



(51) International Patent Classification:

CI2Q 1/6883 (2018.01)

(21) International Application Number:

PCT/GB2021/000101

(22) International Filing Date:

15 September 2021 (15.09.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

20200100574 22 September 2020 (22.09.2020) GR

(71) Applicant: **THE SECRETARY OF STATE FOR DEFENCE DSTL** [GB/GB]; Porton Down, Salisbury, Wiltshire, SP4 0JQ (GB).

(72) Inventors: **LUKASZEWSKI, Roman Antoni**; DSTL IP Group, G02 Building 005, DSTL Porton Down, Salisbury, Wiltshire SP4 0JQ (GB). **PANAGIOTOU, Ioannis**; Knaackstrasse 35, 10405, Berlin (DE). **SCHMIDT-HECK, Wolfgang**; Jenaprießnitzer Straße 8, 07749 Jena (DE).

(74) Agent: **FARNSWORTH, Alastair Graham**; DIPR Formalities Section, Poplar 2#2214, MOD Abbey Wood (South), Bristol BS34 8JH (GB).

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

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

Published:

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

(54) Title: APPARATUS, KITS AND METHODS FOR PREDICTING THE DEVELOPMENT OF SEPSIS

(57) Abstract: The present invention is concerned with kits, methods and apparatus for analysing a biological sample from a subject to predict and/or monitor the development of infection and/or organ dysfunction and/or sepsis utilising groups of nucleic acid markers to predict the development of infection and/or organ dysfunction and/or sepsis.



Apparatus, Kits and Methods for Predicting the Development of Sepsis

The present invention is concerned with kits, methods and apparatus for analysing a biological sample from a subject to predict and/or monitor the development of infection and/or organ dysfunction and/or sepsis utilising groups of nucleic acid markers to predict the development of infection and/or organ dysfunction and/or sepsis.

Following exposure to a microbial pathogen there is often a lag phase before symptoms of infection, which could further result in symptoms of organ dysfunction, and development of sepsis. After the onset of clinical symptoms, the effectiveness of treatment often decreases as the disease progresses, so the time taken to make any diagnosis is critical. It is likely that a detection or diagnostic assay will be the first confirmed indicator of infection, organ dysfunction or sepsis. The availability, rapidity and predictive accuracy of such an assay will therefore be crucial in determining the outcome. Any time saved will speed up the implementation of medical countermeasures and will have a significant impact on recovery.

The development of technologies to facilitate rapid detection of infection, organ dysfunction and sepsis is a key concern for all at risk. During the initial stages of infection many biological agents are either absent from, or present at very low concentrations in, typical clinical samples (e.g. blood). It is therefore likely that agent-specific assays would have limited utility in detecting infection before clinical symptoms arise. Previous studies have shown that infection elicits a pattern of immune response involving changes in the expression of a variety of biomarkers that is indicative of the type of agent. Such patterns of biomarker expression have proven to be diagnostic for a variety of infectious agents. It is now possible to distinguish patterns of gene expression in blood leukocytes from symptomatic patients with acute infections caused by four common human pathogens (Influenza A, *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Escherichia coli*) using whole transcriptome analysis. More recently, researchers have been able to reduce the number of host biomarkers

required to make a diagnosis through use of appropriate bioinformatic analysis techniques to select key biomarkers for the diagnosis of infectious disease.

While host biomarker signatures represent an attractive solution for the prediction of microbial infection, their discovery relies on the exploitation of laboratory models of infection whose fidelity to the pathogenesis of disease in humans varies. An alternative approach for biomarker discovery in humans is to exploit a common sequela of biological agent infection; such as the life-threatening condition sepsis, which now requires organ dysfunction for a positive diagnosis. Sepsis has traditionally been defined as a systemic inflammatory response syndrome (SIRS) in response to infection which, when associated with acute organ dysfunction, may ultimately cause severe life-threatening complications. However, sepsis is now defined as life-threatening organ dysfunction caused by a dysregulated host response to infection, wherein organ dysfunction can be identified as an increase in the total sequential organ failure assessment (SOFA) score of 2 or more points from one day to the next following infection. Severity of organ dysfunction has in the past been assessed with various scoring systems that quantify abnormalities according to clinical findings, laboratory data, or therapeutic interventions. The predominant score in current use is the SOFA (originally the Sepsis-related Organ Failure Assessment). A higher SOFA score is associated with an increased probability of mortality. The score grades abnormality by organ system and accounts for clinical interventions. The baseline SOFA score can be assumed to be zero in patients not known to have pre-existing organ dysfunction. A SOFA score ≥ 2 reflects an overall mortality risk of approximately 10% in a general hospital population with suspected infection.

Sepsis is a major cause of morbidity and mortality in intensive care units (ICU). In the UK, sepsis is believed to be responsible for about 27% of all ICU admissions. Across Europe the average incidence of sepsis in the ICU is about 30%, with a mortality rate of 27%. In the USA, hospital-associated mortality from sepsis ranges between 18 to 30%; an estimated 9.3% of all deaths occurred in patients with

sepsis. Clearly there is a very accessible patient population that could be used to study predictive markers for the onset of sepsis.

Despite greatly improved diagnosis, treatment and support, serious infection and sepsis remain significant causes of death and often result in chronic ill-health or disability in those who survive acute episodes. Although sudden, overwhelming infection is comparatively rare amongst otherwise healthy adults, it constitutes an increased risk in immunocompromised individuals, seriously ill patients in intensive care, burns patients and young children. In a proportion of cases, an apparently treatable infection leads to the development of sepsis; a dysregulated, inappropriate response to infection characterised by progressive circulatory collapse leading to renal and respiratory failure, abnormalities in coagulation, profound and unresponsive hypotension and, in about 30% of cases death. The incidence of sepsis in the population of North America is about 0.3% of the population annually (about 750,000 cases) with mortality rising to 40% in the elderly and to 50% in cases of the most severe form, septic shock.

The ability to detect potentially serious infections and organ dysfunction as early as possible and, especially, to predict the onset of sepsis in susceptible individuals is clearly advantageous.

Although a number of biomarkers (markers), such as nucleic acid markers or protein markers, have been shown to correlate with sepsis and some give an indication of the seriousness of the condition, no single marker or combination of markers has yet been shown to be a reliable diagnostic test, much less a predictor of the development of sepsis, especially sepsis meeting the criteria of the new definition, requiring life-threatening organ dysfunction.

Extracting reliable diagnostic patterns and robust prognostic indications from changes over time in complex sets of variables including traditional clinical observations, clinical chemistry, biochemical, immunological and cytometric data requires sophisticated methods of analysis. The use of expert

systems and artificial intelligence, including neural networks, for medical diagnostic applications has been being developed for some time.

The ability to detect the earliest signs of infection and/or organ dysfunction and/or sepsis has clear benefits in terms of allowing treatment as soon as possible. Indications of the severity of the condition and likely outcome if untreated inform decisions about treatment options. This is relevant both in vulnerable hospital populations, such as those in intensive care, or who are burned or immunocompromised, and in other groups in which there is an increased risk of serious infection and subsequent sepsis. The use or suspected use of biological weapons in both battlefield and civilian settings is an example where a rapid and reliable means of testing for the earliest signs of infection or organ dysfunction (i.e. sepsis) in individuals exposed would also be advantageous.

However, until now the majority of investigations focussed on developing a group of biomarkers and/or a test for sepsis were based on the previous definition of sepsis of systemic inflammatory response syndrome (SIRS) in response to infection, and thus have generally focussed on identifying a group of biomarkers and/or a test to predict SIRS in response to infection. Sepsis is, however, now defined as life-threatening organ dysfunction caused by a dysregulated host response to infection, and thus a group of biomarkers and/or a test is required which is capable of identifying individuals at risk of developing life-threatening organ dysfunction caused by a dysregulated host response to infection. Until now neither a test nor a list of biomarkers has been identified/produced which can detect or predict sepsis (organ dysfunction) with a good/high predictive accuracy (for example with an area under the curve (AUC) > 0.8, but preferably > 0.85, and most preferably > 0.9).

The present invention thus aims to provide biomarker signatures (groups of biomarkers), and methods for classifying biological samples using the biomarker signatures, to predict/detect the development of organ dysfunction, and consequently sepsis, and preferably both infection and organ dysfunction, and sepsis, with a high predictive accuracy.

With this in mind, the applicants have determined several lists/groups of nucleic acid markers (biomarkers) which can be used to predict the development of infection and organ dysfunction, and thus predict the development of sepsis, with the possibility of predicting prior to the onset of symptoms (pre-symptomatic). The Applicant has identified through a comprehensive analysis of the host transcriptome, sourced from blood samples from human patients collected prior to the clinical onset of sepsis, panels of 52 and 67 (which comprises the 52) nucleic acids highly significant to predicting sepsis, with an ability to predict both infection and organ dysfunction, and subsets thereof for predicting development of infection and/or organ dysfunction and/or sepsis.

The 67 nucleic acid markers highly significant to predicting infection and organ dysfunction (combining 52 markers identified through microarray analysis, and 15 key markers from a previous analysis) are:

AFF1, AGTPBP1, ATP9A, ATXN1, B3GNT5, B4GALT5, BIRC5, CAPN15, CD247, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDTP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DMA, HLA-DPA1, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LARP4B, LDHA, LDLR, LGALS2, LIMK2, LINC00999, LOC399744, MED13L, METTL7B, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPL13A, RPL18A, RPS13, RPS14, RRBP1, SGSH, SLC26A6, SLC36A1, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TRPM2, TXNDC3, ZKSCAN1.

The Applicant had in particular identified a list of 52 gene biomarkers from microarray analysis whose expression is in particular effected by infection and organ dysfunction, and consequently sepsis. The group of 52 nucleic acid markers from microarray analysis consists of AFF1, AGTPBP1, ATXN1, B3GNT5, B4GALT5, BIRC5, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDTP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LDLR, LIMK2, LOC399744, MED13L, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPS13, RPS14, RRBP1, SGSH, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TXNDC3, ZKSCAN1.

Details of the particular nucleic acids, from their annotated names (as provided throughout this application), can be found from human gene databases such as GeneCards®: The Human Gene Database (genecards.org), or the Human Gene Resources at the National Center for Biotechnology Information (NCBI) (ncbi.nlm.nih.gov/genome/guide/human/).

- 5 Organ dysfunction and infection are two key indicators of sepsis, and thus there is a need to identify markers which are capable of predicting development of either, though at least organ dysfunction, but preferably both, with good levels of predictivity, and as early as possible, and preferably before symptoms occur. The Applicant has down-selected lists of nucleic acid markers, as detailed above, critical to predicting sepsis with a high level of confidence in the prediction, which are capable of
- 10 providing a prediction up to at least three days in advance of symptoms.

- Subsets of both the groups of 67 and 52 nucleic acid markers, in particular subsets from as few as 4 biomarkers, have been shown to be capable of predicting development of infection versus absence of infection (and development of infection vs development of SIRS), and predicting development of organ dysfunction versus non-complicated infection and/or versus absence of organ dysfunction,
- 15 and thus predict sepsis versus absence of sepsis, through analysis of samples from subjects (for example patients) categorised as progressing to development of sepsis, or not, or as progressing to having an infection, or not.

- Thus in a first aspect the present invention provides a method for predicting the development of infection and/or organ dysfunction and/or sepsis in a subject, the method comprising determining
- 20 levels of at least four nucleic acid markers, or a product expressed by those nucleic acids, such as the corresponding proteins, in a biological sample taken from the subject, wherein the at least four nucleic acid markers are selected from the lists consisting of:

AFF1, AGTPBP1, ATP9A, ATXN1, B3GNT5, B4GALT5, BIRC5, CAPN15, CD247, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDPI, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3,

HLA-DMA, HLA-DPA1, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LARP4B, LDHA, LDLR, LGALS2, LIMK2, LINC00999, LOC399744, MED13L, METTL7B, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPL13A, RPL18A, RPS13, RPS14, RRBP1, SGSH, SLC26A6, SLC36A1, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TRPM2, TXNDC3, ZKSCAN1; or

5 AFF1, AGTPBP1, ATXN1, B3GNT5, B4GALT5, BIRC5, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LDLR, LIMK2, LOC399744, MED13L, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPS13, RPS14, RRBP1, SGSH, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TXNDC3, ZKSCAN1;

10 wherein the levels of the at least four nucleic acid markers, or the products expressed by those nucleic acids, are used to predict the development of infection and/or organ dysfunction and/or sepsis.

The levels of the nucleic acid markers, or the products expressed by those nucleic acids, are preferably collectively, and in combination, used to predict the development of infection and/or

15 organ dysfunction and/or sepsis, which could be through use of a mathematical model applied to the levels of the nucleic acid markers, or the products expressed by those nucleic acids, to provide the prediction.

The method of the first aspect may comprise use of a control in the method from which to establish the level (or expression level) of each nucleic acid, or a product thereof. The control may for
20 example be one or more housekeeping genes/nucleic acids whose expression is predictable or relatively static irrespective of whether infection and/or organ dysfunction may develop For example, the reference gene/nucleic acid may be FAM105B and/or RANBP3.

The term 'biological sample' includes, but not exclusively, blood, serum, plasma, urine, saliva, cerebrospinal fluid or any other form of material, preferably fluid-based or capable of being

converted into a fluid-like state (e.g. tissue which can be broken down or separated in a solution, such as a buffered solution), which can be extracted or collected from a patient.

For prediction of infection versus absence of infection the Applicant has in particular identified 28 markers (through microarray analysis), from the list of 52 nucleic acid markers which are significant to a prediction, with selections of at least 4, though optimally at least 7 or 8, markers from the 28 being particularly effective for providing a prediction of infection versus an absence of infection, with a high level of predictivity.

The group of 28 markers is: AFF1, AGTPBP1, B3GNT5, B4GALT5, BIRC5, CSGALNACT2, CTDTP1, DENND4B, EIF4G3, FBXW2, FKBP5, GYG1, HLA-DPB1, HVCN1, KLHL2, LDLR, LOC399744, PHOSPHO1, RPS13, RPS14, SGSH, SLC41A3, SOLH, SPATA13, STOM, SYCP2, TCEA3, ZKSCAN1.

For example, the following subsets from the list of 28 markers are capable of achieving an AUC > 0.9 for predicting development of infection on specific days prior to development of symptoms:

1. B4GALT5, BIRC5, CSGALNACT2, CTDTP1, RPS13, RPS14, SLC41A3, ZKSCAN1 (Day-1 before symptoms)
2. AFF1, AGTPBP1, DENND4B, FBXW2, GYG1, SGSH, SPATA13, TCEA3 (Day-2 before symptoms)
3. B3GNT5, CTDTP1, HVCN1, KLHL2, LOC399744, PHOSPHO1, SLC41A3, SYCP2 (Day-3 before symptoms)

These subsets of markers may be optimal for a specific day before the onset of symptoms of infection, or sepsis, although may be used in a test to identify development of infection or sepsis at any point before or after onset of symptoms, whereas other subsets of markers from the list of 28 markers may also be used for providing predictions, on individual days before symptoms or multiple days before symptoms, for example other subsets of markers from the list of 28 may be capable of predicting development of infection (or sepsis) with an AUC > 0.9 irrespective of the number of days prior to onset of symptoms, up to at least three days prior to onset of symptoms. For example, the

11 marker subset of B4GALT5, ZKSCAN1, RPS14, BIRC5, STOM, LDLR, LOC399744, SOLH, FKBP5, HLA-DPB1, EIF4G3.

Other biomarker signatures (groups of nucleic acid markers) for use in the present invention (identified through RT-PCR analysis), which can be selected from the 67 markers, and can predict infection with a high predictive accuracy are:

1. ATXN1, DOK3, HLA-DMA, LDLR, NFKBIA, RPS13, SLC26A6, SLC36A1
2. AFF1, ATP9A, B4GALT5, CAPN15, HLA-DPB1, LARP4B, METTL7B, TNFAIP3
3. AFF1, CAPN15, DENND4B, LDHA, QSOX1, RPS14, SLC36A1, TCEA3
4. B4GALT5, SLC36A1, LARP4B, KIF1B, SLC26A6, ATXN1, DENND4B, AFF1

10 Key nucleic acid markers, present in multiple (2 or more) down-selected subsets for prediction of infection versus no-infection are, from microarray analysis, B4GALT5, ZKSCAN1, RPS14 and BIRC5.

Key nucleic acid markers, present in multiple (2 or more) down-selected subsets for prediction of infection versus no-infection are, from PCR analysis, AFF1, ATXN1, SLC26A6, SLC36A1, B4GALT5, CAPN15, LARP4B, DENND4B,

15 The selection of markers thus may comprise one or more of the key nucleic acid markers that recur during down-selection of subsets of markers. A subset may for example comprise B4GALT5, AFF1, and DENND4B, markers that recur in numerous subsets, and may in addition, or substitution, comprise other recurring nucleic acid markers.

For prediction of organ dysfunction versus absence of organ dysfunction the Applicant has in particular identified 26 markers, and a smaller group of 14 markers, which are significant to a prediction, with selections of at least 4, or at least 6, from the 26 (or 14) being particularly effective for providing a prediction of organ dysfunction versus absence of organ dysfunction with a high level of predictivity.

The group of 26 markers is: ATXN1, CD6, CFLAR, CHD7, CHSY1, DOK3, FLT3, HCK, HIPK3, HLA-DOA, IL1R1, KIF1B, LIMK2, MED13L, MIDN, NFKBIA, QSOX1, RNF144B, RPL3L, RPS14, RRBP1, SOS2, STOM, TLR2, TNFAIP3, and TXNDC3.

The group of 14 markers is: CD6, CFLAR, CHD7, CHSY1, FLT3, HIPK3, IL1R1, KIF1B, LIMK2, MED13L, QSOX1, RNF144B, STOM, and TNFAIP3.

For example, the following subsets from the list of 14 markers are capable of achieving an AUC > 0.8 for predicting development of organ dysfunction on specific days prior to development of symptoms:

1. IL1R1, LIMK2, QSOX1, RNF144B, STOM, TNFAIP3 (Day-1 before symptoms)
2. CFLAR, CHD7, KIF1B, MED13L (Day-2 before symptoms)
3. CD6, CHSY1, FLT3, HIPK3 (Day-3 before symptoms)

These subsets of markers may be optimal for a specific day before the onset of symptoms of organ dysfunction (sepsis), although may be used in a test to identify development of infection or sepsis at any point before or after onset of symptoms, whereas other subsets of markers from the list of 14 or 26 markers may also be used for providing predictions, on individual days before symptoms or multiple days before symptoms, for example other subsets of markers from the list of 26 may be capable of predicting development of organ dysfunction, irrespective of the number of days prior to onset of symptoms, up to at least three days prior to onset of symptoms. For example, the 16 marker subset of ATXN1, CFLAR, DOK3, FLT3, HCK, HLA-DOA, MIDN, NFKBIA, RPL3L, RPS14, RRBP1, SOS2, STOM, TLR2, TNFAIP3, and TXNDC3.

Other biomarker signatures (groups of nucleic acid markers) for use in the present invention, identified through PCR analysis, which can be selected from the 67 markers, and can predict organ dysfunction with a high predictive accuracy are:

1. DOK3, HLA-DPA1, RPL13A, RPL18A, SLC41A3, TNFAIP3
2. BIRC5, DENND4B, FLT3, SLC26A6, SLC36A1
3. CD247, DENND4B, LGALS2, LINC00999, MED13L, RPL3L, TRPM2
4. RPL13A, FLT3, DENND4B, DOK3, TNFAIP3, LGALS2, MED13L, MIDN, RPL3L, LINC00999

- 5 Key nucleic acid markers, present in multiple (2 or more) down-selected subsets for prediction of organ dysfunction versus no organ dysfunction are, from microarray analysis, CFLAR, FLT3, STOM and TNFAIP3.

Key nucleic acid markers, present in multiple (2 or more) down-selected subsets for prediction of organ dysfunction versus no organ dysfunction are, from q-PCR analysis, DOK3, TNFAIP3, DENND4B,

- 10 FLT3, LGALS2, MED13L, RPL3L, LINC00999, RPL13A.

The selection of markers thus may comprise one or more of the key nucleic acid markers that recur during down-selection of subsets of markers. A subset may for example comprise FLT3, DENND4B, and TNFAIP3, markers that recur in numerous subsets.

The biomarker signatures (lists of nucleic acid markers) for predicting development of organ

- 15 dysfunction are all capable of achieving an AUC of at least 0.75, with the majority achieving an AUC of > 0.8 or > 0.85 , which is particularly advantageous for predicting organ dysfunction, which until now has been difficult to predict, with the majority of studies in the past focussing on predicting sepsis based on the previous definition of sepsis, i.e. SIRS in response to infection, thus infection versus absence of infection alone.

- 20 The biomarker signatures of the present invention are thus especially valuable as they are capable of providing a test that can predict both infection and organ dysfunction, either together or separately, as part of a test for predicting sepsis, with the added value of better informing treatment and monitoring treatment of a subject.

The at least 4 nucleic acid markers to be determined in the method of the first aspect of the invention may be selected from any of the subsets, or key nucleic acid markers, that have been identified or down-selected for differentiating development of infection from no-infection, development of infection from development or SIRS, or development of organ dysfunction from non-specific infection or no-organ dysfunction, and in any combination, from the list of 67 or 52 markers identified as highly significant to predicting sepsis.

In a second aspect, the present invention provides a method for monitoring a subject at risk of developing infection and/or organ dysfunction and/or sepsis, the method comprising determining levels of at least four nucleic acid markers, or the products expressed by those nucleic acids, in biological samples taken from the subject at multiple time points, wherein the monitored levels of the at least four markers are used to predict development of infection and/or organ dysfunction and/or sepsis, wherein the at least four nucleic acid markers are selected from the list consisting of:

AFF1, AGTPBP1, ATP9A, ATXN1, B3GNT5, B4GALT5, BIRC5, CAPN15, CD247, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1,, HCK, HIPK3, HLA-DMA, HLA-DPA1, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LARP4B, LDHA, LDLR, LGALS2, LIMK2, LINC00999, LOC399744, MED13L, METTL7B, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL13A, RPL18A, RPS13, RPS14, RRBP1, SGSH, SLC26A6, SLC36A1, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TCEA3, TLR2, TNFAIP3, TRPM2, TXNDC3, ZKSCAN1; or

AFF1, AGTPBP1, ATXN1, B3GNT5, B4GALT5, BIRC5, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LDLR, LIMK2, LOC399744, MED13L, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPS13, RPS14, RRBP1, SGSH, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TXNDC3, ZKSCAN1.

The biomarkers of the present invention, or subsets thereof, are particularly advantageous for use in a test for monitoring a subject at risk over several days, which could incorporate different marker sets to identify the most likely day before symptoms, such as Day-1, Day-2, or Day-3 before symptoms, and whether infection and/or organ dysfunction is likely, and thereby consequently inform the best course of treatment to prevent or treat infection/organ dysfunction. The monitoring may comprise comparing dynamic changes in quantitation or rates of change of the biomarkers to derive predictors, such as for whether and when a subject may develop sepsis. For example, monitoring may comprise interrogation of biomarker velocity, such as its rate of change over time. An AUC of 0.8 and above for predicting organ dysfunction is achievable with nucleic acid markers subsets from the list of 67 or 52 markers, which is particularly good for predicting organ dysfunction versus non-specific infection or absence of organ dysfunction, as no biomarker set to date has been identified that is particularly directed to organ dysfunction, since most studies on sepsis have used the previous definition of sepsis, which was concerned with infection alone, and not organ dysfunction, and especially not an approach that could evaluate both infection and organ dysfunction with a single set of biomarkers, or different groups of biomarkers for each of infection and organ dysfunction. Particular subsets of markers are as for the first aspect of the present invention.

In a third aspect, the present invention provides a kit for predicting development of infection and/or organ dysfunction and/or sepsis in a subject, said kit comprising reagents and/or systems for determining levels of at least four markers in a biological sample from the subject, wherein the at least four markers are selected from the list consisting of:

AFF1, AGTPBP1, ATP9A, ATXN1, B3GNT5, B4GALT5, BIRC5, CAPN15, CD247, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1,, HCK, HIPK3, HLA-DMA, HLA-DPA1, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LARP4B, LDHA, LDLR, LGALS2, LIMK2, LINC00999, LOC399744, MED13L, METTL7B, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL13A,

RPL18A, RPS13, RPS14, RRBP1, SGSH, SLC26A6, SLC36A1, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TCEA3, TLR2, TNFAIP3, TRPM2, TXNDC3, ZKSCAN1; or

AFF1, AGTPBP1, ATXN1, B3GNT5, B4GALT5, BIRC5, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LDLR, LIMK2, LOC399744, MED13L, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPS13, RPS14, RRBP1, SGSH, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TXNDC3, ZKSCAN1.

The subject is most likely a human, but may also be an animal, and the biological sample is most likely a blood or serum sample.

10 Particular subsets of markers are as for the first aspect of the present invention.

The kit of the invention may comprise means for detecting levels of a nucleic acid or nucleic acid product. Although nucleic acid expression may be determined by detecting the presence of nucleic acid products including proteins and peptides, such processes may be complex. In a particular embodiment, the means comprises means for detecting a nucleic acid, for example RNA, such as mRNA, or means for detecting a product expressed by the nucleic acid, such as a protein.

The reagents or systems may include use of recognition elements, or microarray based methods. Thus in a particular embodiment, the kit of the invention comprises a microarray on which are immobilised probes suitable for binding to RNA expressed by each nucleic acid of a biomarker signature. Means for detecting a protein may comprise an antibody, which may be a fluorescently-labelled antibody, and may comprise protein recognition elements on a microarray, or other suitable platform, such as lateral flow strips.

In an alternative embodiment, the kit comprises at least some of the reagents suitable for carrying out amplification of nucleic acids of the biomarker signature, or regions thereof.

In one embodiment the reagents or systems use real-time (RT) polymerase chain reaction (PCR). In such cases, the reagents may comprise primers for amplification of said nucleic acids or regions thereof. The kits may further comprise labels in particular fluorescent labels and/or oligonucleotide probes to allow the PCR to be monitored in real-time using any of the known assays, such as TaqMan, LUX, etc. The kits may also contain reagents such as buffers, enzymes, salts such as MgCl etc. required for carrying out a nucleic acid amplification reaction. The reagents, especially for nucleic acid amplification, may comprise for example one or more of fluorescently-labelled oligonucleotide probes or fluorescently-labelled primers, wherein the fluorescently-labelled oligonucleotide probes or fluorescently-labelled primers may consist of probes and primers each capable of specific binding and detection of nucleic acid products of the least 4 markers.

The methods of the first or second aspect may advantageously be computer-implemented to handle the complexity in monitoring and analysis of the numerous biomarkers, and their respective relationships to each other. Such a computer-implemented invention could enable a yes/no answer as to whether infection and/or organ dysfunction and/or sepsis is likely to develop, or at least provide an indication of how likely the development is.

The method preferably uses mathematical tools and/or algorithms to monitor and assess expression of the biomarkers (the nucleic acid markers, or products thereof) both qualitatively and quantitatively. The tools could in particular include support vector machine (SVM) algorithms, decision trees, random forests, artificial neural networks, quadratic discriminant analysis, and Bayes classifiers. In one embodiment the data from monitoring all biomarkers in the biomarker signature is assessed by means of an artificial neural network.

In one embodiment of the first or second aspect the method is a computer-implemented method wherein the monitoring, measuring and/or detecting comprises producing quantitative, and optionally qualitative, data for all markers, inputting said data into an analytical process on the computer, using at least one mathematical method, that may compare the data with reference data,

and producing an output from the analytical process which provides a prediction for the likelihood of developing infection and/or organ dysfunction and/or sepsis, or enables monitoring of the condition.

The reference data may include data from healthy subjects, subjects diagnosed with sepsis (organ dysfunction), subjects with infection, and subjects with SIRS, but no infection.

- 5 The output from the analytical process may enable the time to onset of symptoms to be predicted, such as 1, 2, or 3 days prior to onset of symptoms, and consequently may be particularly valuable and useful to a medical practitioner in suggesting a course of treatment, especially when the choice of course of treatment is dependent on the progression of the disease. The method may also enable monitoring of the success of any treatment, assessing whether the likelihood of onset of symptoms
- 10 decreases over the course of treatment.

In a fourth aspect, the present invention provides an apparatus for analysis of a biological sample from a subject to predict or monitor the development of sepsis comprising means for monitoring, measuring or detecting the expression of at least four markers in a biological sample from the subject, wherein the at least four markers are selected from the list consisting of:

- 15 AFF1, AGTPBP1, ATP9A, ATXN1, B3GNT5, B4GALT5, BIRC5, CAPN15, CD247, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTD1P1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1,, HCK, HIPK3, HLA-DMA, HLA-DPA1, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LARP4B, LDHA, LDLR, LGALS2, LIMK2, LINC00999, LOC399744, MED13L, METTL7B, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL13A, RPL18A, RPS13, RPS14, RRB1, SGSH, SLC26A6, SLC36A1, SLC41A3, SOLH, SOS2,
- 20 SPATA13, STOM, SYCP2, TCEA3, TCEA3, TLR2, TNFAIP3, TRPM2, TXNDC3, ZKSCAN1; or

AFF1, AGTPBP1, ATXN1, B3GNT5, B4GALT5, BIRC5, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTD1P1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LDLR, LIMK2, LOC399744, MED13L, MIDN, NFKBIA, PHOSPHO1, QSOX1,

RNF144B, RPL3L, RPS13, RPS14, RRBP1, SGSH, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TXNDC3, ZKSCAN1;

and means for analysis of data produced from the means for monitoring, measuring or detecting, such as a computer comprising an appropriate mathematical model to analyse the data, such as an artificial neural network, and means for providing an output from the analysis which output provides a prediction of the likelihood of a subject having sepsis, or an output to enable monitoring of infection and/or organ dysfunction and/or sepsis, which output could also be provided by an appropriately programmed computer.

All subsets of markers (biomarker signatures) of the different aspects of the present invention may be used to monitor and/or predict the response of a subject to a particular therapeutic agent, such as a sepsis targeting drug or an antibiotic. For example the expression of particular nucleic acid markers, which may for example be elevated in a subject on a course to develop sepsis, or having already developed sepsis, could be monitored to establish whether the levels are returning to the levels expected for a subject without, or unlikely to develop, sepsis, which could be an indication of the therapeutic agent successfully treating the subject. A therapeutic agent may be one targeted to particular subsets of markers in an attempt to treat the subject, and indeed the choice of therapeutic agent to be used in a subject may be determined by the expression of specific nucleic acid markers which are most affected, or differ most from a control or from a patient not predicted to develop sepsis, as a result of developing sepsis. For example, the elevation of certain markers may suggest the use of one therapeutic agent, whereas elevation of a different subset of markers, may suggest use of another therapeutic agent.

Any feature in one aspect of the invention may be applied to any other aspects of the invention, in any appropriate combination. In particular, method aspects may be applied to use, kit and system aspects and vice versa. The invention extends to methods, uses, kits or systems substantially as herein described, with reference to the Example(s).

In all aspects, the invention may comprise, consist essentially of, or consist of any feature or combination of features.

Example – Development of a predictive panel of pre-symptomatic biomarkers for infection, organ dysfunction, and thereby sepsis

5 The aim of this program of work was to develop a predictive panel of pre-symptomatic biomarkers for infection and organ dysfunction (sepsis), through comprehensive analysis of the host transcriptome, sourced from blood samples from human patients collected prior to the clinical onset of infection or organ dysfunction, and to develop biomarker signatures that may indicate whether and when clinical symptoms will arise. In so doing it would yield a suitably powered bioinformatic model
10 for differentiating sepsis patients from control patients based on transcriptomic biomarker signatures. In turn, this will assist in the development of RT-PCR methods for infection and organ dysfunction prediction, where this capability should provide timely diagnosis and treatment when medical countermeasures are most effective.

We used microarray technology to obtain gene expression data of samples derived from pre-
15 symptomatic patients and control patient samples. An unsupervised bioinformatic approach was used to identify prognostic transcriptomic expression patterns that characterize infection and organ dysfunction before the onset of clinical symptoms. Characteristic biomarker patterns were further identified, analysed and validated using quantitative RT-PCR on the Fluidigm BioMark™ real-time PCR array platform.

20 The Applicant conducted a large prospective, multicenter study in patients undergoing elective major surgery, with daily blood sampling and data recording commencing before the operation and continuing for up to a week after, to enable pre-symptomatic identification of patients developing infection complicated or not by new-onset organ dysfunction (sepsis). Crucially, there was a clinical adjudication panel who independently examined clinical and laboratory data to identify patients

with definite infection ($\pm$ sepsis). Samples from these patients enabled accurate comparison of microarray and RT-qPCR data against cohorts of age-, sex- and procedure-matched patients with non-infective systemic inflammation (SIRS) or an uncomplicated postoperative course. Nucleic acid expression signatures measured in blood samples could identify patients developing infection up to
5 three days prior to clinical presentation, and could differentiate between patients developing uncomplicated infection or sepsis (organ dysfunction).

Patient recruitment

Elective surgery patients were prospectively recruited into this study between November 2007 and February 2017. Patients were enrolled if they gave informed consent, were between 18-80 years of
10 age and undergoing an elective high-risk surgical procedure that placed them at an increased risk of infection $\pm$ sepsis. Recruitment occurred at seven centres in the UK and one in Germany.

Clinical and blood samples were collected from 4,385 patients undergoing high-risk elective surgery. The total number of sample vials received was 72,734, and subsequent sub-aliquoting of key samples for further analysis generated a further 81,800 vials. 155 patients were adjudicated by the
15 clinical advisory panel (CAP) to have definite post-operative infection.

Sampling

Blood sample collection occurred once between 1-7 days before surgery and then daily post-operatively until seven days, hospital discharge (if earlier), or diagnosis of infection or sepsis by the treating clinician.

20 Sepsis patient selection process

Initial diagnosis of infection or sepsis was based on the treating clinician's interpretation of clinical and laboratory markers using the then-extant 'Sepsis-2' definition of sepsis. This described '*sepsis*' as suspected or confirmed infection with two or more systemic inflammatory response syndrome (SIRS) criteria, and '*severe sepsis*' as sepsis plus new-onset organ dysfunction. The new sepsis

definition (Sepsis-3) published in February 2016 rebadged '*sepsis*' as infection plus new organ dysfunction identified by a rise of ≥ 2 points from one day to the next of the patient's SOFA score. To keep in line with modern nomenclature, subsequent analyses and descriptors apply the new definition.

- 5 As initial diagnosis of infection and sepsis is often based on clinical judgement and before any microbiological confirmation, the CAP was formed to adjudicate cases labelled as postoperative infection. A minimum of five specialists in intensive care or microbiology independently reviewed clinical, laboratory and imaging results to give a high confidence diagnosis of infection. These patients were allocated to either an infection, group a non-infective systemic inflammation (SIRS) or
- 10 an uncomplicated post-operative recovery group assuming no significant non-infective issues arose (e.g. hemorrhage, myocardial infarction). The infection group was subsequently divided into 2 sub-groups, those with or without organ dysfunction as defined by an increase of the patient's SOFA score by 2 or more from one day to the next.

Microarray analysis

- 15 Microarray analysis performed on blood samples taken up to 3 days (Days -1, -2 and -3) before clinical diagnosis of infection identified 52 transcripts (nucleic acids) that were taken through to qRT-PCR studies, with 15 additional transcripts (nucleic acids) from a previous analysis. Comparison was made against qRT-PCR measured on the same genes in age-, sex- and procedure-matched patients who either had non-infective systemic inflammation or an uncomplicated recovery.
- 20 For each sample analysed, globin-reduced RNA (GlobinClear™, ThermoFisher) was prepared from total RNA. RNA integrity was measured using an Agilent Bioanalyzer 2100 (Town, State) and concentration using a NanoQuant™ (Tecan, Town). cRNA was prepared by amplification and labeling using the Illumina® TotalPrep™ RNA Amplification Kit (ThermoFisher) and hybridized to Human HT-12v4 Beadarrays (Illumina®, Place, State). An Illumina® HighScanHQ™ then imaged each chip with
- 25 resulting intensities indicating the expression level of each probe's corresponding gene.

The Illumina® Human HT-12v4 beadarrays were preprocessed and background corrected using GenomeStudio™ Software v2011.1 (Illumina®). To obtain genes with the greatest evidence of differential expression, a linear model fit was applied for each gene using the limma package (Doi: 10.1093/nar/gkv007: Ritchie, M.E. et al, Nucleic Acids Res., 2015, 43(7)). Datasets include patient group with or without infection up to three days before diagnosis of infection. Data obtained from non-infected patients were used as a reference. A false positive rate of 0.05 with FDR correction and a fold change greater than 1.3 was taken as the level of significance.

qPCR Methods and Analysis

RNA Isolation

- 10 Total RNAs were extracted from an aliquot of 1,832 samples containing 400 µl whole blood immersed in 1 mL RNeasy lysis buffer using a twostep protocol as implemented in the RiboPure™-Blood Kit (ThermoFisher Scientific). RNeasy lysis buffer was removed and the cells were re-suspended in 800 µl guanidinium-based lysis solution. After addition of 50 µl sodium acetate solution, RNA was extracted with 500 µl acid phenol/chloroform. After phase separation, the aqueous phase was recovered and
- 15 600 µl 100% ethanol added for binding of the RNA on a glass fiber filter spin column. After three wash steps the RNA was eluted in 2x 50 µl elution solution.

The purity and the concentration of the resulting RNA was determined with spectrophotometry on a Nanodrop™ 8000 and the RNA integrity was analyzed on an Agilent Bioanalyzer. The RNA samples were then normalized to 15 ng/µl using 96 well UF plates (Qiagen, Hilden, Germany).

20 Quantitative Gene-Expression

For cDNA synthesis, 100 ng RNA was used as input in a 20 µl reaction using the High Capacity cDNA Kit (ThermoFisher Scientific) and the following incubation program; 25°C 10 min, 37°C 120 min, 85°C 5 min. For the standard curve, 1000 ng reference RNA was used as input in a 20 µl reaction. Negative controls were Non-Enzyme Control (NEC) prepared using 100 ng RNA without reverse transcriptase

and Non-Template Control (NTC) prepared from a reverse transcription with RNase free water instead of RNA.

Specific Target Amplification (STA) reactions were performed in accordance with the Fluidigm Specific Target Amplification Quick Reference [PN 69000133 RevB]. In short, 1.25 µl cDNA was mixed with 2.5 µl TaqMan PreAmp Master Mix (2x) [PN 4391128] and 1.25 µl Pooled assay mix (0.2x) and amplified in a thermal cycler under the following conditions; 95°C 10 min pre-incubation followed by 14 cycles of 95°C 15 sec, 60°C 4 min.

Gene Expression Analysis was measured by multiplex real time PCR on a BioMark HD system using 96.96 Dynamic Arrays (Fluidigm, San Francisco, CA) and TaqMan Gene Expression Assays (Applied Biosystems, Carlsbad, CA) with three technical replicates. The Taqman assays for FAM105B and RANBP3 were used to normalize the relative abundance of transcripts between samples. The optimal set of reference genes was determined using geNorm analysis. The analysis showed that optimal normalization was achieved using the two genes FAM105B and RANBP3 (geNorm V < 0.15 when comparing a normalization factor based on the two or three most stable targets).

Feature selection and learning of predictive models

To improve the performance metric of the predictive models a two-step feature selection was performed. First, the Boruta algorithm, a wrapper method based on Random forest was used for selection of relevant features in the data set. Then a new randomized feature (shadow feature) was added for each feature in the dataset. The classifier was then trained with the dataset and the importance of each feature calculated. Real features that have a significantly higher z-score than the best shadow feature are called relevant features. The Boruta algorithm was applied in a 5-fold cross-validation, repeated 25 times. A feature identified as relevant in at least one model was considered for further evaluation. As a second step, backward elimination was used to determine those features with the most discriminatory power for a particular classification problem. Starting with all relevant features in a 5 fold cross validation repeated 25 times the importance of features was calculated.

This loop was reiterated until a maximum of the assessment index area under the curve (AUC) was found. In each iteration step, the feature with the least importance was removed.

Results

Microarray Gene Expression Analysis of Infected Patients

- 5 Transcriptomic sequencing was carried out on samples from patients taken over the three days preceding clinical presentation of post-operative infection and compared to matched healthy postoperative controls.

Classification of development of infection based on microarray expression and RT-qPCR data

Microarray

- 10 A random forest-based algorithm was used to classify differential gene expression on Days -1, -2, or -3 prior to infection diagnosis against their respective non-infected controls. Random forest reports the most important genes to reach performance next to statistical metrics.

The best identified classification for Day -1 (based on 63 infection plus and 66 infection minus samples) reached an Area under the ROC Curve (AUC) of 0.957 and a positive predictive value (PPV)

- 15 of 0.901 for a set of eight genes (B4GALT5, BIRC5, CSGALNACT2, CTD1P1, RPS13, RPS14, SLC41A3, ZKSCAN1). The best identified classification for Day -2 (based on 48 infection plus and 51 infection minus samples) achieved an AUC of 0.932 (PPV 0.834) using a different set of eight genes from the overall set (AFF1, AGTPBP1, DENND4B, FBXW2, GYG1, SGSH, SPATA13, TCEA3), and the best classification for Day -3 (based on 40 infection plus and 40 infection minus samples) achieved an AUC
- 20 of 0.904 (PPV 0.892) using again eight different genes (B3GNT5, CTD1P1, HVCN1, KLHL2, LOC399744, PHOSPHO1, SLC41A3, SYCP2).

Classification of infection was repeated with a requirement for the same set of genes used by the random forest classifier for each day (Days -1, -2 and -3) prior to infection diagnosis. The best

performing classifiers by random forest required 11 gene features and achieved AUC values of 0.937 for Day -1 (PPV 0.898), 0.922 for Day -2 (PPV 0.866), and 0.900 for Day -3 (PPV 0.847). The 11 markers were B4GALT5, ZKSCAN1, RPS14, BIRC5, STOM, LDLR, LOC399744, SOLH, FKBP5, HLA-DPB1, and EIF4G3. The identified 11 gene set from the random forest classification of infection versus no infection was validated using RT-qPCR data.

RT-qPCR

Again, random forest based classification yielded excellent performance. Using individual gene sets for identifying infection on individual days prior to infection diagnosis (compared to a control, non-infected, non-inflamed group) generated AUC values of 0.930 (PPV 0.884) for Day -1 (based on 59 infection plus and 59 infection minus samples; ATXN1, DOK3, HLA-DMA, LDLR, NFKBIA, RPS13, SLC26A6, SLC36A1), 0.943 (PPV 0.882) for Day -2 (based on 45 infection plus and 45 infection minus samples; AFF1, ATP9A, B4GALT5, CAPN15, HLA-DPB1, LARP4B, METTL7B, TNFAIP3) and 0.925 (PPV 0.858) for Day -3 (based on 35 infection plus and 35 infection minus samples; AFF1, CAPN15, DENND4B, LDHA, QSOX1, RPS14, SLC36A1, TCEA3) based on biomarker sets all having 8 nucleic acid markers. With the requirement for the same gene set being used for each day of infection, AUC values of 0.894 (PPV 0.850) for Day -1, 0.901 for Day -2 (PPV 0.874) and 0.868 for Day -3 (PPV 0.828) were achieved with an eight gene set: B4GALT5, SLC36A1, LARP4B, KIF1B, SCLC26A6, ATXN1, DENND4B, AFF1.

Classification of development of organ dysfunction (sepsis) versus non-complicated infection based on microarray expression and RT-qPCR data

Microarray

A further random forest-based model set was generated using the microarray data for classifying infected patients either with development of organ dysfunction (sepsis) or without organ dysfunction (non-complicated infection). The best classification for Day -1 (based on 28 organ

dysfunction plus and 35 non-complicated infection samples) achieved an AUC of 0.891 (PPV 0.854) based on six genes (IL1R1, LIMK2, QSOX1, RNF144B, STOM, TNFAIP3). The best classification for Day -2 (based on 20 organ dysfunction plus and 28 non-complicated infection samples) was achieved with four genes (CFLAR, CHD7, KIF1B, MED13L), yielding an AUC of 0.844 (PPV 0.830). The classification performance for Day -3 (based on 15 organ dysfunction plus and 25 non-complicated infection samples) yielded an AUC of 0.830 (PPV 0.861) based on four genes (CD6, CHSY1, FLT3, HIPK3).

Following the same procedure as for development of infection, another classification model set based on random forest was created that required the same gene set for each separate day. AUC values of 0.753 (PPV 0.733) for Day -1, 0.756 (PPV 0.739) for Day -2 and 0.758 (PPV 0.733) were achieved with sixteen genes (*ATXN1, CFLAR, DOK3, FLT3, HCK, HLA-DOA, MIDN, NFKBIA, RPL3L, RPS14, RRBP1, SOS2, STOM, TLR2, TNFAIP3, and TXNDC3*).

RT-qPCR

Finally, RT-qPCR based expression was used for classification of sepsis versus uncomplicated infection. Individual gene sets per day prior to infection diagnosis yielded AUC values of 0.829 (PPV 0.837) with 6 genes (based on 27 organ dysfunction plus and 32 non-complicated infection samples; DOK3, HLA-DPA1, RPL13A, RPL18A, SLC41A3, TNFAIP3), 0.844 (PPV 0.830) with 5 genes (based on 18 organ dysfunction plus and 17 non-complicated infection samples; BIRC5, DENND4B, FLT3, SLC26A6, SLC36A1) and 0.902 (PPV 0.915) with 7 genes (based on 11 organ dysfunction plus and 24 non-complicated infection samples; CD247, DENND4B, LGALS2, LINC00999, MED13L, RPL3L, TRPM2) for Days -1, -2 and -3, respectively. Requiring a common gene set for each day prior to infection diagnosis yielded AUC values of 0.778 (PPV 0.756) for Day -1, 0.713 (PPV 0.725) for Day -2 and 0.776 (PPV 0.838) for Day-3 based on a ten gene set (RPL13A, FLT3, DENND4B, DOK3, TNFAIP3, LGALS2, MED13L, MIDN, RPL3L, LINC00999).

Overall 67 nucleic acid markers were identified in the study as being relevant and highly significant to the prediction of development of infection, or development of organ dysfunction. The 67 nucleic acid markers are:

AFF1, AGTPBP1, ATP9A, ATXN1, B3GNT5, B4GALT5, BIRC5, CAPN15, CD247, CD6, CFLAR, CHD7,
 5 CHSY1, CSGALNACT2, CTDP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1,, HCK, HIPK3,
 HLA-DMA, HLA-DPA1, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LARP4B, LDHA, LDLR,
 LGALS2, LIMK2, LINC00999, LOC399744, MED13L, METTL7B, MIDN, NFKBIA, PHOSPHO1, QSOX1,
 RNF144B, RPL13A, RPL18A, RPS13, RPS14, RRB1, SGSH, SLC26A6, SLC36A1, SLC41A3, SOLH, SOS2,
 SPATA13, STOM, SYCP2, TCEA3, TCEA3, TLR2, TNFAIP3, TRPM2, TXNDC3, ZKSCAN1.

- 10 Fifty two (52) of these nucleic acid markers were identified through microarray analysis as being a key nucleic acid marker for predicting infection and/or organ dysfunction:

AFF1, AGTPBP1, ATXN1, B3GNT5, B4GALT5, BIRC5, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDP1,
 DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DOA, HLA-DPB1, HVCN1,
 IL1R1, KIF1B, KLHL2, LDLR, LIMK2, LOC399744, MED13L, MIDN, NFKBIA, PHOSPHO1, QSOX1,
 15 RNF144B, RPL3L, RPS13, RPS14, RRB1, SGSH, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3,
 TLR2, TNFAIP3, TXNDC3, ZKSCAN1.

Of these 52 nucleic acid markers, 28 markers are particularly relevant for predicting infection, as compared to controls:

AFF1, AGTPBP1, B3GNT5, B4GALT5, BIRC5, CSGALNACT2, CTDP1, DENND4B, EIF4G3, FBXW2, FKBP5,
 20 GYG1, HLA-DPB1, HVCN1, KLHL2, LDLR, LOC399744, PHOSPHO1, RPS13, RPS14, SGSH, SLC41A3,
 SOLH, SPATA13, STOM, SYCP2, TCEA3, ZKSCAN1.

Of these 52 nucleic acid markers, 26 markers, with a smaller subset of 14 markers, are particularly relevant for predicting organ dysfunction, as compared to patients developing infection alone.

The group of 26 markers is: ATXN1, CD6, CFLAR, CHD7, CHSY1, DOK3, FLT3, HCK, HIPK3, HLA-DOA, IL1R1, KIF1B, LIMK2, MED13L, MIDN, NFKBIA, QSOX1, RNF144B, RPL3L, RPS14, RRBP1, SOS2, STOM, TLR2, TNFAIP3, and TXNDC3.

The group of 14 markers is: CD6, CFLAR, CHD7, CHSY1, FLT3, HIPK3, IL1R1, KIF1B, LIMK2, MED13L, QSOX1, RNF144B, STOM, and TNFAIP3.

The Applicant has interrogated this nucleic acid marker data, in light of specifically effective subsets for predicting infection or organ dysfunction, and concluded that subsets of 4 or 5 nucleic acid markers (and potentially less) from the 67, 52, 28, 26, or 14, should be capable of predicting infection or organ dysfunction with high AUCs, or accuracy.

Markers that occur in multiple exemplified subsets are more likely to provide further effective down-selected subsets, and thus the Applicant suggests recurring nucleic acid markers should be combined to provide further subsets for predicting infection and/or organ dysfunction. For example the nucleic acid markers B4GALT5, AFF1, and DENND4B occur in multiple down-selected lists for predicting infection, whereas the nucleic acid markers FLT3, DENND4B, and TNFAIP3 occur in multiple down-selected lists for predicting organ dysfunction, and thus the at least 4 nucleic acid markers may comprise one or more of these markers.

This study is the culmination of over 10 years of research into the pathogenesis of sepsis in elective surgery patients. Analysis of the host transcriptome in whole blood samples before and after surgery has led to the identification of a number of key host biomarkers whose expression enables differentiation of sepsis patients from other cohorts. Whilst whole transcriptome studies of sepsis have been widely reported, all have focussed on patients with established symptoms i.e. patients diagnosed with sepsis or infection before sampling began. However, these have limited utility for the early diagnosis of infection, or pre-symptomatic prediction of infection or organ dysfunction. In

addition, the complicated nature of whole transcriptomic data has limited its clinical utility for a number of practical and interpretational reasons.

The work reported here is unique since it has sought to address these two issues. Firstly, it has described the early host response that leads to sepsis through characterisation of the transcriptome of patients that go on to develop sepsis. Secondly, it has down-selected genes capable of discriminating between infection, organ dysfunction (sepsis) and other patient cohorts, and proved that it is possible to identify clinically useful host biomarker signatures (with low numbers of markers, less than twenty, and often less than ten required to provide a prediction) to predict infection and organ dysfunction in advance of symptoms.

In order to understand the host response that leads to sepsis, a unique approach to patient recruitment was adopted. Clarity on the provenance of each patient was considered as important as the clinical data itself for robust modelling of early sepsis pathogenesis. Patients were comparatively well when they entered the study. Apart from the underlying need for surgery, patients were infection free and were in relatively good health. This was underlined by the high rates of uncomplicated recovery observed in the study. However, 3.53% of the patients recruited into this study did develop organ dysfunction (sepsis). The prospective collection of samples before and after surgery enabled the detailed characterisation of changes in gene expression that led to the development of infection or sepsis in this elective surgery cohort. The low incidence of sepsis also gave a large patient cohort from which age/sex/procedure matched patients could be selected for inter-patient comparison. This enabled the effects of age, gender and surgical procedure to be controlled effectively.

Successful down-selection of genes to manageable numbers is vital for transition to a platform. Consequently, a machine learning algorithm approach has been used to select appropriate targets and classify patients based on host gene expression with output compared to clinical diagnosis to determine predictive accuracy. The success of this approach is evidenced by high AUC values and

small biomarker signatures (subsets of nucleic acid markers) of between 4 or 5 markers, up to about 12 markers or more (though less than twenty), when comparing infection and comparator, infection and SIRS and sepsis (organ dysfunction⁺) with infection only patients. These biomarker signatures are significantly much smaller than previously reported signatures for pre-symptomatic prediction of infection, and especially for pre-symptomatic prediction of organ dysfunction.

In trying to answer the more difficult question of when an individual will develop sepsis, this study was able to utilise multiple sample time points collected from each patient.

Claims

1. A method for predicting the development of infection and/or organ dysfunction and/or sepsis in a subject, the method comprising determining levels of at least four nucleic acid markers in a biological sample taken from the subject, wherein the at least four nucleic acid markers are selected from the lists consisting of:

AFF1, AGTPBP1, ATP9A, ATXN1, B3GNT5, B4GALT5, BIRC5, CAPN15, CD247, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DMA, HLA-DPA1, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LARP4B, LDHA, LDLR, LGALS2, LIMK2, LINC00999, LOC399744, MED13L, METTL7B, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPL13A, RPL18A, RPS13, RPS14, RRBP1, SGSH, SLC26A6, SLC36A1, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TRPM2, TXNDC3, ZKSCAN1; or

AFF1, AGTPBP1, ATXN1, B3GNT5, B4GALT5, BIRC5, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LDLR, LIMK2, LOC399744, MED13L, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPS13, RPS14, RRBP1, SGSH, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TXNDC3, ZKSCAN1;

wherein the levels of the at least four markers are used to predict the development of infection and/or organ dysfunction and/or sepsis.

2. Method according to Claim 1 wherein the at least 4 nucleic acid markers are selected from the lists consisting of:

AFF1, AGTPBP1, B3GNT5, B4GALT5, BIRC5, CSGALNACT2, CTDP1, DENND4B, EIF4G3, FBXW2, FKBP5, GYG1, HLA-DPB1, HVCN1, KLHL2, LDLR, LOC399744, PHOSPHO1, RPS13, RPS14, SGSH, SLC41A3, SOLH, SPATA13, STOM, SYCP2, TCEA3, ZKSCAN1, or

ATXN1, CD6, CFLAR, CHD7, CHSY1, DOK3, FLT3, HCK, HIPK3, HLA-DOA, IL1R1, KIF1B, LIMK2,
 MED13L, MIDN, NFKBIA, QSOX1, RNF144B, RPL3L, RPS14, RRBP1, SOS2, STOM, TLR2,
 TNFAIP3, TXNDC3, or

CD6, CFLAR, CHD7, CHSY1, FLT3, HIPK3, IL1R1, KIF1B, LIMK2, MED13L, QSOX1, RNF144B,
 STOM, TNFAIP3.

3. Method according to Claim 1 for predicting the development of infection, wherein the at least 4 nucleic acid markers are selected from the list consisting of AFF1, AGTPBP1, B3GNT5, B4GALT5, BIRC5, CSGALNACT2, CTDPI, DENND4B, EIF4G3, FBXW2, FKBP5, GYG1, HLA-DPB1, HVCN1, KLHL2, LDLR, LOC399744, PHOSPHO1, RPS13, RPS14, SGSH, SLC41A3, SOLH, SPATA13, STOM, SYCP2, TCEA3, ZKSCAN1, and wherein the levels of the at least four markers are used to predict the development of infection.

4. Method according to Claim 3, wherein the method comprises determining levels of at least 7 nucleic acid markers in a biological sample taken from a subject, and wherein the levels of the at least 7 markers are used to predict the development of infection.

5. Method according to Claim 1 for predicting the development of organ dysfunction and/or sepsis, wherein the at least 4 nucleic acid markers are selected from the lists consisting of:

ATXN1, CD6, CFLAR, CHD7, CHSY1, DOK3, FLT3, HCK, HIPK3, HLA-DOA, IL1R1, KIF1B, LIMK2,
 MED13L, MIDN, NFKBIA, QSOX1, RNF144B, RPL3L, RPS14, RRBP1, SOS2, STOM, TLR2,
 TNFAIP3, TXNDC3, or

CD6, CFLAR, CHD7, CHSY1, FLT3, HIPK3, IL1R1, KIF1B, LIMK2, MED13L, QSOX1, RNF144B,
 STOM, TNFAIP3.

and wherein the levels of the at least 4 nucleic acid markers are used to predict the development of organ dysfunction and/or sepsis.

6. Method according to Claim 1 or Claim 3 wherein the at least 4 nucleic acid markers comprises B4GALT5, AFF1, and DENND4B.

5 7. Method according to Claim 1 of Claim 5, wherein the at least 4 nucleic acid markers comprises FLT3, DENND4B, and TNFAIP3.

8. Method according to Claim 1 or Claim 3, wherein the nucleic acid markers determined is one of the following lists of nucleic acid markers:

- i. B4GALT5, BIRC5, CSGALNACT2, CTDP1, RPS13, RPS14, SLC41A3, ZKSCAN1;
- 10 ii. AFF1, AGTPBP1, DENND4B, FBXW2, GYG1, SGSH, SPATA13, TCEA3;
- iii. B3GNT5, CTDP1, HVCN1, KLHL2, LOC399744, PHOSPHO1, SLC41A3, SYCP2;
- iv. B4GALT5, ZKSCAN1, RPS14, BIRC5, STOM, LDLR, LOC399744, SOLH, FKBP5, HLA-DPB1, EIF4G3;
- v. ATXN1, DOK3, HLA-DMA, LDLR, NFKBIA, RPS13, SLC26A6, SLC36A1;
- 15 vi. AFF1, ATP9A, B4GALT5, CAPN15, HLA-DPB1, LARP4B, METTL7B, TNFAIP3;
- vii. AFF1, CAPN15, DENND4B, LDHA, QSOX1, RPS14, SLC36A1, TCEA3;
- viii. B4GALT5, SLC36A1, LARP4B, KIF1B, SCLC26A6, ATXN1, DENND4B, AFF1.

9. Method according to Claim 1 or Claim 5 wherein the nucleic acid markers determined is one of the following lists of nucleic acid markers:

- 20 i. IL1R1, LIMK2, QSOX1, RNF144B, STOM, TNFAIP3;
- ii. CFLAR, CHD7, KIF1B, MED13L;
- iii. CD6, CHSY1, FLT3, HIPK3;
- iv. ATXN1, CFLAR, DOK3, FLT3, HCK, HLA-DOA, MIDN, NFKBIA, RPL3L, RPS14, RRBP1, SOS2, STOM, TLR2, TNFAIP3, TXNDC3;

- v. DOK3, HLA-DPA1, RPL13A, RPL18A, SLC41A3, TNFAIP3;
 - vi. BIRC5, DENND4B, FLT3, SLC26A6, SLC36A1;
 - vii. CD247, DENND4B, LGALS2, LINC00999, MED13L, RPL3L, TRPM2;
 - viii. RPL13A, FLT3, DENND4B, DOK3, TNFAIP3, LGALS2, MED13L, MIDN, RPL3L,
- 5 LINC00999.

10. A method for monitoring a subject at risk of developing infection and/or organ dysfunction and/or sepsis, the method comprising determining levels of at least four nucleic acid markers in biological samples taken from a subject at multiple time points, wherein the monitored levels of the at least four markers are used to predict development of infection and/or organ dysfunction and/or sepsis., wherein the at least four nucleic acid markers are selected from the lists consisting of:

AFF1, AGTPBP1, ATP9A, ATXN1, B3GNT5, B4GALT5, BIRC5, CAPN15, CD247, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTD1P1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DMA, HLA-DPA1, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LARP4B, LDHA, LDLR, LGALS2, LIMK2, LINC00999, LOC399744, MED13L, METTL7B, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPL13A, RPL18A, RPS13, RPS14, RRB1, SGSH, SLC26A6, SLC36A1, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TRPM2, TXNDC3, ZKSCAN1; or

AFF1, AGTPBP1, ATXN1, B3GNT5, B4GALT5, BIRC5, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTD1P1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LDLR, LIMK2, LOC399744, MED13L, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPS13, RPS14, RRB1, SGSH, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TXNDC3, ZKSCAN1.

11. A kit for predicting development of infection and/or organ dysfunction and/or sepsis in a subject, said kit comprising reagents and/or systems for determining levels of at least four

markers in a biological sample from the subject, wherein the at least four markers are selected from the lists consisting of:

AFF1, AGTPBP1, ATP9A, ATXN1, B3GNT5, B4GALT5, BIRC5, CAPN15, CD247, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTD1P1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DMA, HLA-DPA1, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LARP4B, LDHA, LDLR, LGALS2, LIMK2, LINC00999, LOC399744, MED13L, METTL7B, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPL13A, RPL18A, RPS13, RPS14, RRB1, SGSH, SLC26A6, SLC36A1, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TRPM2, TXNDC3, ZKSCAN1; or

AFF1, AGTPBP1, ATXN1, B3GNT5, B4GALT5, BIRC5, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTD1P1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LDLR, LIMK2, LOC399744, MED13L, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPS13, RPS14, RRB1, SGSH, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TXNDC3, ZKSCAN1.

12. An apparatus for analysis of a biological sample from a subject to predict or monitor the development of sepsis comprising means for monitoring, measuring or detecting the expression of at least four markers in a biological sample from the subject, wherein the at least four markers are selected from the lists consisting of:

AFF1, AGTPBP1, ATP9A, ATXN1, B3GNT5, B4GALT5, BIRC5, CAPN15, CD247, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTD1P1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DMA, HLA-DPA1, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LARP4B, LDHA, LDLR, LGALS2, LIMK2, LINC00999, LOC399744, MED13L, METTL7B, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPL13A, RPL18A, RPS13, RPS14, RRB1, SGSH, SLC26A6, SLC36A1, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TRPM2, TXNDC3, ZKSCAN1; or

AFF1, AGTPBP1, ATXN1, B3GNT5, B4GALT5, BIRC5, CD6, CFLAR, CHD7, CHSY1, CSGALNACT2, CTDP1, DENND4B, DOK3, EIF4G3, FBXW2, FKBP5, FLT3, GYG1, HCK, HIPK3, HLA-DOA, HLA-DPB1, HVCN1, IL1R1, KIF1B, KLHL2, LDLR, LIMK2, LOC399744, MED13L, MIDN, NFKBIA, PHOSPHO1, QSOX1, RNF144B, RPL3L, RPS13, RPS14, RRBP1, SGSH, SLC41A3, SOLH, SOS2, SPATA13, STOM, SYCP2, TCEA3, TLR2, TNFAIP3, TXNDC3, ZKSCAN1,

and means for analysis of data produced from the means for monitoring, measuring or detecting, such as a computer comprising an appropriate mathematical model to analyse the data, and means for providing an output from the analysis which output provides a prediction of the likelihood of a subject having sepsis, or an output to enable monitoring of infection and/or organ dysfunction and/or sepsis, which output could also be provided by an appropriately programmed computer.

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2021/000101

A. CLASSIFICATION OF SUBJECT MATTER**INV. C12Q1/6883****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)

C12Q

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

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

EPO-Internal, BIOSIS, EMBASE, 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	WO 2015/121605 A1 (SECR DEFENCE [GB]) 20 August 2015 (2015-08-20) abstract tables 1, 5, 20-21, 25 claims 1-2, 12, 20 -----	1, 6-12
X	WO 2006/113529 A2 (BECTON DICKINSON CO [US]; GARRETT JAMES A [US] ET AL.) 26 October 2006 (2006-10-26) abstract tables 1, 31-34 paragraphs [0010], [0103], [0106], [0148] claims 73, 81-82 ----- -/-	1-11



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

10 December 2021

Date of mailing of the international search report

04/01/2022

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

Barz, Wolfgang

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2021/000101

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GAVIN C. K. W. KOH ET AL: "Host Responses to Melioidosis and Tuberculosis Are Both Dominated by Interferon-Mediated Signaling", PLOS ONE, vol. 8, no. 1, 1 January 2013 (2013-01-01) , page e54961, XP055376146, DOI: 10.1371/journal.pone.0054961 abstract figure 4 discussion -----	1-11
X	Illumina: "HumanHT-12 v4 BeadChip Product Information", , 1 January 2010 (2010-01-01), pages 1-2, XP055178072, Retrieved from the Internet: URL:http://www.illumina.com/documents/products/product_information_sheets/product_info_humanht-12.pdf [retrieved on 2015-03-20] the whole document -----	11
X	Ambion: "Illumina TotalPrep TM RNA Amplification Kit", , 1 January 2011 (2011-01-01), pages 1-30, XP055178075, Retrieved from the Internet: URL:http://tools.lifetechnologies.com/content/sfs/manuals/IL1791ME.pdf [retrieved on 2015-03-20] the whole document -----	11

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2021/000101

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2015121605 A1	20-08-2015	AU 2015216742 A1	18-08-2016
		CA 2938761 A1	20-08-2015
		CN 106103744 A	09-11-2016
		EP 3105342 A1	21-12-2016
		GB 2527164 A	16-12-2015
		GB 2559070 A	25-07-2018
		JP 6681337 B2	15-04-2020
		JP 2017506510 A	09-03-2017
		US 2017009297 A1	12-01-2017
		US 2020172978 A1	04-06-2020
		WO 2015121605 A1	20-08-2015
WO 2006113529 A2	26-10-2006	AU 2006236588 A1	26-10-2006
		BR PI0609302 A2	11-10-2011
		CA 2605143 A1	26-10-2006
		EP 1869463 A2	26-12-2007
		JP 2008538007 A	02-10-2008
		KR 20080006617 A	16-01-2008
		US 2006246495 A1	02-11-2006
		US 2011105350 A1	05-05-2011
		US 2014141435 A1	22-05-2014
		US 2016168638 A1	16-06-2016
		US 2021087632 A1	25-03-2021
		WO 2006113529 A2	26-10-2006