [0027] Advantageously, the cartridge of the invention not only automatically prepares an NGS library but also is capable of concomitantly providing quantity and quality information related to said library, and/or also performing some clinical characterization of the sample from which the library is being generated. To achieve this, in the present invention further provides a cartridge where in parallel with the library preparation, a real-time PCR (qPCR) can be thermocycled and monitored e.g. through a  
15 transparent wall. Thus, in a preferred embodiment a cartridge is provided that further comprises at least one thermocycling qPCR compartment positioned downstream of the sample compartment and of the means for liberating or purifying nucleic acid, and configured to receive at least a portion of the liberated or purified nucleic acid or at least a portion of the nucleic acid library prepared in the library compartment (e.g. by pumping through channels), wherein said thermocycling qPCR compartment is  
20 suitable for amplifying nucleic acids and allowing detection of signals generated during such amplification and comprises reagents for performing qPCR, such as at least one set of primers, a signal-generating dye or probe, a DNA polymerase, dNTPs, buffer, etc.

[0028] Preferably, the qPCR performed in the thermocycling qPCR compartment is a multiplex qPCR, i.e. a qPCR simultaneously amplifying and detecting multiple sequences in a single reaction. A  
25 multiplex qPCR uses multiple primer sets in a single qPCR mixture and thus generates multiple products in one tube, chamber, or other type of a qPCR thermocycling compartment. Multiplex qPCR using two primer sets (usually pairs) is called a duplex, often denoted 2plex. Similarly, a multiplex qPCR with three primer sets is a triplex or a 3plex. Preferred in the present invention multiplex qPCR arrangements include a 2plex, a 3plex, a 4plex, a 5plex, a 6plex, a 7plex, or more.

30 [0029] Therefore, preferably, the reagents for performing PCR comprise a plurality of primer sets, each directed to a different amplicon, wherein said plurality is preferably 2, 3, 4, 5, 6, 7, or more.

[0030] It is well known in the art that designing a robust multiplex qPCR is not easy as the assay requires that the multiplicity of individual primer sets (usually primer pairs) will specifically target their unique amplicons in one reaction tube and thus under a single set of reaction conditions. Primer  
35 design will typically take into account primer purity (17 to 30 bases in length); balance G/C and A/T-rich domains (20 to 70% G+C); set melting temperature between 55-80°C; avoid creating

complementary 3'end base pairs; avoid primer self-complementary; and avoid 1 or more C's or G's at the 3' end of primers, especially when the multiplex is to be performed on complex samples such as eukaryotic genomic DNA. Several web-based primer design software tools are available that help design PCR primers (e.g. Primer3Plus). Other factors, such as the relative concentration of the primers, the right concentration of the PCR buffer components, balance between the magnesium chloride and deoxynucleotide concentrations, cycling temperatures, and DNA thermocycling polymerase etc. also often have to be fine-tuned for a successful multiplex qPCR. Notably, finding an optimal combination of the annealing temperature and buffer concentration is essential in multiplex PCR to obtain highly specific amplification products. Magnesium chloride concentration needs to be proportional to the amount of dNTP, while primer concentration for each target should be relatively robust. A choice of a proper polymerase can also have impact on the outcome of the reaction. In theory, multiplex PCR can be performed with standard PCR polymerase; however in practice it is preferred that highly processive and sensitive DNA polymerases such as GoTaq or AmpliTaq are used. For particularly sensitive applications, further modifications of multiplex qPCR such as the ones using DNAzymes or MNAzymes may be advantageous for the application in the present invention. The list of the mentioned-herein factors and multiplex optimization strategies is by no means to be interpreted as extensive, which will be immediately appreciated by any person skilled in the art.

[0031] In preferred embodiments of the invention, the quantification of the qPCR in the systems and methods of the invention is based on the standard curve method; however, other quantification strategies can also be used, as will be immediately appreciated by any skilled person.

[0032] Similarly, many different general types of fluorescent dyes or detection probes can be used in the systems and methods of the present invention. In preferred embodiments sequence specific probes will be used, e.g. selected from exonuclease probes, hybridization probes, or molecular beacons. Such probes not only add specificity to the assay, but are also key for enabling multiplex applications. As shown in the examples, the methods of the invention were successfully applied with the use of Taqman probes but other probes would be suitable as well.

[0033] Advantageously, together with the NGS library preparation, the cartridge of the present invention provides for simultaneous qPCR screening of the provided therein clinical sample for the presence of therapy actionable alleles or disease-associated sequences. Some examples of such sequences, among many others, include mutant gene sequences associated with cancer, or exogenous pathogenic sequences indicative of an ongoing infection with a pathogenic organism. In line with this, in a preferred embodiment, a cartridge is provided, wherein the reagents for performing qPCR comprise at least one set of primers, preferably a plurality of sets of primers for amplifying at least one disease-associated target nucleic acid sequence.

[0034] In a further aspect of the present invention, the provided herein cartridge can also advantageously comprise means for assessing the quality of the purified nucleic acids or the already

obtained library. Thus, in another preferred embodiment, a cartridge is provided wherein the reagents for performing qPCR comprise at least one set of primers, preferably a plurality of sets of primers for performing a quality control (QC) qPCR. In principle, for performing QC qPCR, targets can be selected from any genes or genomic regions. For diagnosis of genetically unstable conditions like cancer, it is better however, to avoid disease-sensitive regions that are likely to have their sequence mutated or change in copy number. Therefore, in a preferred embodiment, the targets of the quality control multiplex qPCR are selected from intra-exon sequences of single copy genes, such as a housekeeping gene. A further advantage of said solution is that the same target sequences directed to intra-exon regions can be used in assessment of both DNA as well as in RNA quality as a library material. Because of the latter, such design of the quality control multiplex qPCR multiplex of the present invention is particularly useful when it is desired to sequence both the genome and the transcriptome from one sample, which requires construction of and thus also quality verification for both DNA- and RNA-based. Thus, in one preferred embodiment, the quality control multiplex qPCR to be used in systems and method of the invention, targets at least one intra-exon sequence in a single copy gene. Possibly, one, two, three, four, five, six, seven or more intra exon sequence in a one, two, three, four, five, six, seven or more single copy genes are targeted. Alternatively to housekeeping genes, in DNA library construction, repetitive sequences (e.g. LINEs, or SINEs such as Alu elements) may be the targets of interest. As a non-limiting example shown in the example section, the methods of the invention are successfully practiced on intra-exon sequences of the RNaseP, HPRT1, Beta-Actin, TRFC and ABCB1 genes.

[0035] In one preferred embodiment, the cartridge of the invention can perform QC qPCR in addition to an assay qPCR that amplifies target disease-associated nucleic acid sequences. Thus, in an advantageous embodiment, a cartridge is provided further comprising at least a second one thermocycling qPCR compartment positioned downstream of the sample compartment and of the means for liberating or purifying nucleic acid, the at least second one thermocycling qPCR compartment configured to receive at least a portion of the liberated or purified nucleic acid or a portion of the nucleic acid library prepared in the library compartment, and comprising a at least one set of primers for performing a nucleic acid quality control (QC) qPCR.

[0036] It has been observed that a multiplex qPCR generating amplicons of different sizes is particularly suitable for assessing nucleic acid quality. Consequently, in another preferred embodiment, a cartridge is provided comprising a plurality of sets of primers for generating amplicons of different sizes. The size range of the amplicons varies from a smaller size Ax to a higher size Ay (with x<y), typically in the size range of 50-600 bp. Preferably the smaller Ax size ranges from 50 to 110 bp and can be of any length in between that range. Preferably Ax is between 60 bp +/- 10 bp. As shown further herein in Examples, Ax is 63 bp or 105 bp. Preferably the larger Ay size ranges from 300 to 550 bp and can be any length in between that range. As shown in further in the Examples, Ay

is 504 bp (e.g. for the TRFC gene) or alternatively 318 bp (e.g. for the beta-actin gene). In case of RNA sequencing, particularly in applications focusing on microRNA (e.g. 15-35 bp size range), the smaller Ax size range may need to be lowered accordingly. As will be appreciated by any skilled artisan, during the multiplex qPCR design for the purposes of the present invention, the preferred 5 amplicon sizes may advantageously be selected in accordance with the NGS application of choice in order to provide the most adequate estimation about the NGS coverage according to amplicon length.

[0037] In preferred embodiments, the biological sample is a human clinical sample. The clinical sample can be a fresh sample, a fresh frozen sample, a fine needle aspirate, a sample that has been treated for preservation and may contain cross-linking of reactive sites due to fixation treatment, a 10 wax-contacted or wax-embedded sample, an FFPE sample in the form of an FFPE slice, a liquid sample such as a urine sample, a blood sample, a serum sample, or any other clinical sample.

[0038] Once introduced to the cartridge of the invention, a biological sample will usually be processed by contacting it with a composition that provides for releasing of nucleic acids. In preferred 15 embodiments, the composition is optimized for use in microfluidic analyzers and preferably contains surfactants rather than organic solvents. Mixing with said composition usually also facilitates transporting of such processed sample through a microfluidic system. In embodiments wherein the sample is an FFPE sample, the surfactant comprised in said composition will preferably be non-ionic. Nucleic acids obtained from FFPE samples typically contain nucleotide-to-nucleotide and nucleotide-to-protein cross-links, base modifications and other chemical modifications that affect the integrity of 20 the nucleic acid. Preferred methods of the present invention incorporate a non-ionic surfactant and permit automated removal of embedded wax and liberation of the components without use of organic solvents. This is particularly beneficial because it puts the liberated nucleic acids in a condition and environment that interfaces with downscale applications requiring enzymatic activity such as nucleic acid amplification via PCR. In one embodiment, the lysate and/or components released from the 25 sample will be further processed in diagnostic analyzers using microfluidic systems.

[0039] Optionally, the liberated nucleic acid is provided in a form sufficiently pure for being directly used as a template for a qPCR and a nucleic acid library construction according to the invention. In a preferred embodiment, the nucleic acid-liberating means performs its function in a fluidic or microfluidic arrangement. In such instance, the elements forming such nucleic acid-liberating means 30 may comprise a series of consecutive or otherwise fluidly interconnected compartments, like chambers or channels, at least some of which being supplied with reagents like lysis buffers, enzyme solutions, extraction buffers, binding buffers and/or wash buffers; or optionally comprising any of the known in the art physical barriers, such as filters or high-affinity resins, that facilitate processes like mechanical sample clearing or nucleic acid binding, washing, and releasing. Such and alternative 35 means for liberating nucleic acids are well known in the art and therefore will not be discussed herein in greater detail.

[0040] Usually, the fluidic cartridge of the invention will comprise means for dividing nucleic acids liberated and/or purified from a received therein sample between at least the library compartment and the thermocycling qPCR compartment. Such means could e.g. comprise two separate channels extending from the nucleic acid source-receiving compartment or the compartment whereto the 5 liberated from said source nucleic acid is deposited in the last step of the nucleic acid liberation process, into the thermocycling qPCR compartment and into the library compartment, respectively. In order to actively transport the nucleic acid between the compartments of choice, the automated system of the invention could provide a pressure gradient capable of pushing or pulling a desired amount of fluid into prescribed direction. Generation of such pressure gradients by means of pumps, 10 suction devices, manifolds etc. is widely employed in contemporary microfluidic systems and thus well known in the art.

[0041] As will be understood by any skilled person, fluidic or microfluidic cartridges suitable for the present invention may contain at least two reaction chambers comprising the thermocycling qPCR compartment and the library compartment, and one or more fluid chambers. Some of the fluid 15 chambers may hold fluid which is used for producing lysate from the sample. Other chambers may hold fluids such as washing fluids and amplification solution. Separate reaction chambers are used as the thermocycling qPCR compartment and the library compartment. The chamber configured to serve as the thermocycling qPCR compartment comprises a number of primer sets, along with other amplification reagents and enzymes required for performing a qPCR. The other chamber configured to 20 serve as the library compartment is adapted to performing the steps of constructing a nucleic acid library for an NGS application of choice. Parts of the sample will be transferred to the reaction chambers and to make such transfer possible, chambers are connected to one or more fluid channel. In at least one, but preferably each of these fluid channels a valve means may be provided, which valve means preferably normally closes the fluid channel, but opens the fluid channel upon actuation 25 of the valve means therewith placing the respective two chambers in fluid communication. The valve means may be designed as a one-way valve.

[0042] In another embodiment, the library compartment and the one or more of the thermocycling qPCR compartments of the present cartridge can be operated in the automated system engaging the cartridge simultaneously or sequentially. This means that three modes of operation can be envisaged: 30 (i) both compartments operate once nucleic acid is fed into them, thus library preparation is independent of and proceeds in parallel with qPCR; and (ii) first, qPCR is made, then library is made; like this the decision to prepare a library for sequencing is made once the results of the qPCR are known and may depend on these results; (iii) first, the library is made, then qPCR is performed on the library for the verification of the library quality.

35 [0043] The option (i) above describes the simultaneous operation in which the preparation of a library for NGS from a part of a nucleic acid sample is run simultaneously with a qPCR assay on another part

of the same nucleic acid sample, and both are preferably performed in parallel in a cartridge-based microfluidic system. As used herein the term “*simultaneously*” or “*in parallel*” refers to happening or being done at the same time. In such arrangement, the qPCR is being performed on a part from a nucleic acid sample at the same time as the nucleic acid library for NGS application is being 5 constructed from another part from the same nucleic acid sample. In other words, during the simultaneous operation, both the sample analysis via qPCR and library construction are executed by the automated system of the invention at the same time.

[0044] Conversely, the options (ii) and (iii) above can both be described as operating “*sequentially*”. In one possible embodiment of the sequential operation at least two thermocycling qPCR 10 compartments operate on the automated system of the invention. For example, first qPCR can be done to read the expression of interesting markers and verify the quality of nucleic acid source fed into the system. Following this first qPCR (sequentially), or to save time in parallel with said first qPCR (simultaneously), a library is constructed. Then, a second or control qPCR can be performed on the thus constructed library to verify whether its quality is sufficient for subjecting it to further applications, 15 such as sequencing.

[0045] With regard to the NG sequencing library preparation or construction, currently there exist many different ways of generating a sequencing-ready library, and their choice naturally depends on which NGS strategy is intended to be performed. In general, NGS library generation involves 20 generation of nucleic acid fragments, which are compatible with given NGS. Therefore, in a preferred embodiment, a cartridge is provided wherein the reagents for preparing a nucleic acid library in the library allow:

- generation of nucleic acid fragments from the portion of the liberated or purified nucleic acid received into the library compartment, and
- attaching oligonucleotide adapters to at least one, preferably both ends of said generated 25 nucleic acid fragments.

[0046] For most of the commercially available NGS platforms, amplification of nucleic acid fragments is necessary to generate sufficient copies of sequencing templates. Thus, preferably, a cartridge is provided wherein the generation of nucleic acid fragments is performed by a PCR, further referred to as “library PCR”, and wherein the attaching oligonucleotide adapters to the library-PCR-generated 30 nucleic acid fragments is performed by including an adapter sequence in a sequence of at least one primer used in said library PCR. Suitable library PCRs are known in the art and include methods such as bridge amplification or emulsion PCR.

[0047] As already state above, nucleic acid library construction is needed for DNA sequencing, RNA sequencing, and other applications such as sequencing-based methylation analysis. RNA sequencing 35 (RNA-seq) is a method of investigating the transcriptome of an organism using deep-sequencing

techniques. Total RNA generally contains only a very small percentage of coding or functional RNA; ribosomal RNA (rRNA: up to 80–90% of the total RNA), and to a lesser degree transfer RNA (tRNA), make up the majority of the RNA in a sample. Often, in order not to use 80–90% of one's sequencing capacity on repetitive rRNA sequences, rRNA can be removed from the sample prior to sequencing.

5 The RNA after removal of rRNA is made into a library. This involves creating double-stranded cDNA through reverse transcription from the RNA (or fragmented RNA). This double-stranded cDNA may then be handled as normal genomic DNA throughout the remaining library construction process, including linking it with appropriate NGS-strategy specific adapters..

[0048] As the cartridges of present invention provide a convenient way for generating parallel datasets from several assays including NGS, sample quality assessment by QC qPCR, and possibly several target specific qPCR assays, in a particularly practical embodiment, cartridges of the present invention may advantageously be equipped with an element, i.e. an identifier, capable of storing cartridge- or patient-specific information, which would allow to easily keep track of all the assays performed on one cartridge or on samples from one patients. Such identifier could for example be a cartridge-unique code or a barcode containing data like a cartridge number, its batch or series number, production and expiry dates, types of target specific assays that the cartridge carries etc. Alternatively, the identifier could be not only readable but also writable storage medium like a disc or a microchip, where a patient-specific data (like name, age, date of the assay etc.) could be introduced, stored, and then retrieved at will in order to facilitate automated patient identification and minimize mixing of samples. The identifier could be conveniently scanned or otherwise introduced into the automated system engaging with and processing such cartridge, which would save the time of the operator performing the run on typing the required data and also eliminate a stage in the workflow where a human error can be introduced. Furthermore, thanks to the association of runs with one specific cartridge-identifier or patient-identifier, different tests performed on samples from the same patient could become easily traceable, retrievable, and associated with one another also in an automated way. Thus in an advantageous embodiment, the present invention provides a cartridge in accordance with the previous embodiments, further comprising a cartridge-specific identifier like a unique code, barcode, or a microchip etc., for storing cartridge-specific information.

[0049] In another aspect, a cartridge is provided further comprising a recovery compartment positioned downstream of the sample compartment, said recovery compartment providing for recovering any of the following:

- a portion of the biological sample received into the sample compartment;
- a portion of the nucleic acid liberated or purified from the biological sample received into the sample compartment;
- at least a portion of the nucleic acid library prepared in the library compartment

[0050] Such recovery compartment may comprise or simply be made of another chamber wherein no

reaction takes place during the operation of the system of the invention. Such recovery would preferably be easily accessible from outside the present automated system or a cartridge of the automated system. For example, it could comprise a wall made of a pierceable material (e.g. a foil or a film) that can be pierced by a needle of a syringe or a pipette, allowing aspiration of its contents.

5 Alternatively, the recovery compartment could be selectively brought in fluid communication and filled in with any of the above by means on pumping and following instructions given by the user through an interface of the automated system of the invention.

[0051] In a possible embodiment, such recovery compartment could be an external container e.g. plastic tube or a vial, engageable with or connectable to the automated system of the invention. In 10 such instance, any of the compartments as follows:

- a compartment housing at least a part of the biological sample introduced into the cartridge or at least a part of the nucleic acid liberated from said sample;
- the library compartment;
- at least one thermocycling qPCR compartment;

15 could comprise a structure (e.g. an extension like a channel or an zone engageable with an element forming a channel) capable of brining it in fluid communication with the recovery compartment by any means capable of transporting at least a part of the content comprised in any one of said above-listed compartments into the recovery compartments.

[0052] In an advantageous embodiment, the library compartment can comprise a structure capable of 20 brining it in fluid communication directly with a compartment where NGS is performed, possibly wherein said compartment is comprised in another subsystem comprising an automatic sequencer being part of the automated system processing the cartridge.

[0053] In a further aspect, the present invention also provides an automated method for preparing a 25 next-generation sequencing (NGS) nucleic acid library from a biological sample, said method comprising the steps of:

- a) receiving a biological sample;
- b) liberating or purifying at least a part of nucleic acids from the received biological sample;
- c) providing at least a portion of the liberated or purified nucleic acid into a library compartment for preparing an NGS nucleic acid library and comprising reagents for preparing a nucleic acid library; and
- d) preparing an NGS nucleic acid library;

30 wherein at least the steps b) to d) are performed on an automated system without any human manual intervention, at least after the biological sample has been received in said automated system.

[0054] In a preferred embodiment, the step of preparing an NGS nucleic acid library is performed by 35 PCR.

[0055] In another particularly preferred embodiment in accordance with the above embodiments of the method of the invention, the method further comprises the step of performing qPCR on said automated system either sequentially or simultaneously with respect to the step of preparing the NGS nucleic acid library, wherein said qPCR is performed on a second portion of the liberated or purified nucleic in a thermocycling qPCR compartment suitable for amplifying nucleic acids and allowing detection of signals generated during such amplification.

[0056] As explained above, one of the aspects of the invention involves preparing an NGS library concomitantly with running either an assay qPCR and/or a QC qPCR using nucleic acid obtained from the same sample. Consequently, another embodiment, a method is provided wherein said at least one qPCR amplifies at least one disease-associated target nucleic acid sequence, or is a quality control (QC) qPCR.

[0057] In a particularly preferred embodiment, the method of the invention is performed on a cartridge engageable with the automated system. Thus, most preferably the afore mentioned steps of obtaining nucleic acid from a sample, NGS library preparation, assay qPCR, and QC PCR are performed in one cartridge, preferably being a fluidic or microfluidic cartridge engageable with an analyzer-type apparatus so that the cartridge is a self-contained disposable platform for performing the steps of the method according to the present invention. In such advantageous embodiment, all of the reagents required for performing the method of the invention are pre-positioned within such cartridge, for storage considerations preferably in a dried-down or a lyophilized form.

[0058] In another preferred embodiment, for the advantageous considerations as described above, desirably a method is provided wherein the cartridge comprises a cartridge-specific identifier for storing cartridge-specific information and wherein the method further comprises the step of introducing to the automated system the information stored in the cartridge-specific identifier.

[0059] In summary, the present invention provides an effective automation of workflow with integrated “sample-in, quality checked nucleic acid library-out” strategy. The presented herein approach has a great potential for providing a minimal turn-around times, lower costs, and improved NGS success rates. Furthermore, the cartridges, automated systems, and methods of the invention not only afford for simplification and acceleration of the library construction workflow, but they are also universally adaptable for the use with different sample types, even the challenging ones like FFPE samples. Thus, in a final aspect of the invention, use is provided of cartridges, automated systems, and methods of the invention for automated preparation of a next-generation sequencing (NGS) library from a biological sample.

[0060] It is to be understood that both the foregoing general description and detailed description are only exemplary and explanatory and are not restrictive for the invention as claimed. In this application, the use of singular includes the plural unless specifically stated otherwise. In this application the use of

"or" means "and/or" unless stated otherwise. The use of the terms "including", "includes" or "included" is not limiting.

5 **EXAMPLES**

***Example 1: Development of a QC qPCR for an automated sample-to-output assessment of nucleic acids***

[0061] First, a quality control (QC) qPCR assay was developed for the purpose of assessing the amount and quality of nucleic acids present in a sample in a fully automated manner. The present QC 10 qPCR tests for the presence of amplicons of various lengths, each derived from a different single copy human gene, and serves to assess nucleic acid suitability for NGS application. The amplicons and their lengths are as follows: (1) 63bp fragment from human RNaseP gene; (2) 105bp fragment from HPRT; (3) 149bp fragment from TFRC; (4) 213bp fragment from ABCB; and (5) 318bp fragment from  $\beta$ -actin. The amplification of the fragments in one PCR reaction (5plex) was initially verified using a 15 qPCR performed on a liquefied FFPE sample (Horizon FFPE sample) with a GoTaq polymerase and Taqman probes (composition as specified by the supplier). The qPCR programme was 5' hold 95°C, followed by 50 cycles of 5" 95°C 44" 64°C. Figure 1 shows the obtained fragments (left lane) next to a DNA ladder (right lane) on a SYBR green stained 10% polyacrylamide gel following electrophoresis in TBE. The corresponding qPCR profile of the same sample shown in Figure 2 (sizes of amplicons 20 indicated next to the corresponding curves). The Cq values determined with the regression algorithm contained within the Biorad CFX Manager 3.1 are : (1) 26.1 for the 63bp fragment (RNaseP); (2) 25.6 for the 105bp fragment (HPRT); (3) 26.0 for the 149bp fragment (TFRC); (4) 26.2 for the 213bp fragment (ABCB); and (5) 27.4 for the 318bp fragment ( $\beta$ -actin).

[0062] Next, the 5plex performance on unfragmented human genomic DNA was assessed to obtain 25 standard curves with R squared values for each of the 5 amplicons. To do so, non-fragmented human genomic DNA at 173 $\mu$ g/ml (Promega) was used as a substrate and the copy number was deduced using 3.3pg/haploid genome as a premise. 4 replicates at 24000 copies per PCR, 4 replicates at 6000 copies, 8 replicates at 1500 copies, 8 replicates at 375 copies, 12 replicates at 94, 16 replicates at 23 copies, 20 replicates at 5.9 copies and 24 replicates at 1.5 copies per PCR were amplified and the 30 Cqs determined as described above. The median Cq values were determined while omitting non-amplifications (so called flatliners). The histogram representing this experiment is shown in Figure 3. The standard curves were deduced for each amplicon using logarithmic regression and the R squared value was determined as exemplified in Figure 4 for the complete dataset and the dataset without the 35 Cq values from both 5.9 and 1.5 copies per PCR, respectively. As expected, the R squared values approach 1 better with only data points in double digits copy number. For Cq values below 34, equations with the highest R squared values were used. Notably, for Cq values above 34, the

quantification is known to be less accurate due to stochastic effects. The thus calculated according to said equations copy numbers of each of the 5 amplicons allow for the determination of both the useful DNA content and the degree of nucleic acid fragmentation in a given sample. An analysis of the direction coefficient of the linear regression between the log<sub>2</sub>(copy number input) and the C<sub>q</sub> provides 5 further indication of the efficiency of amplification in 5plex qPCR of each of the amplicons. As known in the art, a perfect qPCR would be assumed to double amount of amplicon (and hence also the net Taqman fluorescence) per cycle, leaving an absolute direction coefficient of 1. In line with this, as shown in Figure 4, the absolute direction coefficients for all amplicons except for the largest amplicon of 318bp are 0.9 or higher, indicating a robust amplification close to doubling Taqman probe 10 degradation per PCR cycle. The largest amplicons size shows >0.8 absolute direction coefficient indicating that the largest fragment amplifies slower and that there is little point in designing a PCR QC test with longer amplicons for this type of Taqman probe-based assay.

15 ***Example 2. Automated FFPE sample processing, nucleic acid quality assessment, target actionable marker screening, and library construction***

[0063] For the purpose of demonstrating the feasibility of the present invention, a set of Biocartis Idylla cartridges was prepared, each cartridge comprising in separate PCR chambers: (i) reagents for performing the above-described QC 5plex qPCR, (ii) reagents for performing target qPCR for 20 detecting wt and V600M/R mutant BRAF, and (iii) reagents for constructing a DNA library compatible with Illumina MiSeq sequencer. Next, a set of FFPE samples to be analyzed on said cartridges were 25 spiked with plasmids encoding for human BRAF. To simulate clinical reality, different BRAF sequences were used including a fragment containing a wild type (wt) BRAF sequence, a fragment encoding for a V600M mutation, and a fragment encoding for two mutations V600K and T149C. The two mutated fragments were spiked in different amounts with respect to the amount of the wild type 30 copies present in the FFPE samples to obtain a relative concentration of 10% and 5%, respectively. Each of the different BRAF-spiked FFPE samples was introduced into a separate cartridge and processed in a fully automated manner on the Biocartis Idylla instrument. In brief, the processing involved sample liquefaction (as described in e.g. WO2014128129), followed by nucleic acid 35 purification on a silica membrane provided in the cartridge, and then followed by three independent and individually-controlled PCR reactions performed in parallel, which included: (i) verification of the quality of the purified nucleic acids via quality control 5plex QC q PCR, as described above; (ii) real time detection of selected BRAF targets (qPCR for target actionable mutations); and (iii) construction of a DNA library using BRAF-specific or standard random-priming primers comprising linkers compatible with Illumina MiSeq sequencer. The latter library-construction PCR was performed using a Q5 high fidelity hot start polymerase (New England Biolabs) and cycled according to the following programme: 5' at 95°C and 50 cycles of 90" at 60°C, 5" at 94°C.

[0064] Figure 5 shows the results of the 5plex QC qPCR on the different FFPE samples, which provide information with regard to the integrity of the DNA present in said samples. Panel A shows three examples of FFPE tissue samples that contain relatively intact DNA. Panel B shows three other examples that have a slightly higher degree of DNA fragmentation. Lastly, Panel C shows 6 examples of FFPE tissue samples that contain heavily fragmented DNA, which is a counter-indication for subjecting such samples to further analysis by NGS.

[0065] Based on the results of the 5plex QC qPCR, three samples with relatively intact DNA and containing three different forms of spiked BRAF (wt, V600M mutant, or double mutant V600K + T149C) were selected further investigation. Firstly, the results obtained from the assay qPCR capable of detecting wt BRAF and V600M BRAF mutation were checked to confirm the presence of the correct BRAF form. The results are shown in Figure 6. They demonstrate that in all of the screened three FFPE samples, wt BRAF signal could be detected (Figure 6, left column, the term "target" refers to wt BRAF sequence). This result was expected as all FFPE samples prior to spiking with different BRAF plasmids were known to contain wt genomic BRAF sequence. Concerning the detection of the V600M mutant (Figure 6, right column, the term "target" refers to V600M/K BRAF sequence), as expected, in the sample spiked only with the wt BRAF-encoding plasmid, no V600M mutant could be detected (Figure 6, top right pane; flat signal line for the target). However, in the FFPE samples spiked with either the V600M mutant, or the double mutant V600K + T149C, the mutation V600M was correctly detected at the expected amounts (Figure 6, right column, bottom and middle pane). Because the used-herein BRAF-specific qPCR did not include a specific probe for the T1794C mutation, said mutation could not be detected in the double mutant BRAF-spiked sample.

### ***Example 3. Sequencing of the selected NGS libraries***

[0066] To confirm the results of the BRAF-specific qPCR and to also detect the presence of the undetected T1794C mutation, the NGS libraries constructed from the same three selected FFPE samples plasmid (wt, V600M, or V600K and T149C) were then subjected to Illumina MiSeq sequencing. The library PCR used for constructing these libraries is schematically shown in Figure 7 and, as mentioned above, was performed on the same cartridges as and in parallel with the 5plex QC PCR and the BRAF-specific assay qPCR. The library PCR included 50 cycles and used simplified BRAF-specific fusion primers (also known as tailed primers). The fusion primers (shown in middle pane of Figure 7) in addition to the target (BRAF)-specific sequence also contained sequencing primer sequence and a tag (P5 and P7) for flow cell attachment. In addition, the reverse primer also contained a barcode (or and index) that during a sequencing run allows to discern between samples obtained from different sources. The reason for introducing such barcode is that typically, libraries constructed from different samples or patients are pooled and sequenced on the same NGS instrument. Thus, in order to differentiate between libraries obtained from the three different FFPE samples, each cartridge contained a slightly different reverse primer having a unique barcode sequence.

[0067] Before sequencing, the three NGS libraries were recovered from respective cartridges using a needle syringe, after which the samples were pooled and purified further on the bench to remove any unreacted primers and primer-dimers. It should be noted that the latter purification step can also be performed automatically. Finally, the purified NGS-libraries were loaded into a flow cell of the MiSeq 5 Illumina instrument and sequenced.

[0068] The results of the sequencing run are shown in Figure 8. In line with the afore-described results of the BRAF-specific qPCR, in the first sample that contained only wt BRAF no mutations were detected. For the two other samples, all the expected BRAF gene mutations were correctly identified during sequencing, even if they could not be captured in the BRAF-specific qPCR. In particular, NGS 10 not only detected the T1794C BRAF mutation missed on the target-specific qPCR, but also allowed to discriminate between the V600M and V600K mutations in mutant- and double mutant-spiked FFPE samples, respectively, thus providing even more exact identification of the already detected mutations. The present results demonstrate the unprecedented robustness of the present invention, wherein 15 desired results are not only provided in a fast and efficient way, but also can be successfully followed up at will if deeper insight is desired.

[0069] For the fuller appreciation of the present invention, the above-described workflow is schematically illustrated in Figure 9. It starts from providing an FFPE sample into a cartridge, after which the subsequent steps till obtaining of the final qPCR results (of both QC qPCR and target actionable marker qPCR, here BRAF) and ready-to-use library are performed in a fully automated and 20 rapid manner (real time frames provided). It should be noted that all these results and the library are obtained from the same and identically-processed sample, which ensures high-comparability of the data from different assays. Notably, by concomitant NGS library construction and providing information with regard to said library's quality, the present approach not only allows to quickly subject a given 25 sample to an NGS clinical follow-up, but also allows to decide whether such rather costly follow-up is feasible in view of that sample's quality. In view of the above, the present invention opens new possibilities in the current diagnostic practice.

**CLAIMS**

1. A cartridge for automated processing a biological sample, the cartridge comprising
  - a sample compartment for receiving a biological sample;
  - means for liberating or purifying nucleic acid from the biological sample received in the sample compartment, said means capable of entering in fluid communication with said sample compartment;
  - a library compartment for preparing a nucleic acid library positioned downstream of the sample compartment and of the means for liberating or purifying nucleic acid, the library compartment configured to receive at least a portion of the liberated or purified nucleic acid and comprising reagents for preparing a nucleic acid library.
2. Cartridge according to claim 1, said cartridge further comprising at least one thermocycling qPCR compartment positioned downstream of the sample compartment and of the means for liberating or purifying nucleic acid, and configured to receive at least a portion of the liberated or purified nucleic acid or at least a portion of the nucleic acid library prepared in the library compartment, wherein said thermocycling qPCR compartment is suitable for amplifying nucleic acids and for allowing detection of signals generated during such amplification and comprises reagents for performing qPCR.
3. Cartridge according to claim 2, wherein the reagents for performing qPCR comprise at least one set of primers, preferably a plurality of sets of primers for amplifying at least one disease-associated target nucleic acid sequence.
4. Cartridge according to any of the preceding claims, said cartridge further comprising at least a second one thermocycling qPCR compartment positioned downstream of the sample compartment and of the means for liberating or purifying nucleic acid, the at least second one thermocycling qPCR compartment configured to receive at least a portion of the liberated or purified nucleic acid or a portion of the nucleic acid library prepared in the library compartment, and comprising a at least one set of primers for performing a nucleic acid quality control (QC) qPCR.
5. Cartridge according to claim 4, comprising a plurality of sets of primers for generating amplicons of different sizes.
6. Cartridge according to any of the preceding claims, wherein the reagents for preparing a nucleic acid library allow:
  - generation of nucleic acid fragments from the portion of the liberated or purified nucleic acid received into the library compartment, and

- attaching oligonucleotide adapters to at least one, preferably both ends of said generated nucleic acid fragments.

7. Cartridge according to claim 6, wherein the generation of nucleic acid fragments is performed by a PCR, further referred to as "library PCR", and wherein the attaching oligonucleotide adapters to the library-PCR-generated nucleic acid fragments is performed by including an adapter sequence in a sequence of at least one primer used in said library PCR.

5

8. Cartridge according to any of the preceding claims, further comprising a cartridge-specific identifier for storing cartridge-specific information.

10

9. Cartridge according to any of the preceding claims, further comprising a recovery compartment positioned downstream of the sample compartment, said recovery compartment providing for recovering any of the following:

- a portion of the biological sample received into the sample compartment;
- a portion of the nucleic acid liberated or purified from the biological sample received into the sample compartment;
- at least a portion of the nucleic acid library prepared in the library compartment.

15

10. An automated method for preparing a next-generation sequencing (NGS) nucleic acid library from a biological sample, said method comprising the steps of:

20

- e) receiving a biological sample;
- f) liberating or purifying at least a part of nucleic acids from the received biological sample;
- g) providing at least a portion of the liberated or purified nucleic acid into a library compartment for preparing an NGS nucleic acid library and comprising reagents for preparing a nucleic acid library; and

25

- h) preparing an NGS nucleic acid library;

wherein at least the steps b) to d) are performed on an automated system.

11. Automated method according to claim 10, the method further comprising the step of performing qPCR on said automated system either sequentially or simultaneously with respect to the step of preparing the NGS nucleic acid library, wherein said qPCR is performed on a second portion of the liberated or purified nucleic acid in a thermocycling qPCR compartment suitable for amplifying nucleic acids and allowing detection of signals generated during such amplification.

30

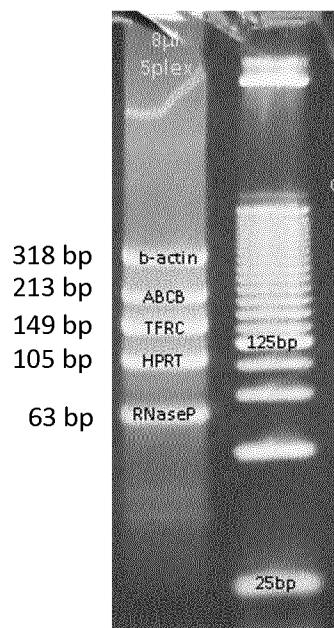
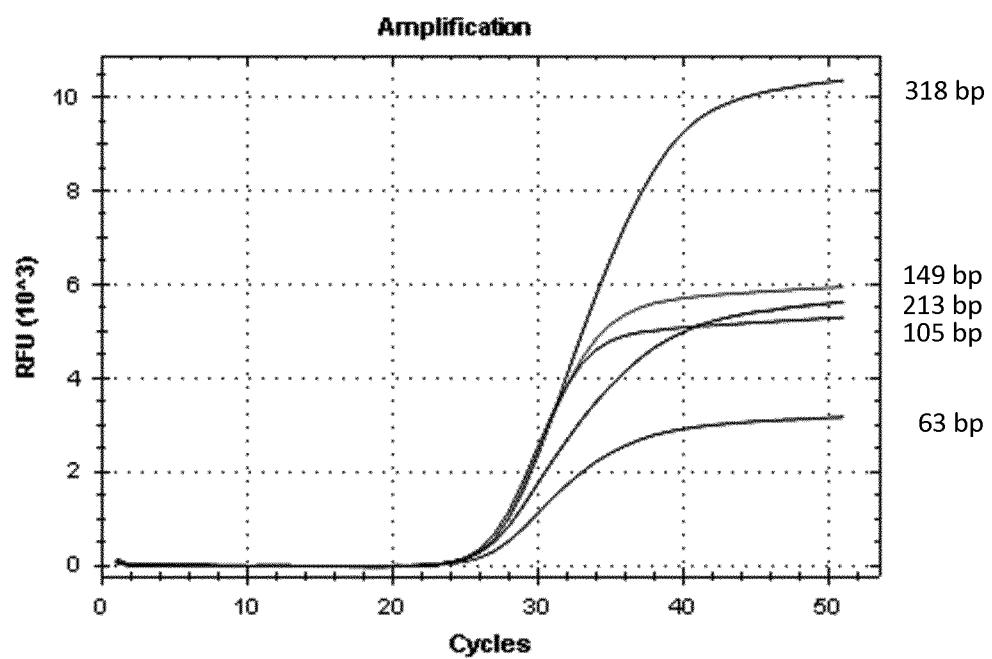
12. Automated method according to claim 11, wherein said at least one qPCR amplifies at least one disease-associated target nucleic acid sequence, or is a quality control (QC) qPCR.

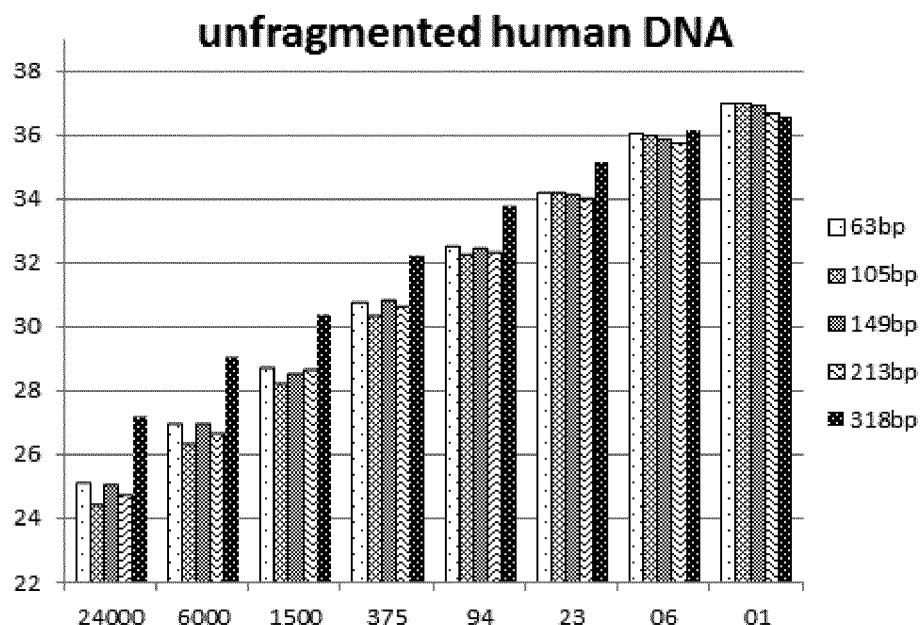
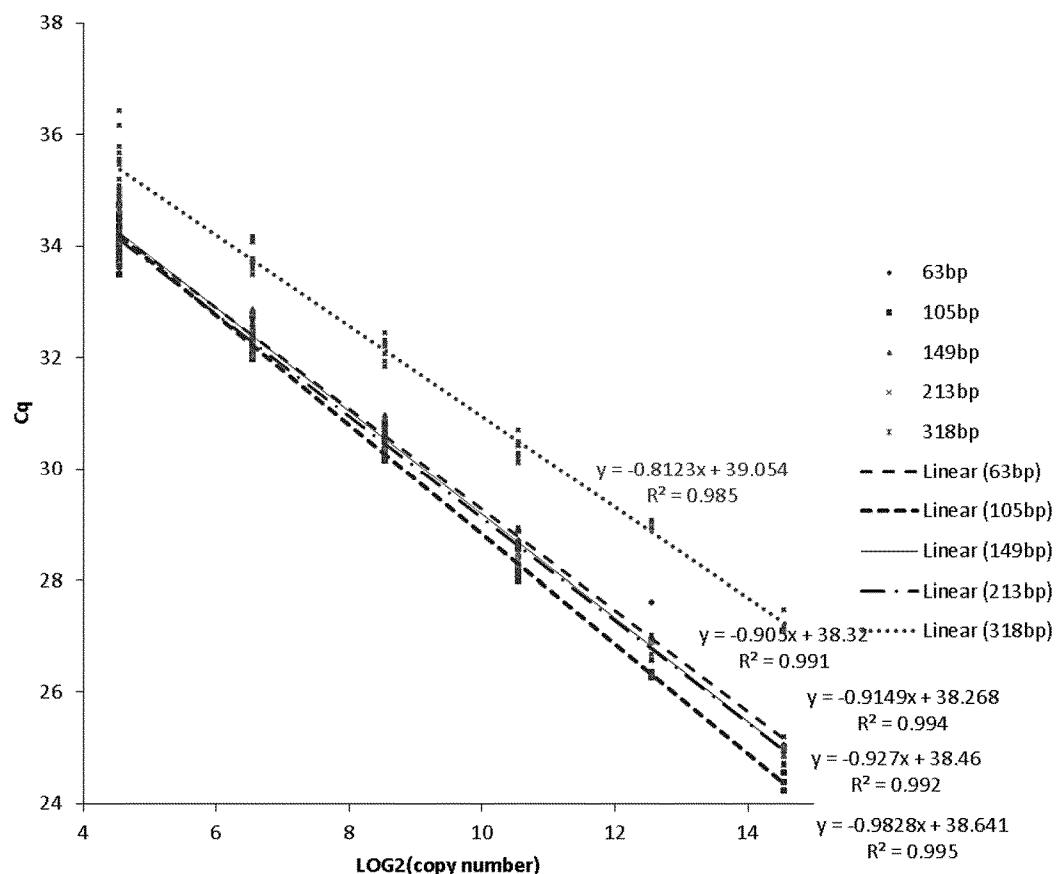
35

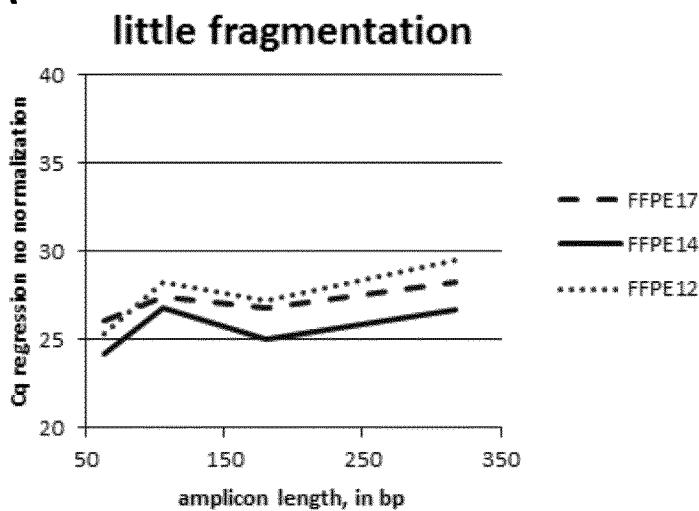
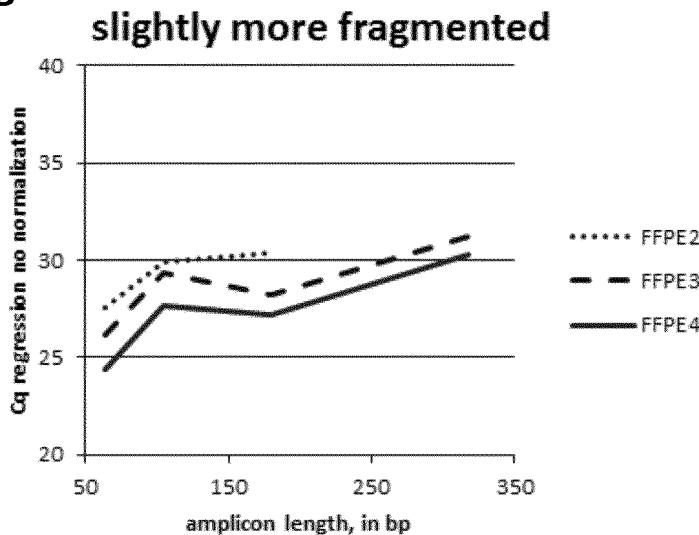
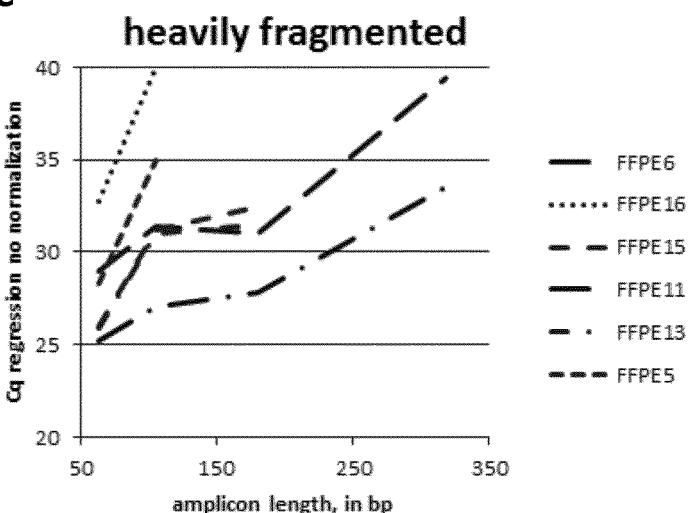
13. Automated method according to any of the claims 10-12, the method performed on a cartridge

engageable with the automated system.

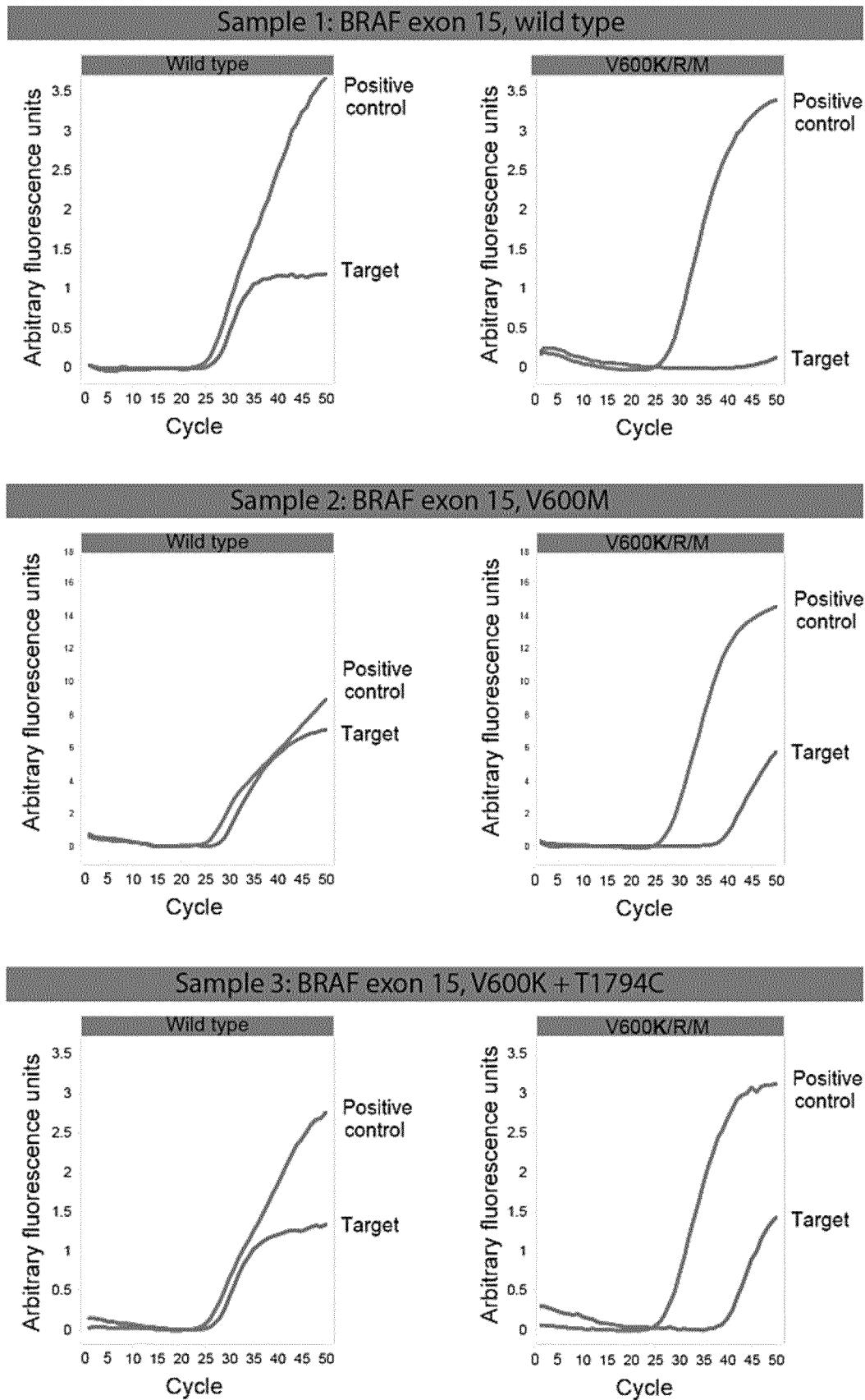
14. Automated method according to claim 13, wherein the cartridge comprises a cartridge-specific identifier for storing cartridge-specific information and wherein the method further comprises the step of introducing to the automated system the information stored in the cartridge-specific identifier.
- 5
15. Use of a cartridge according to claims 1-9 for automated preparation of a next-generation sequencing (NGS) library from a biological sample.

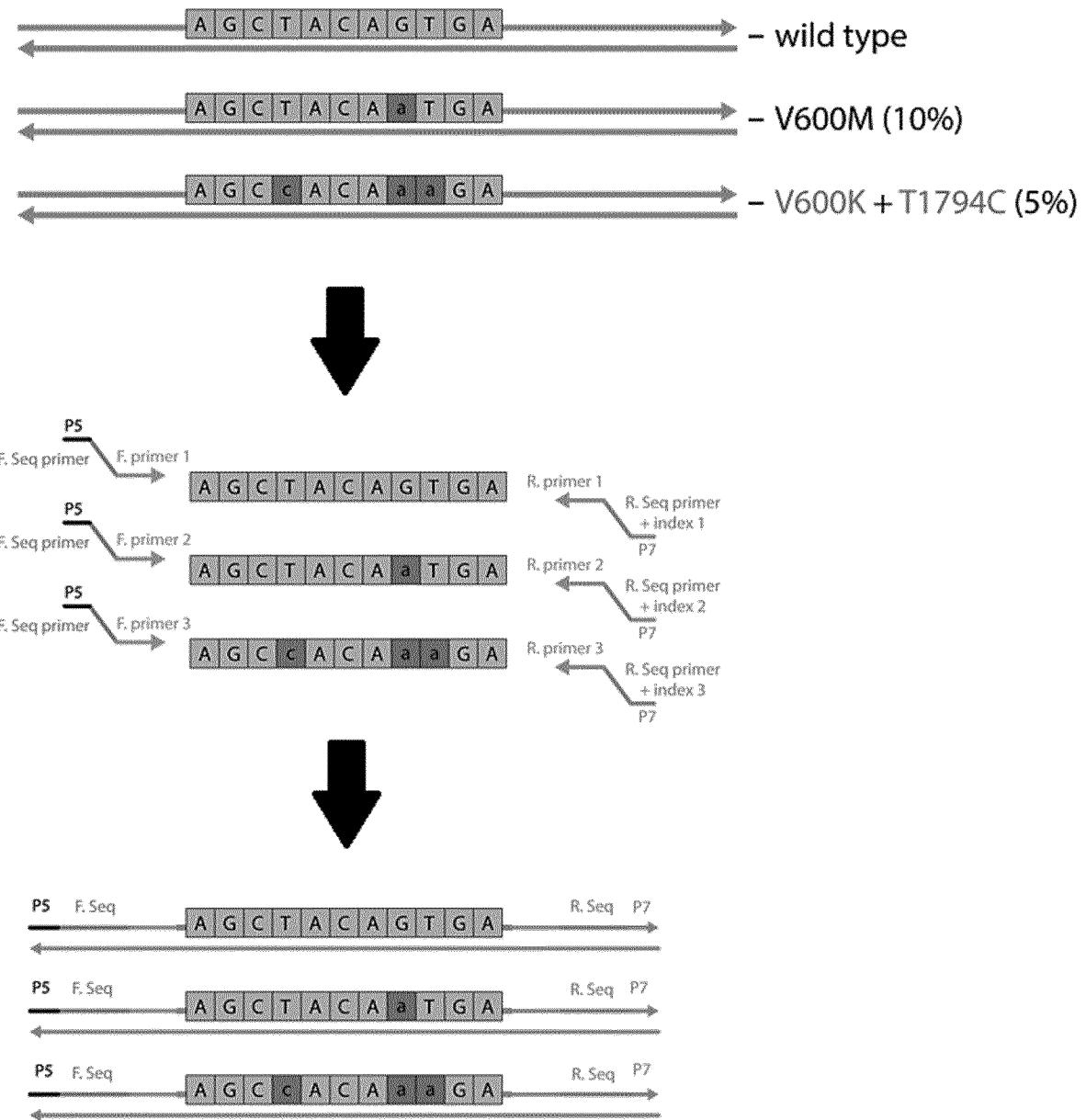
**Figure 1****Figure 2**

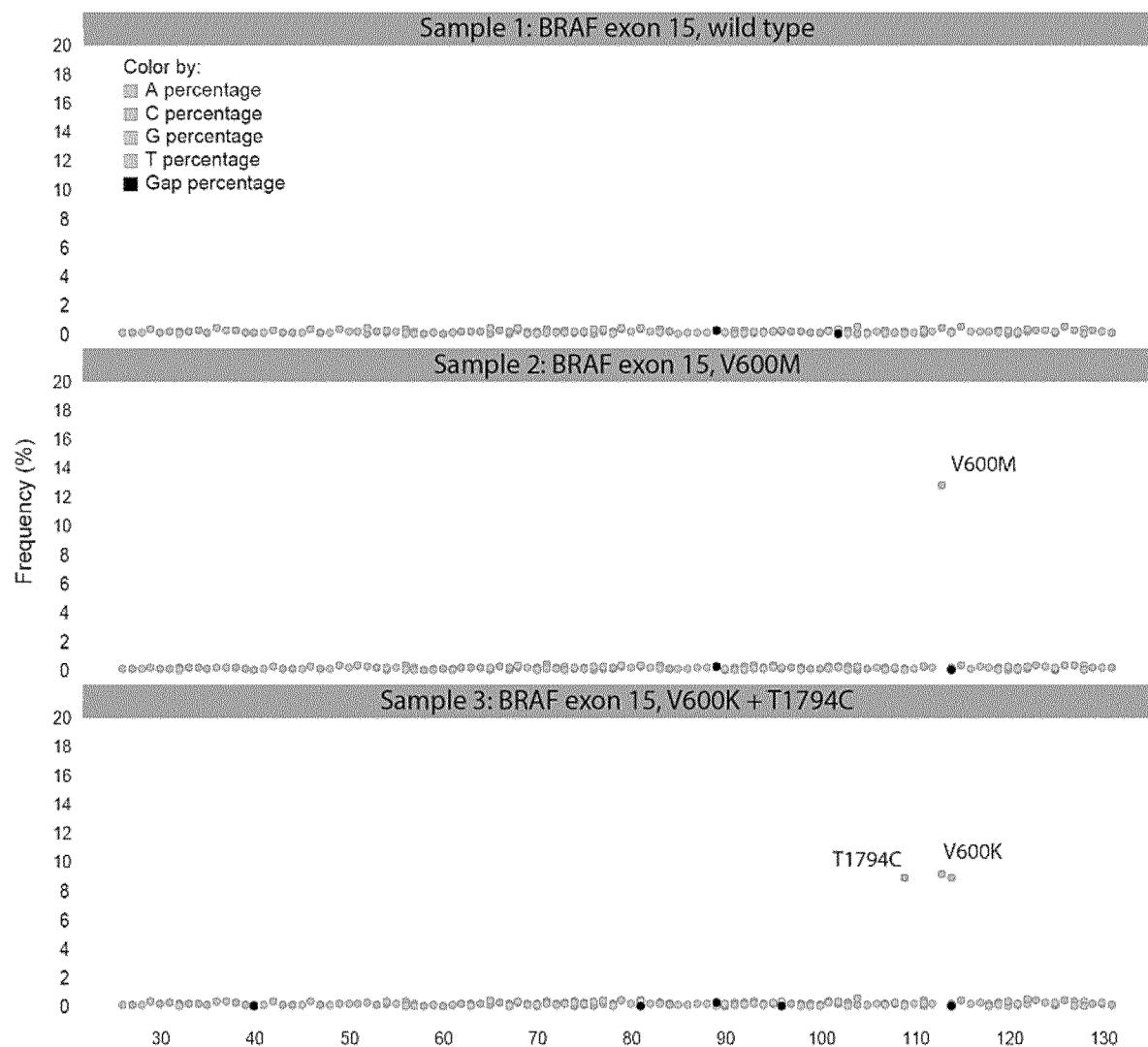
**Figure 3****Figure 4**

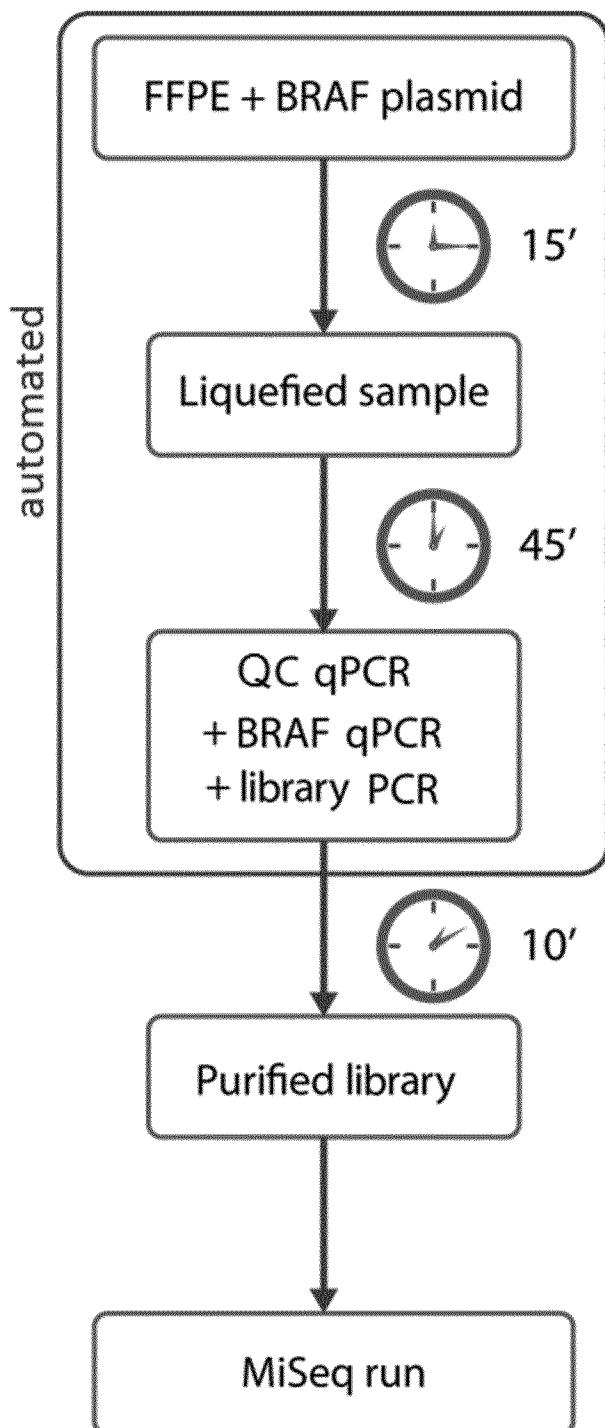
**Figure 5****A****B****C**

**Figure 6**



**Figure 7**

**Figure 8**

**Figure 9**

# INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2016/067149

**A. CLASSIFICATION OF SUBJECT MATTER**  
INV. B01L3/00  
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)  
B01L

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

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

EPO-Internal, BIOSIS, CHEM ABS Data, EMBASE, FSTA

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Anna-Iris Schiefer: "Evaluation of a novel fully automated PCR-based technique for detection of BRAF-status in FFPE-tissues", 1 January 2014 (2014-01-01), XP055229934, Retrieved from the Internet: URL: <a href="https://media.biocartis.com/biocartis/documents/141103_AMP-Poster_BRAF-Study_AT-Wien_AKH_Pathologie_Birner_FINALFINAL.pdf">https://media.biocartis.com/biocartis/documents/141103_AMP-Poster_BRAF-Study_AT-Wien_AKH_Pathologie_Birner_FINALFINAL.pdf</a> [retrieved on 2015-11-20] the whole document	1,10,15
Y	----- -/-	2-9, 11-14

Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 September 2016	17/10/2016
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  Leber, Thomas

## INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/067149

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	"Beckman Coulter SPRIworks Systems", , 1 January 2011 (2011-01-01), XP055229930, Retrieved from the Internet: URL: <a href="https://www.beckmancoulter.com/wsrporta1/bibliography?docname=BR-15981.pdf">https://www.beckmancoulter.com/wsrporta1/bibliography?docname=BR-15981.pdf</a> [retrieved on 2015-11-20] the whole document	1,10,15
Y	----- BUEHLER B ET AL: "Rapid quantification of DNA libraries for next-generation sequencing", METHODS, ACADEMIC PRESS, US, vol. 50, no. 4, 1 April 2010 (2010-04-01), pages S15-S18, XP026966692, ISSN: 1046-2023 [retrieved on 2010-03-07] items 1. and 5.	2-9, 11-14
Y	----- ALISON S. DEVONSHIRE ET AL: "Application of next generation qPCR and sequencing platforms to mRNA biomarker analysis", METHODS, vol. 59, no. 1, 1 January 2013 (2013-01-01), pages 89-100, XP055228123, US ISSN: 1046-2023, DOI: <a href="https://doi.org/10.1016/j.ymeth.2012.07.021">10.1016/j.ymeth.2012.07.021</a> items 4.1 and 4.1.5	2-9, 11-14
Y	----- "The SPRIworks Fragment Library System Simplifying Next Generation Sequencing Workflows SPRIworks Fragment Library System II Fully Automated Fragment Library System for the Roche GS FLX* DNA Sequencer", , 1 January 2010 (2010-01-01), XP055229937, Retrieved from the Internet: URL: <a href="https://www.beckmancoulter.com/wsrporta1/bibliography?docname=DS-14103.pdf">https://www.beckmancoulter.com/wsrporta1/bibliography?docname=DS-14103.pdf</a> [retrieved on 2015-11-20] page 1, drawing for "SPRIworks HT"	2-9, 11-14
A	----- WO 2012/012779 A2 (BECKMAN COULTER INC [US]; WILSON BRIAN D [US]; ALARURI SAMI D [US]; AN) 26 January 2012 (2012-01-26) page 53 "Assay Cartridge"; Fig. 4a-1 and 4a-2	1-15

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2016/067149

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2012012779	A2	26-01-2012	
		BR 112013001647 A2	24-05-2016
		CN 103119451 A	22-05-2013
		CN 103540517 A	29-01-2014
		CN 103540518 A	29-01-2014
		CN 103540530 A	29-01-2014
		CN 103543281 A	29-01-2014
		CN 103543282 A	29-01-2014
		CN 103551212 A	05-02-2014
		CN 103555558 A	05-02-2014
		CN 103555815 A	05-02-2014
		CN 103589627 A	19-02-2014
		CN 103657754 A	26-03-2014
		CN 103675303 A	26-03-2014
		CN 103725591 A	16-04-2014
		CN 104345161 A	11-02-2015
		CN 104345165 A	11-02-2015
		EP 2596135 A2	29-05-2013
		EP 2743704 A2	18-06-2014
		EP 2743705 A2	18-06-2014
		EP 2746777 A2	25-06-2014
		EP 2749356 A2	02-07-2014
		EP 2749887 A2	02-07-2014
		EP 2752245 A2	09-07-2014
		EP 2752246 A2	09-07-2014
		EP 2752247 A2	09-07-2014
		EP 2752668 A2	09-07-2014
		EP 2752670 A2	09-07-2014
		EP 2752671 A2	09-07-2014
		EP 2754496 A2	16-07-2014
		JP 5948271 B2	06-07-2016
		JP 2013150623 A	08-08-2013
		JP 2013535193 A	12-09-2013
		JP 2014193114 A	09-10-2014
		JP 2016036345 A	22-03-2016
		KR 20130029127 A	21-03-2013
		KR 20130029128 A	21-03-2013
		KR 20130041255 A	24-04-2013
		US 2013130369 A1	23-05-2013
		US 2013132006 A1	23-05-2013
		US 2013136670 A1	30-05-2013
		US 2013137087 A1	30-05-2013
		US 2013137109 A1	30-05-2013
		US 2013137110 A1	30-05-2013
		US 2013137164 A1	30-05-2013
		US 2013137166 A1	30-05-2013
		US 2013142711 A1	06-06-2013
		US 2013143199 A1	06-06-2013
		US 2013196422 A1	01-08-2013
		US 2013203046 A1	08-08-2013
		US 2013209334 A1	15-08-2013
		US 2015111288 A1	23-04-2015
		US 2015217291 A1	06-08-2015
		US 2015337400 A1	26-11-2015
		WO 2012012779 A2	26-01-2012